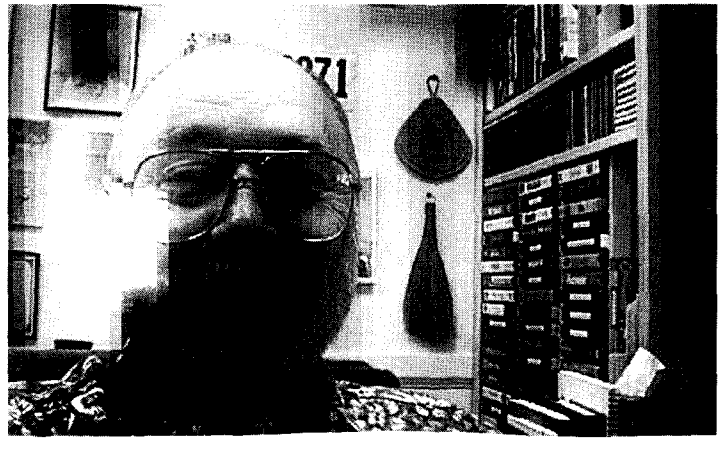
Od autora

Pisanie to moja radość, socjologia to moja pasja. Zachwyca mnie składanie słów tak, by skłaniały ludzi do nauki lub śmiechu – lub jednego i drugiego zarazem. Socjologia jest także zbiorem słów. Jest ona symbolem ostatniej nadziei na globalne wychowanie naszego gatunku i odnalezienie sposobów na wspólne życie. Fakt, że mogę być świadkiem tego, jak socjologia wchodzi nareszcie w centrum zainteresowania jako idea, której czas już nadszedł, jest dla mnie szczególnie podniecający.

Wyrosłem w małym miasteczku Vermont w New Hampshire. Kiedy oznajmiłem, że chcę zostać mechanikiem samochodowym jak mój tata, nauczyciel powiedział mi, żebym lepiej poszedł na studia. Kiedy Malcolm X oznajmił, że chce zostać prawnikiem, jego nauczyciel powiedział, że kolorowy chłopiec powinien być raczej kimś w rodzaju stolarza. Różnica w naszych doświadczeniach mówi coś istotnego o koncepcji poziomu odniesienia tła społecznego. Nierówności między grupami etnicznymi sięgają bardzo głęboko.

Wydostałem się na świat zewnętrzny poprzez Harvard, amerykańską piechotę morską, Uniwersytet w Berkeley i dwanaście lat nauczania na Uniwersytecie Hawajskim. W międzyczasie ożeniłem się z Sheilą (dwa miesiące po naszej pierwszej randce) i trzy lata później stworzyliśmy Aarona – dwie najrozsądniejsze rzeczy, jakie kiedykolwiek zrobiłem. Zrezygnowałem z pracy nauczyciela w 1980 roku i pracowałem przez siedem lat jako pełnoetatowy pisarz, póki zew sali wykładowej nie stał się zbyt głośny, by go zignorować. Uczenie jest dla mnie jak jazz; nawet kiedy grasz ten sam numer wciąż od nowa, nigdy nie wychodzi ci dokładnie to samo i nigdy nie wiesz dokładnie, jak będzie brzmiało – póki nie usłyszysz. Uczyć – znaczy pisać głosem.

Dojrzałem ostatecznie na tyle, że każdego lata odkrywam na nowo i doceniam moje korzenie w Vermont. Odbieram to nie jako powrót w przeszłość, ale raczej jako kolejny zakręt na drodze wzdłuż rozszerzającej się spirali. Nie mogę się doczekać, by zobaczyć, co kryje się za następnym.



Badania społeczne w praktyce

EARL BABBIE

Przełożyli

WITOLD BETKIEWICZ
MARTA BUCHOLC
PRZEMYSŁAW GADOMSKI
JACEK HAMAN
AGNIESZKA JASIEWICZ-BETKIEWICZ
AGNIESZKA KLOSKOWSKA-DUDZIŃSKA
MICHAŁ KOWALSKI
MAJA MOZGA

Redakcji naukowej dokonała oraz przekład przejrzała
AGNIESZKA KLOSKOWSKA-DUDZIŃSKA

WYDAWNICTWO NAUKOWE PWN
WARSZAWA 2004

Dane o oryginale:

Earl Babbie *The Practice of Social Research*

9th Edition. Copyright © 2001 Wadsworth/Thomson Learning. All Rights Reserved
ISBN 0-534-57474-2

Zespół tłumaczy:

Witold Betkiewicz (rozdz. 10, 11)
Marta Bucholc (rozdz. 5, 6, 14–16, Od autora, Wprowadzenie, Podziękowania, Przegląd holograficzny, Załączniki A–H)
Przemysław Gadomski (rozdz. 4, Załącznik I)
Jacek Haman (rozdz. 17–19, Załącznik J, Słowniczek)
Agnieszka Jasiewicz-Betkiewicz (rozdz. 8)
Agnieszka Kloskowska-Dudzińska (rozdz. 7, 9)
Michał Kowalski (rozdz. 12, 13)
Maja Mozga (rozdz. 1–3)

Pomoc w tłumaczeniu – **Krzysztof Jasiewicz**

Współpraca przy redakcji naukowej (rozdz. 14–19, Załączniki A–H) – **Marta Bucholc**

Konsultacja naukowa – **Antoni Sułek**

Współpraca przy opracowaniu Załącznika I – **SPSS Polska**

Projekt okładki i stron tytułowych – **Maryna Wiśniewska**

Redaktor – **Agnieszka Betlejewska**

Redaktor techniczny – **Teresa Skrzypkowska**

258935

Copyright © for the Polish edition by Wydawnictwo Naukowe PWN SA
Warszawa 2003

ISBN 83-01-14068-2

Wydawnictwo Naukowe PWN SA
00-251 Warszawa, ul. Miodowa 10
tel.: (0-22) 695-43-21; faks: (0-22) 826-71-63
e-mail:pwn@pwn.com.pl; http://www.pwn.pl

Spis treści

Wprowadzenie 3
Podziękowania 9
Przegląd holograficzny 11

Część 1

Wprowadzenie do badań 25

1 Działalność badawcza człowieka a nauka 27

Wprowadzenie 28
W poszukiwaniu rzeczywistości 29
Zwykła działalność badawcza człowieka 29
Tradycja 30
Autorytet 31
Błędy w badaniach oraz jak ich uniknąć 31
Co rzeczywiście jest rzeczywiste? 33
Podstawy nauk społecznych 36
Teoria zamiast filozofii czy wiary 36
Prawidłowości społeczne 37
Zbiorowości zamiast jednostek 39
Język zmiennych 40
Trochę dialektyki badań społecznych 45
Wyjaśnianie idiograficzne i nomotetyczne 45
Teoria indukcyjna i dedukcyjna 46
Dane jakościowe i ilościowe 48
Etyka w badaniach społecznych 50
Dobrowolne uczestnictwo 50
Nie krzywdzić badanych 50

Główne zagadnienia 51
Kluczowe pojęcia 52
Pytania i ćwiczenia 52
Dodatkowe lektury 52
Socjologiczna strona w sieci 52
Infotrac College Edition 53

2 Paradygmaty, teoria i badania społeczne 54

Wprowadzenie 55
Kilka paradygmatów nauk społecznych 56
Makroteoria i mikroteoria 57
Wczesny pozytywizm 57
Darwinizm społeczny 58
Paradygmat konfliktu 59
Symboliczny interakcjonizm 60
Etnometodologia 60
Funkcjonalizm strukturalny 61
Paradygmaty feministyczne 62
Racjonalna obiektywność raz jeszcze 63
Elementy teorii społecznej 66
Powrót do dwóch systemów logicznych 67
Tradycyjny model nauki 67
Rozumowanie dedukcyjne i indukcyjne – zilustrowane na przykładzie 70
Kontrast ujęty graficznie 73
Budowa teorii dedukcyjnej 75
Zaczynamy 75
Tworzenie własnej teorii 75
Przykład teorii dedukcyjnej: sprawiedliwość dystrybutywna 76
Budowa teorii indukcyjnej 78
Przykład teorii indukcyjnej: dlaczego ludzie palą marihuanę? 78
Związki między teorią a badaniami 79

Główne zagadnienia 81
Kluczowe pojęcia 82
Pytania i ćwiczenia 82
Dodatkowe lektury 82
Socjologiczna strona w sieci 83
Infotrac College Edition 83

3 Pojęcie przyczynowości w badaniach społecznych 84

Wprowadzenie 85
Determinizm a nauki społeczne 85

Przyczynowość w naukach przyrodniczych 85
Znajdowanie przyczyn w naukach społecznych 85
Nie ma powodu bez powodu 86
Determinizm we właściwych proporcjach 87
Przyczynowość w idiograficznych i nomotetycznych modelach wyjaśniania 88
Kryteria przyczynowości 91
Przyczyny konieczne i wystarczające 96
Błędy w rozumowaniu na temat związków przyczynowych 97
Prowincjonalizm 98
Pochopne wnioski 99
Wątpliwa przyczyna 99
Przemilczane dowody 99
Fałszywy dylemat 99
Połączenie pomiaru z ustalaniem związków 100
Praktyczne problemy przy pomiarze zmiennych i ustalaniu związków między nimi 100
Przykład pomiaru i związku między zmiennymi 102

Główne zagadnienia 103
Kluczowe pojęcia 103
Pytania i ćwiczenia 103
Dodatkowe lektury 104
Socjologiczna strona w sieci 104
Infotrac College Edition 105

Część 2

Struktura procesu badawczego 107

4 Plan badań 109
Wprowadzenie 110
Trzy cele badań 110
Eksploracja 111
Opis 112
Wyjaśnianie 113
Jednostki analizy 113
Jednostki indywidualne 115
Grupy 115
Organizacje 118
Wytwory społeczne 118
Przegląd jednostek analizy 119
Błędne wnioskowanie o jednostkach analizy: błąd ekologiczny oraz redukcjonizm 119
Wymiar czasowy 121
Badania przekrojowe 121
Badania dynamiczne 122
Badania zbliżone do dynamicznych 125
Przykłady strategii badawczych 127
Jak zaplanować projekt badawczy 127
Zaczynamy 129
Konceptualizacja 130
Wybór metody badawczej 130
Operacjonalizacja 131
Populacja i dobór próby 131
Obserwacje 132
Przetwarzanie danych 132
Analiza 132
Zastosowanie 132
Przegląd procesu planowania badań 133
Oferta badawcza 134
Elementy oferty badawczej 134

Główne zagadnienia 135
Kluczowe pojęcia 136
Pytania i ćwiczenia 136
Dodatkowe lektury 137
Socjologiczna strona w sieci 138
Infotrac College Edition 138

5 Konceptualizacja, operacjonalizacja i pomiar 139
Wprowadzenie 140
Mierzenie wszystkiego, co istnieje 140
Koncepcje, pojęcia i rzeczywistość 141
Pojęcia jako konstrukty 142
Konceptualizacja 144
Wskaźniki i wymiary 144
Wzajemna wymienialność wskaźników 145
Definicje realne, nominalne i operacyjne 146
Tworzenie ładu pojęciowego 147
Przykład konceptualizacji: pojęcie anomii 148
Definicje w badaniach opisowych i wyjaśniających 151
Wybory dokonywane w procesie operacjonalizacji 153
Zakres zmienności 153
Zmienność między skrajnościami 154
Uwaga na temat wymiarów 155
Definiowanie zmiennych i wartości 155
Poziomy pomiaru 156
Implikacje poziomów pomiaru 159
Wskaźniki proste i złożone 160
Kilka przykładów możliwych wyborów w procesie operacjonalizacji 161
Operacjonalizacja trwa 162
Kryteria jakości pomiaru 162
Precyzja i poprawność 162
Rzetelność 163
Trafność 166
Kto decyduje, co jest trafne? 167
Napięcie między rzetelnością a trafnością 168

Główne zagadnienia 169
Kluczowe pojęcia 169

Pytania i ćwiczenia 170
Dodatkowe lektury 170
Socjologiczna strona w sieci 170
Infotrac College Edition 171

6 Indeksy, skale i typologie 172

Wprowadzenie 173
Indeksy a skale 173
Budowa indeksu 176
Dobór pytań 176
Badanie zależności empirycznych 177
Punktacja indeksu 183
Braki danych 184
Walidacja indeksu 186
Status kobiety: przykład procesu tworzenia indeksu 189
Budowa skali 190
Skala dystansu społecznego Bogardusa 190
Skale Thurstone'a 191
Skala Likerta 191
Dyferencjał semantyczny 192
Skala Guttmana 193
Typologie 196

Główne zagadnienia 197
Kluczowe pojęcia 198
Pytania i ćwiczenia 198
Dodatkowe lektury 199
Socjologiczna strona w sieci 199
Infotrac College Edition 199

7 Logika doboru próby 200

Wprowadzenie 201
Krótka historia doboru próby 202
Prezydent Alf Landon 202
Prezydent Thomas E. Dewey 203
Dwa typy doboru próby 204
Próby nieprobabilistyczne 204
Dobór oparty na dostępności badanych 204
Dobór celowy lub arbitralny 205
Metoda kuli śnieżnej 205
Dobór kwotowy 206
Wybór informatorów 207
Teoria i logika doboru losowego 208
Świadome i nieświadome obciążenie próby 209
Reprezentatywność i prawdopodobieństwo wyboru 210
Dobór losowy 212
Rachunek prawdopodobieństwa, rozkład z prób i szacowanie błędu z próby 212
Populacje i operaty losowania 220
Przegląd populacji i operatów losowania 223
Typy doboru próby 223
Prosty dobór losowy 224
Dobór systematyczny 225
Dobór warstwowy 227
Ukryty podział na warstwy w doborze systematycznym 229
Przykład: dobór próby studentów uniwersytetu 229
Dobór próby wielostopniowy grupowy 230
Schematy doboru wielostopniowego a błąd z próby 231
Warstwowanie w wielostopniowej próbie grupowej 234
Dobór proporcjonalny 234
Dobór nieproporcjonalny i ważenie 236
Przykład: dobór próby kobiet chodzących do kościoła 237
Podsumowanie probabilistycznego doboru próby 239

Główne zagadnienia 239
Kluczowe pojęcia 240
Pytania i ćwiczenia 241
Dodatkowe lektury 241
Socjologiczna strona w sieci 241
Infotrac College Edition 241

Część 3

***Typy obserwacji* 243**

8 Eksperyment 245

Wprowadzenie 246
Tematy odpowiednie dla eksperymentu 246
Eksperyment klasyczny 247
Zmienne niezależne i zależne 247
Pretest i posttest 247
Grupy eksperymentalne i kontrolne 248
Eksperyment podwójnie ślepy 249
Dobór uczestników 250
Dobór losowy 250
Randomizacja 251
Dobieranie grup przez dopasowanie 251
Dobieranie grup przez dopasowanie czy randomizacja? 252
Odmiany schematów eksperymentalnych 253
Preeksperymentalne schematy badawcze 253
Trafność w badaniach eksperymentalnych 255
Przykłady eksperymentów 259
Eksperymenty „naturalne" 262
Mocne i słabe strony metody eksperymentalnej 264

Główne zagadnienia 265
Kluczowe pojęcia 265

Pytania i ćwiczenia 265
Dodatkowe lektury 266
Socjologiczna strona w sieci 266
Infotrac College Edition 266

9 Badania sondażowe 267

Wprowadzenie 268
Tematy odpowiednie dla badań sondażowych 268
Wskazówki do zadawania pytań 270
Wybierz właściwą formę pytania 270
Spraw, by pytania były jasne 271
Unikaj podwójnych pytań 271
Respondenci muszą być w stanie udzielić odpowiedzi 273
Respondenci muszą chcieć odpowiadać 273
Pytania powinny być adekwatne 274
Najlepsze są krótkie pytania 274
Unikaj form przeczących 274
Unikaj obciążonych pytań i pojęć 275
Budowa kwestionariusza 276
Ogólny format kwestionariusza 276
Formaty odpowiedzi 276
Pytania warunkowe 277
Pytania tabelaryczne 279
Kolejność pytań w kwestionariuszu 279
Instrukcje do kwestionariusza 280
Pilotaż kwestionariusza 281
Całościowy przykład 282
Kwestionariusz do samodzielnego wypełniania 282
Dystrybucja pocztowa i zwroty 285
Monitorowanie zwrotów 286
Listy przypominające 286
Wystarczający poziom zwrotów 287
Opis przypadku 288
Kwestionariusz wypełniany przez ankietera 290
Rola ankietera w badaniach sondażowych 290
Ogólne zasady pracy ankieterskiej 291
Koordynacja i kontrola 293
Sondaże telefoniczne 295
Wywiad telefoniczny wspomagany komputerowo (CATI) 297
Nowe technologie a badania sondażowe 298
Porównanie różnych metod badań sondażowych 300
Mocne i słabe strony badań sondażowych 301
Analiza wtórna 303

Główne zagadnienia 304
Kluczowe pojęcia 305
Pytania i ćwiczenia 305
Dodatkowe lektury 306
Socjologiczna strona w sieci 307
Infotrac College Edition 307

10 Jakościowe badania terenowe 308

Wprowadzenie 309
Tematy odpowiednie dla badań terenowych 309
Rzeczy, które należy brać pod uwagę w jakościowych badaniach terenowych 312
Różne role obserwatora 312
Stosunek do badanych 314
Niektóre paradygmaty jakościowych badań terenowych 316
Naturalizm 316
Etnometodologia 317
Teoria ugruntowana 318
Studium przypadku i metoda rozszerzonych przypadków 320
Etnografia instytucjonalna 322
Badania uczestniczące 323
Prowadzenie jakościowych badań terenowych 325
Przygotowanie do pracy w terenie 325
Wywiady jakościowe 327
Grupy fokusowe 330
Rejestrowanie obserwacji 331
Mocne i słabe strony jakościowych badań terenowych 334
Trafność 334
Rzetelność 335
Etyka badawcza w jakościowych badaniach terenowych 336

Główne zagadnienia 336
Kluczowe pojęcia 337
Pytania i ćwiczenia 337
Dodatkowe lektury 338
Socjologiczna strona w sieci 339
Infotrac College Edition 339

11 Badania niereaktywne 340

Wprowadzenie 341
Analiza treści 342
Tematy odpowiednie dla analizy treści 342
Dobór próby w analizie treści 343
Kodowanie w analizie treści 347
Przykład analizy treści 351
Mocne i słabe strony analizy treści 352
Analiza istniejących danych statystycznych 353
Badania nad samobójstwami Durkheima 353
Jednostki analizy 355
Problem trafności 356
Problem rzetelności 356
Źródła istniejących danych statystycznych 358
Analizy historyczno-porównawcze 360
Przykłady analiz historyczno-porównawczych 360
Źródła danych historyczno-porównawczych 363
Techniki analityczne 364

Główne zagadnienia 367
Kluczowe pojęcia 367
Pytania i ćwiczenia 367
Dodatkowe lektury 368
Socjologiczna strona w sieci 368
Infotrac College Edition 368

12 Badania ewaluacyjne 369

Wprowadzenie 370
Tematy odpowiednie dla badań ewaluacyjnych 371
Sformułowanie problemu: zagadnienia związane z pomiarem 372
Określenie wyników 373
Mierzenie tła eksperymentu 374
Określenie interwencji 375
Określenie populacji 375
Nowe mierniki a mierniki już istniejące 375
Operacjonalizacja sukcesu lub porażki 376
Typy schematów badań ewaluacyjnych 377
Schematy eksperymentalne 377
Schematy quasi-eksperymentalne 378
Ewaluacja jakościowa 383
Kontekst społeczny 383
Problemy logistyczne 383
Niektóre zagadnienia etyczne 385
Wykorzystanie wyników badań 387
Badania wskaźników społecznych 391
Kara śmierci i odstraszanie 392
Symulacja komputerowa 393

Główne zagadnienia 394
Kluczowe pojęcia 394
Pytania i ćwiczenia 395
Dodatkowe lektury 395
Socjologiczna strona w sieci 395
Infotrac College Edition 396

Część 4

Analiza danych 397

13 Analiza danych jakościowych 399

Wprowadzenie 400
Powiązanie teorii z analizą 400
Odkrywanie prawidłowości 400
Metoda teorii ugruntowanej 402
Semiotyka 403
Analiza konwersacyjna 405
Przetwarzanie danych jakościowych 406
Kodowanie 406
Zapisywanie notatek 411
Tworzenie mapy pojęciowej 412
Programy komputerowe dla danych jakościowych 412
Księga Kapłańska widziana oczami programu NUD*IST 414
Sandrine Zerbib: zrozumieć kobietę reżysera 417

Główne zagadnienia 422
Kluczowe pojęcia 422
Pytania i ćwiczenia 422
Dodatkowe lektury 423
Socjologiczna strona w sieci 423
Infotrac College Edition 423

14 Kwantyfikacja danych 424

Wprowadzenie 425
Komputery w badaniach społecznych 425
Kodowanie 428
Tworzenie kategorii kodowych 429
Budowa książki kodowej 431
Sposoby kodowania i wprowadzania danych 432
Arkusze kodowe 432
Kodowanie brzegowe 433
Bezpośrednie wprowadzanie danych 433
Wprowadzanie danych przez ankieterów 433
Kodowanie na karty do skanera optycznego 433
Bezpośrednie użycie kart do skanera optycznego 434
Powiązanie z programem do analizy danych 434
Czyszczenie danych 434
Czyszczenie dotyczące dopuszczalnych kodów 434
Czyszczenie logiczne 435

Główne zagadnienia 435
Kluczowe pojęcia 436
Pytania i ćwiczenia 436
Dodatkowe lektury 437
Socjologiczna strona w sieci 437
Infotrac College Edition 437

15 Podstawy analizy ilościowej 438

Wprowadzenie 439
Analiza jednozmiennowa 439
Rozkłady 439
Tendencja centralna 441
Rozproszenie 443
Zmienne ciągłe i dyskretne 444
Szczegółowość a czytelność 445
Porównywanie podgrup 445
Łączenie kategorii odpowiedzi 446
Sposoby postępowania z kategorią „nie wiem" 447
Opisy liczbowe w badaniach jakościowych 448
Analiza dwuzmiennowa 449
Procentowanie w tabeli 450

Tworzenie i odczytywanie tabel dwuzmiennowych 453
Wprowadzenie do analizy wielozmiennowej 454
Główne zagadnienia 456
Kluczowe pojęcia 457
Pytania i ćwiczenia 457
Dodatkowe lektury 457
Socjologiczna strona w sieci 458
Infotrac College Edition 458

16 Model analizy rozbudowanej 459

Wprowadzenie 460
Początki modelu analizy rozbudowanej 460
Paradygmat analizy rozbudowanej 465
Replikacja 466
Eksplanacja 466
Interpretacja 470
Specyfikacja 470
Doskonalenie paradygmatu 471
Analiza rozbudowana i stawianie hipotez *post factum* 474
Główne zagadnienia 475
Kluczowe pojęcia 476
Pytania i ćwiczenia 476
Dodatkowe lektury 476
Socjologiczna strona w sieci 477
Infotrac College Edition 477

17 Statystyka w naukach społecznych 478

Wprowadzenie 479
Niebezpieczeństwo sukcesów matematycznych 479
Statystyka opisowa 480
Redukcja danych 481
Miary siły związku 481
Analiza regresji 486
Inne techniki wielozmiennowe 491
Analiza ścieżkowa 491
Analiza szeregów czasowych 492
Analiza czynnikowa 494
Wnioskowanie statystyczne 497
Wnioskowanie jednozmiennowe 497
Testy istotności statystycznej 498
Logika istotności statystycznej 499
Test chi-kwadrat 503
Główne zagadnienia 508
Kluczowe pojęcia 509
Pytania i ćwiczenia 509
Dodatkowe lektury 509
Sociologiczna strona w sieci 509
Infotrac College Edition 510

Część 5

Społeczny kontekst badań 511

18 Etyka i polityka w badaniach społecznych 513

Wprowadzenie 514
Zagadnienia etyczne w badaniach społecznych 515
Dobrowolne uczestnictwo 515
Nie krzywdzić badanych 516
Anonimowość i poufność 518
Oszukiwanie badanych 520
Analiza i prezentacja wyników 521
Instytucjonalne organa kontrolne 523
Kodeksy etyki zawodowej 523
Dwie kontrowersje etyczne 523
Kłopoty w kafejce 525
Obserwacja posłuszeństwa 526
Polityczne aspekty badań społecznych 527
Obiektywność a ideologia 528
Polityka przez małe „p" 531
Polityka we właściwych proporcjach 532
Główne zagadnienia 533
Kluczowe pojęcia 534
Pytania i ćwiczenia 534
Dodatkowe lektury 536
Socjologiczna strona w sieci 536
Infotrac College Edition 536

19 Zastosowania badań społecznych 537

Wprowadzenie 538
Teoria – zbieranie danych – analiza 538
Wykorzystanie umiejętności badawczych na codzień 539
Poradnik konsumenta badań społecznych 541
Plan badań 541
Pomiar 542
Dobór próby 542
Eksperyment 543
Badania sondażowe 543
Badania terenowe 544
Analiza istniejących danych statystycznych 544
Badania ewaluacyjne 544
Analiza danych 545
Raportowanie 545
Zakończenie 545
Socjologiczna strona w sieci 546
Infotrac College Edition 546

Załączniki 547

A Jak korzystać z biblioteki 548

B Badania społeczne w cyberprzestrzeni 554

C Raport z badań 563

D Kwestionariusz struktury gospodarstwa domowego z badania GSS 570

E Liczby losowe 580

F Rozkład chi-kwadrat 582

G Powierzchnie pod krzywą normalną 683

H Szacunkowy błąd z próby 584

I Podręcznik SPSS 585

J Dwadzieścia pytań, które dziennikarz powinien zadać, zapoznając się z wynikami sondażu opinii 616

Słowniczek 623

Bibliografia 634

Bibliografia uzupełniająca 643

Indeks osób 649

Indeks rzeczowy 653

Wprowadzenie

„Kilka" lat temu (nie chciałbym mówić, ile dokładnie) pierwszy raz prowadziłem zajęcia z metod badań społecznych. Tematem moich zajęć były przede wszystkim metody badań sondażowych, a uczestniczyło w nich tylko sześciu studentów. W miarę upływu czasu rozluźniłem się nieco jako nauczyciel i wkrótce zaczęliśmy się spotykać ze studentami w moim pokoju, gdzie mogłem im przynosić i pożyczać książki, skoro okazało się w czasie zajęć, że mogą się im przydać.

Stanąłem wtedy wobec naglącego problemu braku odpowiedniego podręcznika na temat badań sondażowych. Wszystkie dostępne książki dzieliły się na dwie kategorie. Jedne z nich prezentowały teoretyczną logikę metod badawczych w języku tak abstrakcyjnym, że nie sądziłem, by studenci byli w stanie zastosować którąkolwiek z tych ogólnych reguł w praktycznym świecie „robienia" badań. Pozostałe książki, nazywane często „książkami kucharskimi", wręcz przeciwnie – zawierały szczegółowe instrukcje, jak krok za krokiem przeprowadzić sondaż. Niestety w ten sposób studenci przygotowywali się przede wszystkim do prowadzenia badań podobnych do opisanych przez autorów. Ani podejście abstrakcyjne, ani „książka kucharska" nie wydawały się szczególnie użyteczne z punktu widzenia studentów i ich nauczycieli.

Pewnego dnia zacząłem układać spis treści mojego wymarzonego podręcznika metod badawczych. Był on zorgnizowany wedle podziału na trzy zasady teoretyczne, na których opierają się badania społeczne:

1. Rozumienie *teoretycznych* podstaw badań naukowych.
2. Świadomość stosowania tych podstawowych zasad w ustalonych technikach *prowadzenia* badań.
3. Gotowość do poszukiwania odpowiednich *kompromisów*, kiedy warunki realizacji nie pozwalają na rutynowe zastosowanie ustalonych technik.

Następnego dnia dostałem niespodziewany telefon z Wadswortha – prosili mnie, żebym napisał dla nich jakiś tekst metodologiczny! *Metody badań sondażowych* (*Survey Research Methods*) zostały opublikowane w 1973 roku. Zarówno wydawca, jak i ja otrzymaliśmy potem kilka dobrych wiadomości, kilka złych i dodatkowo jeszcze kilka dobrych. Pierwszą dobrą wiadomością było, że wszyscy nauczyciele metod badawczych pokochali od razu tę książkę i wydawało się, że używa się jej niemal na każdym kursie badań sondażowych w całym kraju. Złą wiadomością było natomiast, że nie ma wcale zbyt wielu tego typu kursów.

Dodatkową dobrą wiadomością było jednak, że wielu nauczycieli akademickich prowadzi bardziej ogólne zajęcia na temat badań społecznych, w toku których badania sondażowe omawiane są wraz z innymi metodami badawczymi. Prowadzący ci używali na ogół naszej książki, uzupełniając ją innymi podręcznikami na temat badań terenowych, eksperymentu itp. Korzystając z naszej specjalistycznej książki, wielu nauczycieli sugerowało Wadsworthowi, że dobrze byłoby, gdyby „ten sam facet" napisał coś ogólnie o badaniach społecznych.

W przedmowie do pierwszego wydania *Badań społecznych w praktyce* (1975) dziekuję za współpracę całemu tuzinowi nauczycieli badań społecznych od Kalifornii po Florydę. Książka była w istocie owocem ścisłej współpracy, mimo że na okładce figurowało tylko moje nazwisko i to ja byłem za nią ostatecznie odpowiedzialny.

w stanie Waszyngton, by zilustrować różnice między średnią a medianą.

Rozdział 18: Rozszerzyłem jeszcze bardziej rozważania o etyce dzięki wykorzystaniu nowych przykładów, wzbogaciłem też rozważania na temat ochrony prywatności badanych.

Rozdział 19: Poświęciłem więcej uwagi podejściu jakościowemu.

Jak zawsze, uaktualniłem materiały użyte w książce. Starałem się przyzwyczaić do dat publikacji, które nie zaczynają się od 19. Teraz, kiedy kończę już wprowadzać poprawki, coraz bardziej przekonuję się do licznych materiałów pomocniczych dostępnych przez Internet, które i tobie mogą się wydać interesujące.

Jako nauczyciel poszukiwałem zawsze nowych, lepszych sposobów wyjaśnienia studentom, na czym polegają badania społeczne – wiele spośród tych wyjaśnień przybrało formę schematów graficznych. W tym wydaniu znajdziesz wiele nowych ilustracji. Również tu starałem się zastąpić starzejące się przykłady badań (z wyjątkiem klasycznych) nowymi. Pominąłem też kilka fragmentów, które – jak sądzę – nie przydają się już studentom.

Podobnie jak w przypadku każdego nowego wydania, będę wdzięczny za wszelkie uwagi dotyczące możliwości ulepszenia tej książki. Jej ponad 25-letnia ewolucja odzwierciedla niezliczone sugestie studentów i pracowników naukowych. Jestem im wdzięczny za współpracę przy tworzeniu tej książki.

Walory pedagogiczne

Studenci i ich nauczyciele informowali mnie co prawda, że poprzednie wydania tej książki były świetnym narzędziem do nauki metod badawczych, postanowiłem jednak wykorzystać okazję, by odświeżyć ją także z pedagogicznego punktu widzenia, dostrajając pewne elementy, dodając inne. Oto zestaw, który mamy w dziewiątym wydaniu.

Przegląd holograficzny: Podobnie jak w poprzednich wydaniach przegląd holograficzny przed rozdziałem 1 prezentuje całą książkę w jednym zdaniu, w jednym akapicie, na mniej więcej jednej stronie i wreszcie w kilku akapitach omówienia dla każdego rozdziału. Zawsze uważałem, że poszczególne elementy badań społecznych są najbardziej zrozumiałe, kiedy najpierw zna się całość. Ta część książki funkcjonuje nie tylko jako wstępny przegląd, ale i jako powtórzenie po zakończeniu lektury rozdziału lub całej książki.

Przegląd rozdziału: Idea „przeglądu holograficznego" powraca na początku każdego rozdziału jako nieduży, zwarty akapit objaśniający najistotniejsze elementy treści rozdziału.

Wprowadzenie do rozdziału: Każdy rozdział zaczyna się wprowadzeniem, które wymienia główne tematy danego rozdziału i – co nie mniej ważne – wiąże je z treścią pozostałych rozdziałów książki. Wprowadzenia są kontynuacją pomysłu zainicjowanego w przeglądzie holograficznym, by sensownie powiązać ze sobą poszczególne idee w miarę, jak student zagłębia się w lekturze.

Jasne i prowokujące przykłady: Studenci często mówią mi, że przykłady – prawdziwe i hipotetyczne – pomogły im zrozumieć trudne i/lub abstrakcyjne koncepcje. To wydanie zawiera wiele nowych przykładów obok tych, których szczególna przydatność została dowiedziona w poprzednich wydaniach.

Szata graficzna: Od czasu, kiedy pierwszy raz uczestniczyłem w kursie metod badawczych, większość kluczowych pojęć wydawała mi się najbardziej zrozumiała w formie graficznej. Moje zadanie polegało na przełożeniu obrazów mentalnych na słowa, ale i na ilustracje graficzne zawarte w tej książce. Postęp grafiki komputerowej ułatwił mi wyjaśnienie artystom z wydawnictwa Wadsworth, co widzę oczyma duszy i co chciałbym przekazać studentom. Jestem wręcz zachwycony nową szatą graficzną tego wydania.

Przykłady i rozważania w ramkach: Studenci mówią mi też, że podobają im się materiały w ramkach, które podkreślają wagę poszczególnych koncepcji i badań, przy okazji zaś urozmaicają format książki. W wydaniu dziewiątym zacząłem stosować ramki

dla tych części materiału, które skupiają się na użyciu i nadużywaniu badań społecznych przez media.

Zasoby internetowe: Rola Internetu rośnie z każdym kolejnym wydaniem. Również i to wydanie kontynuuje ruch w kierunku cyberprzestrzeni. W samej książce znajduje się wiele użytecznych ćwiczeń i cytowań wykorzystujących Internet, a strona www poświęcona tej książce (http://sociology.wadsworth.com) została znacznie wzbogacona zarówno z punktu widzenia studentów, jak i nauczycieli.

Główne zagadnienia: Krótkie zestawienie głównych zagadnień na końcu każdego rozdziału to zarazem podsumowanie rozdziału i skuteczna pomoc w powtarzaniu. Główne zagadnienia pozwalają studentom zorientować się, na jakich problemach powinni się skupić w każdym rozdziale.

Kluczowe pojęcia: Jako nowość pojawia się w tym wydaniu zestawienie kluczowych pojęć, następujące po głównych zagadnieniach. Zestawienie to ułatwia studentom przyswajanie niezbędnego słownictwa. W treści rozdziału znajdują się definicje kontekstowe nowych pojęć występujących w zestawieniu. Pojęcia te zostały w tekście wyróżnione pogrubionym drukiem, ich definicje zaś znajdują się w słowniczku na końcu książki.

Pytania i ćwiczenia: W tym wydaniu poszerzono o nowe pytania i ćwiczenia tę część, służącą jako pomoc w powtórzeniu. Co najmniej jedno pytanie w zestawie wymaga skorzystania z Internetu.

Załączniki: Podobnie jak w poprzednich wydaniach, zestaw załączników pomaga studentom radzić sobie z biblioteką (materialną i wirtualną), pisaniem raportów oraz cyberprzestrzenią.

Jasny i przystępny styl: Jest to być może najistotniejszy „walor pedagogiczny" książki. Wiem oczywiście, że wszyscy starają się pisać jasno i przystępnie, i jestem dość dumny, że ta „cecha szczególna" mojej książki była jedną z jej najbardziej chwalonych zalet w ośmiu poprzednich wydaniach – to o tym właśnie studenci piszą najczęściej. W wydaniu dziewiątym poświęciłem szczególnie wiele uwagi, by dokładnie przejrzeć dosłownie każdą linijkę tekstu. Przycinałem, wygładzałem, ozdabiałem, a niekiedy zmieniałem całą strukturę tekstu, by uczynić go jak najbardziej „przyjaznym" czytelnikowi. Niezależnie od tego, czy jesteś nowym czytelnikiem, czy też znasz dokładnie poprzednie wydania, zachęcam, byś otworzył książkę na którymkolwiek rozdziale i samodzielnie ocenił ją pod kątem stylu.

Dodatki

Strona www

Jako że coraz więcej studentów i osób prowadzących zajęcia dysponuje dostępem do Internetu, starałem się też wykorzystać to medium jako dodatkową pomoc przy korzystaniu z książki. Obecnie mamy już stronę www poświęconą niniejszej książce, pod adresem http://sociology.wadsworth.com/. Mimo że jest tam już wiele pożytecznych materiałów, zamierzam w najbliższej przyszłości skupić wszystkie swoje siły na tworzeniu materiałów elektronicznych, które w znacznym stopniu ułatwią naukę i stosowanie metod badawczych.

Dostępne elementy kolekcji[2]

Surfing Sociology

Ten poręczny, kolorowy folder zawiera adresy www rozmaitych centrów, stowarzyszeń i organizacji, gromadzących zasoby socjologiczne.

Sociology Internet Companion

Ten dysk kompatybilny z IBM PC pracuje łącznie z przeglądarką internetową, zapewniając dostep do stron www związanych z socjologią. Zawiera też linki do ponad 800 stron.

[2] Tylko po angielsku (przyp. red.).

The Practice of Social Research Video

Film ten zawiera sześć dziesięciominutowych części, omawiających tradycyjnie problematyczne pojęcia, jak operacjonalizacja, dobór próby, schematy eksperymentalne, sformułowanie teorii, etykę, indeksy i tabele procentowe oraz zmienne.

Investigating Change in American Society: oprogramowanie i podręcznik

Ten dodatek pozwala studentom uczestniczyć w analizach i przetwarzaniu zbiorów danych pochodzących ze spisu ludności Stanów Zjednoczonych. Zawiera on ćwiczenia dla każdego zestawu danych oraz łatwy w użyciu program statystyczny z przewodnikiem po analizie danych.

Podziękowania

Nie jestem w stanie podziękować należycie wszystkim ludziom, którzy wpłynęli na kształt tej książki. Mój poprzedni tekst metodologiczny, *Metody badań sondażowych*, zadedykowałem Samuelowi Stoufferowi, Paulowi Lazarsfeldowi i Charlesowi Glockowi. Po raz kolejny pragnąłbym wyrazić im swoją wdzięczność.

Chciałbym też powtórzyć moje podziękowania pod adresem tych kolegów, którym dziękowałem już za uwagi podczas pisania pierwszego, drugiego i trzeciego wydania tej książki. Poniższa praca, mimo że wielokrotnie zmieniana, nadal odzwierciedla ich wkład. W udoskonalaniu tej książki pomogło mi także wielu innych kolegów. Szczególnie chciałbym podziękować tym nauczycielom akademickim, którzy przejrzeli rękopis obecnego wydania i zgłosili cenne propozycje:

Anne Baird, Morehouse College
Rae Banks, Syracuse University
Joseph Fletcher, University of Toronto
Roland Hawkes, Southern Illinois University
Daniel J. Klenow, North Dakota State University
Cecilia Menjívar, Arizona State University
Alisa Potter, Concordia College
Zhenchoa Qian, Arizona State University
Robert W. Reynolds, Weber State University
Laurie K. Scheuble, Doane College
Matthew Sloan, University of Wisconsin-Madison
Bernard Sorofman, University of Iowa
Robert Wolf, Eastern Connecticut State University

Chciałbym też podziękować Rolandowi Chiltonowi z University of Massachusetts w Amherst, M. Richardowi Cramerowi z University of North Carolina w Chapel Hill, Shaulowi Gabbayowi z University of Illinois w Chicago, Marcii Ghidina z University of North Carolina w Asheville, Jeffreyowi Jacquesowi z Florida A&M University, Wandzie Kosinski z Ramapo College w New Jersey, Joan Morris z University of Central Florida, Beth Anne Shelton z University of Texas w Arlington, Ronowi Stewartowi i Randy'emu Stokerowi z University of Toledo, Theodore'owi Wagenaarowi z Miami University w Ohio oraz Jerome'owi Wolfe'owi z University of Miami.

Z biegiem lat coraz większe wrażenie wywiera na mnie waga roli wydawcy w przypadku książek takich jak ta. Nazwisko autora pojawia się co prawda na grzbiecie książki, jednak jej szkielet czerpie siłę od wydawcy. Od roku 1973 współpracowałem w wydawnictwie Wadsworth z sześcioma redaktorami do spraw socjologii, co wymagało mniej więcej takiego samego poziomu dopasowania charakterów, jak w sześciu kolejnych małżeństwach. Na szczęście to wydanie książki bardzo skorzystało na mojej współpracy z Ewą Howard. Ewa, oprócz bogactwa doświadczeń edytorskich, ma również wyczucie nowych technik i pedagogiki, i potrafi z nich korzystać.

Poniższe wydanie zyskało szczególnie dzięki pracy dwóch innych redaktorów. Po pierwsze, John Bergez podjął się dokonania całościowego przeglądu tekstu i wychwycenia wszystkich luźnych myśli i nieścisłości, które zawsze wkradają się do wielokrotnie poprawianej książki. John wyszukał także wszystkie przykłady, które wymagały uaktualnienia, argumenty domagające się uściślenia, sformułowania, które mogły być zgrabniejsze, oraz dołożył wszelkich innych starań, by ulepszyć tę książkę. Poza tym John zgłosił też kilka uwag, które pozwolily wprowadzić inne poprawki.

Molly Roth to osoba, z którą powinno się porównywać wszystkich redaktorów świata, aczkolwiek stawiałoby to być może poprzeczkę zbyt wysoko. Molly i ja pracowaliśmy już razem nad kilkoma książkami – ona jest po prostu najlepsza. Doskonale porusza się wzdłuż cienkiej granicy między niechęcią do stwierdzenia, że autor zawiódł, a upodobaniem do tego typu

stwierdzeń. Nigdy nie miałem uczucia, że będzie mi w jakikolwiek sposób pobłażać, nigdy też nie spotkałem się z jej strony z niczym innym, niż całkowite poparcie dla moich zamiarów. Molly w jakiś sposób orientuje się, co chcę powiedzieć, i często potrafi znaleźć sposób, by powiedzieć to dobitniej.

Ted Wagenaar również wniósł wielki wkład w tę książkę. Ted i ja jesteśmy współautorami uzupełniających książkę ćwiczeń pod tytułem *Practicing Social Research*, ale to tylko wierzchołek góry lodowej. Ted jest świetnym kolegą, cenionym krytykiem, dobrym przyjacielem i w ogóle porządnym człowiekiem.

Proces wprowadzania poprawek skomplikował się, gdy okazało się, że mam raka. Teraz już go nie mam. Dzięki wysiłkom Jean deKernion, Mitchell Rausch i centrum medycznego Uniwersytetu Kalifornijskiego w Los Angeles udało mi się wyleczyć chorą prostatę i powrócić niedługo potem do nauczania i poprawiania książki. Moja wdzięczność sięga oczywiście daleko poza możliwość poprawiania książki.

Przy okazji tego wydania książki możesz poznać młodą socjolog, o której usłyszysz jeszcze w przyszłości: Sandrine Zerbib. Sandrine, która kończy właśnie doktorat w University of California w Irvine, jest pierwszorzędnym metodologiem i naukowcem pracującym zarówno po jakościowej, jak i ilościowej stronie barykady. Jest szczególnie wyczulona na perspektywę feministyczną, przy czym jej kobiece doświadczenie wnosi nowe spojrzenie do naszych wspólnych zainteresowań metodologicznych. Wysiłek Sandrine jest szczególnie widoczny w rozdziałach 10 i 14, ale wpłynęła ona także na kształt innych części książki. Współpracowaliśmy więc po partnersku i oczekuję dalszego ciągu tej współpracy.

Zadedykowałem tę książkę mojej żonie, Sheili, która przyczyniła się w wielkim stopniu do jej powstania i ewolucji. Sheila i ja spotkaliśmy się, gdy została przydzielona do asystowania mi przy projekcie badawczym, który nadzorowałem w Centrum Badań Sondażowych przy Uniwersytecie Kalifornijskim w Berkeley*. W trakcie naszego ponadtrzydziestoletniego małżeństwa pracowaliśmy przy licznych projektach i przypuszczam, że będzie ich jeszcze więcej w przyszłości. Jestem jednak wdzięczny Sheili za znacznie więcej niż tylko naszą wspólną pracę badawczą. Jest ona bowiem wspaniałą partnerką życiową. Jej zrozumienie i wsparcie prowadzą mnie zawsze do kresu moich zainteresowań i pozwalają mi spojrzeć dalej. Nie mam słów, by jej za to należycie podziękować.

---

* Zawsze uważałem za zabawne powtarzanie, że Sheila wyszła za swojego szefa. Po 35 latach mam wrażenie, że to ja ożeniłem się ze swoim.

Przegląd holograficzny

Badania społeczne składają się z wielu elementów, które niełatwo wyłożyć w prostym porządku chronologicznym. W toku długich lat nauczania metod badawczych w salach wykładowych i pisania podręczników na ten temat nigdy nie udało mi się zadowalająco rozwiązać problemu porządku prezentacji materiału. Na ogół starałem się uporządkować wykłady i rozdziały zgodnie z kolejnością kroków, które podejmujemy, prowadząc badania, choć nie każdy temat zamyka się całkowicie w jednym stadium tego procesu. Dobrym przykładem jest tu etyka badawcza. Dotyczy ona niemal każdego rozdziału tej książki, jednak szczegółowe omówienie etyki nabiera znaczenia dopiero, gdy rozumie się już całą resztę książki.

Chciałbym, żebyś spojrzał na badania społeczne jak na hologram – obraz powstający w rezultacie dość dziwnego procesu fotograficznego. Gdybyśmy chcieli zrobić sobie razem zwykłe zdjęcie, wykonalibyśmy odbitkę z negatywu i gotowe: mielibyśmy wspólne zdjęcie. Gdybyśmy przecięli negatyw na pół i zrobili odbitkę z jednej części, tylko jedno z nas byłoby na zdjęciu. Negatyw holograficzny jest jednak zadziwiająco inny. Gdybyśmy przecięli go na pół i zrobili odbitkę, powstały w ten sposób obraz byłby dalej obrazem nas obu, tyle że trochę mniej wyraźnym. Przeciąwszy z kolei ten negatyw na pół, otrzymalibyśmy znowu cały obraz, jeszcze mniej wyraźny.

Niniejszy podręcznik – a zwłaszcza jego prezentacja w tym wstępie – skonstruowany został na wzór hologramu. Za chwilę przedstawię ci cały jego temat w jednym zdaniu. Ponieważ zaś pokażę ci całość bardzo szybko, obraz przedmiotu będzie nieco niewyraźny.

Po tym pierwszym krótkim omówieniu przedstawię ci wszystko w jednym akapicie, następnie zaś na kilku stronach. Potem omówię jeszcze raz wszystko nieco bardziej szczegółowo. Reszta książki stanowi kolejną prezentację całego tematu – ze wszystkimi szczegółami, które składają się na gruntowną wiedzę o badaniach społecznych.

Przekonasz się, że w taki sam sposób skonstruowane są poszczególne rozdziały. Na pierwszej stronie każdego z nich znajdziesz „Przegląd holograficzny", który przedstawia sedno rozdziału – całą jego treść – w jednym lub dwóch zdaniach. Następnie „Wprowadzenie" do rozdziału przedstawia temat nieco bardziej szczegółowo. W dalszej części rozdziału znajdziesz pełne rozwinięcie wszystkich zagadnień tematu. Część końcowa rozdziału, zatytułowana „Główne zagadnienia", stanowi coś w rodzaju holograficznego podsumowania, które pomoże ci powtórzyć materiał. W powtórzeniu przyda ci się także umieszczone na końcu rozdziału zestawienie kluczowych pojęć, które zostały wytłuszczone w tekście i zdefiniowane w słowniczku na końcu książki.

Poniższy przegląd holograficzny książki może ci być pomocny na wiele sposobów. Po pierwsze, wstępna lektura stworzy kontekst dla zapamiętania i zrozumienia rozmaitych szczegółowych informacji dotyczących badań społecznych. Po drugie, możesz powracać do tego rozdziału co pewien czas w miarę postępów lektury. Odwoływanie się do tego rozdziału pomoże ci nie tracić z oczu lasu, w miarę zapoznawania się z poszczególnymi drzewami. Po trzecie, kiedy już przeczytasz całą książkę, możesz wykorzystać ten rozdział jako pomoc w powtórzeniu. Możesz wrócić do całości obrazu, umieścić wszystkie elementy w kontekście i na nowo ocenić sposób ich wzajemnego powiązania.

Możesz wreszcie użyć fragmentów tego rozdziału, by zaspokoić ciekawość rodziny i przyjaciół, wypytujących, co też tak sympatyczna osoba robi na takich zajęciach.

Całość: Hologram 1

Badania społeczne polegają na systematycznej obserwacji życia społecznego w celu odkrycia i zrozumienia prawidłowości, które wykazuje przedmiot obserwacji.

Całość: Hologram 2

Ludzie wchodzą ze sobą w interakcje i tworzą struktury tych interakcji. Przedstawiciele nauk społecznych starają się odkryć naturę relacji międzyludzkich, podobnie jak biologowie starają się zrozumieć funkcjonowanie organizmów żywych, a astronomowie – ruchy ciał niebieskich. Na ogół gdy ludzie obserwują to, co się wokół nich dzieje, i starają się to zrozumieć, wpływ co najmniej trzech czynników kładzie się cieniem na ich wysiłkach. Owe czynniki to: (1) opinie na temat tego, jak być powinno, (2) błędne mniemania i przesądy, których nabyli, dorastając oraz (3) niestaranne i błędne obserwacje. Badacze społeczni natomiast podejmują obserwację jako działanie świadome i przemyślane – w istocie często jako działanie zespołowe. Wyniki swoich obserwacji analizują następnie starannie i obiektywnie, by określić, jak *jest*, abstrahując od własnych odczuć na temat tego, jak *powinno* być.

Całość: Hologram 3

Część 1 tej książki stwarza kontekst dla zrozumienia logiki i procedur właściwych badaniom społecznym. Będę rozróżniał badania naukowe i to, co zachodzi w życiu codziennym, zwracając uwagę na niektóre powszechne błędy w obserwacji i interpretacji. Przekonamy się, że badania społeczne zmierzają do odkrycia i zrozumienia prawidłowości w tym, co wydaje się niekiedy chaosem codziennego życia, oraz zbadamy rolę zmiennych w tym przedsięwzięciu.

Paradygmaty i teorie kształtują cel naszych poszukiwań i interpretację ich wyników. W sposób bezwiedny zachodzi to w codziennych procesach poznawczych, jednak badacze społeczni robią to bardziej otwarcie – omówimy kilka najważniejszych spośród stosowanych przez nich paradygmatów i teorii. W rozdziale 3 zajmiemy się ideą przyczynowości – kolejną, którą przyjmujemy za pewnik w życiu codziennym, a która staje się precyzyjnym narzędziem w rękach badaczy społecznych.

Już na początku przekonasz się, że badania społeczne są ograniczone i uwarunkowane rozmaitymi kwestiami etycznymi. W rozdziale 1 otrzymasz ogólne informacje o kwestiach etycznych, abyś był w stanie ocenić ich konsekwencje w miarę zapoznawania się z kolejnymi technikami badawczymi. Pod koniec książki, w rozdziale 18 zajmiemy się etyką badawczą bardziej gruntownie.

Jak wszystkie dziedziny nauki, badania społeczne koncentrują się wokół dwóch czynności: pomiaru i interpretacji. W części 2 książki rozważymy, jak badacze społeczni dokonują pomiaru różnych aspektów rzeczywistości społecznej i jak wyciągają wnioski co do znaczenia wyników swoich pomiarów.

W tej książce używam terminów *pomiar* i *obserwacja* w pewnym sensie zamiennie, jako że pomiar jest w istocie specyficznym typem obserwacji. Chociaż wszyscy ciągle obserwujemy, w każdym razie w sposób nieformalny, termin *pomiar* odnosi się do bardziej przemyślanej i uregulowanej formy aktywności. Uczeni podejmują świadomą decyzję, że będą obserwować, równie świadomie wybierają przedmiot obserwacji, zabezpieczają się na wypadek błędnych obserwacji, wyniki zaś swoich obserwacji rejestrują starannie w formie pomiarów. Pomiary bywają niekiedy liczbowe lub ilościowe: „W czasie egzaminu studenci wychodzili do toalety średnio 2,4 raza". Bywają też jakościowe: „Kierowca wyraził swoją irytację, gdy dowiedział się, że pomoc drogowa skończyła pracę o 17.00".

Pomiar zaczyna się od obserwacji, ale pierwszy surowy materiał wymaga często dopracowania i wyklarowania. W rozdziale 5 zajmiemy się wstępną strukturyzacją systematycznej obserwacji – na przykład w trakcie projektowania kwestionariusza. Wrócimy do tego tematu w rozdziale 6, kiedy to zbadamy sposo-

by tworzenia miar złożonych, takich jak indeksy, skale i typologie. Tego typu złożone mierniki pozwalają połączyć kilka szczegółowych informacji o ludziach, aby następnie opisać ich w kategoriach ogólnych tendencji, na przykład uprzedzeń, poziomu intelektualnego, alienacji czy poglądów politycznych.

W naukach społecznych, bardziej niż w przyrodniczych, to, *co* mierzymy, jest równie ważne, jak to, *jak* mierzymy. Na przykład w sondażu wyborczym musimy pytać o opinie badanych w taki sposób, by nie wpływać na ich odpowiedzi, bardzo ważne jest jednak również, by wybrać do badania właściwych wyborców. Samo przebadanie ludzi w zamożnej dzielnicy podmiejskiej nie pozwoli na trafne przewidywanie wyników głosowania całego elektoratu, podobnie jak badanie wyłącznie mieszkańców zubożałego miejskiego getta. Dobór próby to precyzyjna naukowa dziedzina, która pozwala badaczom przepytać około 1500 głosujących, by przewidzieć w miarę trafnie, jak będzie głosować około 100 milionów ludzi. Dowiesz się, jak to zrobić, w rozdziale 7. Podobnie badacz dokonujący obserwacji uczestniczącej, poszukując wskazówek ze strony informatorów, musi być pewien, że wybiera takich, którzy są w stanie dostarczyć odpowiedniej informacji na temat badanej sytuacji.

W naukach przyrodniczych używa się suwmiarek, spektrografów i innych specjalistycznych narzędzi pomiarowych, w badaniach społecznych zaś rozmaitych technik, z których niektóre poznasz w części 3 tej książki. Badacze społeczni używają niekiedy kwestionariuszy w badaniach sondażowych. Eksperymentatorzy tworzą kontrolowane sytuacje i obserwują zachowania badanych. W obserwacji uczestniczącej badacze zanurzają się w sytuacje życia codziennego, które zamierzają zbadać, i obserwują je od wewnątrz. Prowadząc analizę treści, badacze mogą siedzieć przed telewizorem i używać kart, na których notują częstotliwość występowania różnych elementów. Przekonasz się, że rozmaitość technik pomiarowych możliwych do zastosowania w starannie prowadzonych badaniach społecznych jest niemal nieograniczona – jedyne granice stawia ci własna wyobraźnia.

Pomiar jest jednak tylko jednym z kluczy do badań społecznych. Drugim jest interpretacja, niekiedy nazywana *analizą danych*. Część 4 książki poświęcona jest wyłącznie analizie danych. Zobaczysz, jak na masie obserwacji skupia się logika, a niekiedy też matematyka, by odkryć znaczące prawidłowości. Niekiedy zależy nam na odkryciu prostych prawidłowości opisowych: jaki odsetek wyborców zagłosuje na konkretnego kandydata? W których miastach współczynnik przestępczości jest najwyższy?

Analiza w badaniach społecznych posługuje się tak zwanym językiem zmiennych. Mimo że przedmiotem zainteresowania badaczy społecznych są ludzie, starają się oni zrozumieć zmienne, które charakteryzują lub opisują tych ludzi i grupy, w których żyją. *Płeć* na przykład jest zmienną, w kategoriach której pewni ludzie są mężczyznami, inni zaś kobietami. Badacze mogą więc zająć się na przykład „rozkładem płci" w grupie młodocianych przestępców w danym mieście: jaka jest wśród nich proporcja dziewcząt, a jaka chłopców? Mogą też poświęcić szczególną uwagę zachowaniom przywódców demonstracji studenckich, aby rozróżnić i zrozumieć różne *style przywództwa* – czyli inną zmienną społeczną.

Oprócz opisywania ludzi w kategoriach zmiennych badacze społeczni często zainteresowani są odkrywaniem zależności między zmiennymi. W ten sposób badacze polityczni odkryli na przykład zależność między identyfikacją religijną a zachowaniami wyborczymi w Stanach Zjednoczonych. Rzymscy katolicy i żydzi częściej niż protestanci głosują na kandydatów partii demokratycznej. Jako przykład z innej dziedziny: badacze mówią niekiedy o różnych stylach przywódczych typowych dla kobiet i mężczyzn.

Przekonasz się, że interpretacja obserwacji wymaga nie tylko pomiaru – opiera się ona także na logicznym wyjaśnieniu przyczyn zależności między zmiennymi. Należy to do dziedziny teorii, którą zajmiemy się na początku książki, i do której będziemy w naszej podróży co pewien czas wracać.

Ostatecznie celem badań społecznych jest odkrywanie prawidłowości i zasad rządzących interakcjami międzyludzkimi w społeczeństwie – tym, co *jest*. Pod tym względem różnią się one

od religii czy innych systemów moralnych, które ustalają, jak być powinno. Nie oznacza to jednak, że badania społeczne są wolne od wartościowania, czy też że odbywają się one w próżni moralnej – zagadnienia te omawiam w części 5. Badania mają swoją własną listę „powinności", które dotyczą etycznego traktowania badanych. W rozdziale 18 podsumowuję krótko dotychczasową prezentację etyki badawczej i omawiam ten temat szczegółowo. Przedstawiam też wpływ nacisków politycznych na badania społeczne.

Co więcej, mimo że badania same w sobie nie są moralne ani niemoralne, można ich używać w sposób moralny i niemoralny (zakładam naturalnie, że jesteś zainteresowany tym pierwszym użyciem). Mądrze wykorzystane wyniki badań społecznych są skutecznym narzędziem kształtowania lepszego społeczeństwa: na przykład odkrycie przyczyn uprzedzeń może wskazać sposoby ich likwidacji. Niektórzy badacze studiują problem uprzedzeń właśnie z tego powodu. Można też używać wyników badań społecznych, by po prostu podejmować lepsze decyzje jako obywatel. Wszyscy stajemy na co dzień wobec rozmaitych oczekiwań co do rzeczywistości społecznej – znajomość metod badawczych może nam pomóc w mądrej ocenie tych roszczeń. Rozdział 19 poświęcony jest zastosowaniom badań społecznych.

Całość: Hologram 4

Rozdział 1. Działalność badawcza człowieka a nauka

Obserwacja i próby interpretacji tego, co obserwowane, są naturalną ludzką formą aktywności i podstawą naszego przetrwania w świecie. W życiu codziennym jesteśmy jednak często nieuważnymi i nieświadomymi obserwatorami i interpretatorami, błądzimy więc w jednej i drugiej dziedzinie. Uczeni czynią obserwację i interpretację aktami świadomymi i przemyślanymi.

Nie ma jednej metody prowadzenia badań społecznych. Badacze społeczni poszukują niekiedy ogólnego – acz cząstkowego – wyjaśnienia prawidłowości zachowań społecznych, innym razem skupiają się bardziej na głębokim wglądzie w specyfikę konkretnego przypadku. Badacze zaczynają niekiedy od ogólnej teorii i sprawdzają ją w toku kontrolowanych obserwacji. W innych przypadkach wybierają poszukiwanie ogólnej prawidłowości w obrębie tego, co obserwowane w drodze bardziej otwartego i swobodnego badania. Obiektywność jest co prawda istotnym pojęciem w każdym badaniu naukowym, jednak dla socjologa również subiektywność może stanowić decydujące narzędzie.

Rozdział 1 wprowadza zagadnienie etyki badawczej. Badacze społeczni są szczególnie wyczuleni na ewentualny wpływ, który ich badania wywierają na badanych ludzi, i zobowiązują się do ochrony dobra badanych. Ten istotny temat będzie przedmiotem szczegółowych rozważań w rozdziale 18, jednak w różnych miejscach książki – jeśli się to wyda stosowne – będziemy poruszać tematy etyczne.

Rozdział 2. Paradygmaty, teoria i badania społeczne

Omawianie metod badawczych koncentruje się na ogół na technikach pomiaru i analizy, wszystkie takie techniki zaś są wykorzystywane w systemach logicznych, zwanych paradygmatami. Te modele czy też systemy odniesienia ukierunkowują uwagę badaczy w trakcie prowadzenia pomiarów i kształtują teorie, które tworzą oni w celu ogólnego wyjaśniania.

Na przykład paradygmat funkcjonalistyczny każe nam traktować społeczeństwo jako system, złożony ze składników, z których każdy pełni w całości systemu jakąś funkcję. Orkiestra symfoniczna to dobry praktyczny przykład tego poglądu. Każdy z członków orkiestry ma jakieś konkretne zadanie do wykonania: dyrygent, perkusista, wiolonczelista. Kiedy słuchamy złożonego efektu ich wysiłków, rozumiemy, dlaczego każdy z nich robi to, co robi. Wyobraź sobie, że próbujesz zrozumieć zachowanie dyrygenta samo w sobie. Bywa, że zamyka się ludzi za tego typu zachowania.

Paradygmat interakcjonistyczny natomiast skupia naszą uwagę na typach interakcji i komunikacji międzyludzkiej, na wpływie działań jednej osoby na reakcje drugiej. Obserwując zespół jazzowy, możemy zainteresować się, w jaki sposób rytm podyktowany przez sekcję rytmiczną (bębny, kontrabas, fortepian) stał się podstawą dla solo saksofonu altowego, i jak z kolei inni muzycy podchwycili temat stworzony przez solistę. Taki sposób słuchania improwizacji Charliego Parkera i Milesa Davisa wydaje się bardziej rozsądny, niż badanie poszczególnych funkcji kolejno saksofonu i trąbki.

Badacze społeczni mogą opierać swoje badania na dowolnej liczbie paradygmatów. Żaden z nich nie jest prawdziwy lub fałszywy, co najwyżej bardziej lub mniej użyteczny w określonych okolicznościach. Każdy z nich kształtuje typ teorii, tworzonej na użytek ogólnego wyjaśniania. Na przykład teorie Karola Marksa na temat relacji między klasami społecznymi oparte były na paradygmacie konfliktu, który koncentruje się na wysiłkach jednej grupy, mających na celu zdominowanie innych grup i uniknięcie ich dominacji. W tym kontekście Marks mówił wiele na temat „walki klas", pojęcia, które miałoby o wiele mniejsze znaczenie w teorii funkcjonalistycznej, zajmującej się rolami poszczególnych klas w funkcjonowaniu całego społeczeństwa.

Związek między teorią a badaniami może być postrzegany w kontekście dwóch odrębnych, acz powiązanych wzajemnie, modeli. Metoda indukcyjna wychodzi od konkretnych, szczegółowych obserwacji i ma na celu odkrycie ogólnych zasad rządzących tym, co obserwowane. Moglibyśmy na przykład rozpocząć badanie przekraczania ulicy w niedozwolonych miejscach od szczegółowej obserwacji i rejestrowania zachowań pieszych na skrzyżowaniach. Następnie przeanalizowalibyśmy dane, by przekonać się, czy pewne typy ludzi częściej niż inne przechodzą przez ulicę w niedozwolonych miejscach.

Metoda dedukcyjna wychodzi natomiast od zasad ogólnych (teorii), następnie zaś przechodzi do obserwacji, by sprawdzić trafność przewidywań teoretycznych. Wychodząc od teorii oczekiwań co do ról płciowych w społeczeństwie amerykańskim, moglibyśmy na przykład wysnuć hipotezę, że mężczyźni powinni przechodzić w niedozwolonych miejscach częściej niż kobiety. Następnie przeszlibyśmy do obserwacji prawidłowości w zachowaniu pieszych, by określić, czy nasz teoretyczny wniosek był poprawny. Proces ten nazywamy *testowaniem hipotezy.*

Zarówno metoda indukcyjna, jak i dedukcyjna są w nauce uprawnione. Studiowanie poszczególnych tematów na ogół wymaga w praktyce łączenia ich obu: obserwacje mogą sugerować wyjaśnienia teoretyczne, wyjaśnienia te zaś sugerują, jakie inne prawidłowości powinniśmy móc zaobserwować. Kolejne obserwacje mogą stanowić test naszych oczekiwań, którego wyniki prowadzą do modyfikacji teorii i do kolejnych oczekiwań.

Rozdział 3. Pojęcie przyczynowości w badaniach społecznych

Niektóre badania społeczne są poświęcone przede wszystkim dokładnemu opisowi faktycznego stanu rzeczy. Jednak często badacze społeczni interesują się żywo *przyczynami* tego stanu rzeczy. Innymi słowy, starają się oni zrozumieć, co powoduje obserwowane zachowania i zdarzenia.

Rozważanie logiki przyczynowości w badaniach wymaga użycia języka zmiennych – mówimy, że jedne zmienne determinują inne lub są ich przyczyną. Badacze mogliby na przykład dojść do wniosku, że wyższy poziom wykształcenia powoduje zmniejszenie uprzedzeń, że przynależność religijna wpływa na zachowania wyborcze, czy też, że płeć determinuje postawy wobec kwestii napastowania seksualnego w miejscu pracy.

Jeśli bezwyjątkowy charakter powyższych zdań budzi twój sprzeciw, twoje intuicje są prawidłowe. Wykształcenie nie jest w końcu jedynym czynnikiem wpływającym na uprzedzenia, głosowanie odzwierciedla nie tylko przekonania religijne, a niektórzy mężczyźni mogą być równie zagorzałymi przeciwnikami napastowania seksualnego, jak niektóre kobiety.

W przyczynowym języku badań społecznych rzadko zdarza się, że jedna zmienna wa-

runkuje drugą w sposób zupełny i całkowity. Zachowanie wyborcze na przykład determinowane jest oczywiście przez identyfikację partyjną i wyznawaną filozofię polityczną, lecz również przez wykształcenie, dochód, region kraju, religię i inne zmienne. Zatem wyjaśnianie w naukach społecznych często bierze pod uwagę liczne zmienne wyjaśniające. Nie różni się to właściwie od sytuacji, w której wiele czynników może spowodować, że się przeziębisz – jeśli na przykład będziesz siedział w przeciągu, czy ktoś na ciebie kichnie, albo będziesz przemęczony itp.

Rozdział 4. Plan badań

Podstawowym celem badań społecznych jest obserwacja i interpretacja wyników, jednak cel ten można osiągnąć na wiele sposobów. Jednym z pierwszych kroków w planowaniu projektu badawczego jest sprecyzowanie *jednostki analizy* – przedmiotu naszego zainteresowania. W badaniach społecznych jednostką analizy jest na ogół pojedynczy człowiek, możemy jednak doprecyzować, że chodzi nam wyłącznie o obywateli Stanów Zjednoczonych, studentów wyższych uczelni, członków związków wyznaniowych, wyborców w Kalifornii czy inny podzbiór ludzkości. Jednostkami analizy mogą być również rodziny, organizacje, miasta, narody, filmy, wiersze czy jakakolwiek inna rzecz mająca odniesienia społeczne i obserwowalna. Moglibyśmy na przykład obejrzeć wiele teledysków rockowych, obserwując płeć wykonawców i mierząc stopień zawartej w ich występach wrogości lub przemocy. Następnie moglibyśmy określić, czy istnieje zależność między tymi dwiema zmiennymi.

Rozdział 5. Konceptualizacja, operacjonalizacja i pomiar

Pomiar jest jednym z filarów badań społecznych. W tym rozdziale zajmiemy się dwoma kluczowymi aspektami pomiaru: konceptualizacją i operacjonalizacją. W rozdziale 6 będziemy kontynuować rozważania nad pomiarem, analizując proces tworzenia złożonych mierników, takich jak indeksy, skale i typologie.

Konceptualizacja odnosi się do procesu określania i precyzowania pojęć, czyli naszych idei na temat natury rzeczy. Używaliśmy już w tym przeglądzie takich pojęć, jak uprzedzenia, wykształcenie, religia itp.

Gdybyśmy na przykład chcieli studiować zjawisko fanatyzmu, należałoby zapewne wyróżnić stereotypy (przekonania), dyskryminację (działania) i uprzedzenia (odczucia). Poza tym należałoby wyjaśnić, jakie obiekty fanatyzmu nas interesują: grupy rasowe lub etniczne, kobiety, mniejszości religijne, geje itp. Ten poziom specyfikacji jest niezbędny, jeśli chcemy stworzyć projekt badawczy; pozwoliłby nam również odkryć w dalszym ciągu badania, że przyczyny dyskryminacji gejów mogą być inne niż przyczyny dyskryminacji kobiet. Moglibyśmy też wykryć jakieś wspólne prawidłowości.

Żadne z pojęć, które określamy i badamy, nie ma „prawdziwej" definicji – mają one tylko dokładnie te znaczenia, które zgadzamy się im przypisać. Istotne jest, że w pewnym sensie znaczenie pojęcia *uprzedzenia* zależy od dokonanego przez nas wyboru – trudność zadania polega na wyszczególnieniu tych obserwacji, które możemy w sposób użyteczny objąć tym samym pojęciem.

Konceptualizacja dotyczy intelektualnego wyklarowania pojęć dla celów pomiaru, *operacjonalizacja* zaś jest konstruowaniem faktycznych, konkretnych technik pomiaru. Polega ona dosłownie na stworzeniu „operacji", które umożliwią przeprowadzenie pożądanych pomiarów.

Niekiedy satysfakcjonujący pomiar postaw i opinii jest możliwy z użyciem jednego pytania w kwestionariuszu. Oto prosty sposób poznania opinii ludzi na temat kary śmierci: „Czy jest Pan(i) zwolennikiem czy przeciwnikiem kary śmierci dla osób skazanych za morderstwo?". (Tak właśnie mierzy się tę zmienną w Generalnym Sondażu Społecznym [*General Social Survey – GSS*], ogólnonarodowym sondażu przeprowadzanym co rok lub dwa lata przez Krajowe Centrum Badania Opinii przy Uniwersytecie Chicago).

Weźmy zupełnie inny przykład – załóżmy, że zajmowalibyśmy się historyczno-porównaw-

czym badaniem zależności między wydatkami na zbrojenia a ubóstwem, w którym jednostkami analizy byłyby państwa. Musielibyśmy dokładnie określić, jak chcemy mierzyć każdą z tych zmiennych, poza tym zaś inne zmienne, jak liczbę ludności i produkt krajowy brutto. Być może korzystalibyśmy z oficjalnych statystyk badanych krajów lub też z danych ONZ. W rozdziale 5 omawiam szczegółowo kwestie, związane z operacjonalizacją zmiennych.

Rozdział 6. Indeksy, skale i typologie

Operacjonalizacja i konceptualizacja większości zmiennych w badaniach społecznych wymagają dużo refleksji i pomysłowości. Weź na przykład pod uwagę postawy wobec feminizmu. Konceptualizując zmienną *feminizm*, moglibyśmy zdecydować się na ograniczenie naszych zainteresowań do „politycznej równości mężczyzn i kobiet". Moglibyśmy zoperacjonalizować tę zmienną na wiele sposobów, na przykład pytając respondentów, czy zgadzają się lub nie zgadzają z każdym z następujących stwierdzeń:

1. Chciałbym, żeby kobieta została prezydentem.
2. Chętnie głosowałbym na kompetentną kobietę kandydującą do Kongresu.
3. Szczerze mówiąc, nie jestem pewien, czy kobiety mają niezbędne cechy fizyczne i psychiczne, by zajmować stanowiska polityczne.
4. Ogólnie rzecz biorąc, byłoby zapewne lepiej dla wszystkich, gdyby kobiety zostawiły politykę mężczyznom.

Każde z powyższych stwierdzeń jest wyrazem pewnej postawy wobec obecności kobiet w polityce. Pierwsze dwa sformułowane są pozytywnie (popierają udział kobiet w polityce), kolejne dwa – negatywnie. Moglibyśmy przeprowadzić operacjonalizację na podstawie reakcji respondenta na te cztery zdania, przyznając na przykład po jednym punkcie za każdą odpowiedź feministyczną. W ten sposób respondent, który zgodziłby się ze zdaniami (1) i (2), a nie zgodziłby się z (3) i (4), uzyskałby cztery punkty. Inny respondent, który nie zgodziłby się ze wszystkimi czterema zdaniami, otrzymałby dwa punkty.

Kombinacja kilku wskaźników postaw wobec feminizmu stanowiłaby *indeks*. Moglibyśmy też stworzyć *skalę*, która uwzględnia fakt, że niektóre wskaźniki zmiennej mają bardziej skrajny charakter niż inne. Sprzeciw wobec przyznania kobietom prawa głosu byłby na przykład bardzo silnym wskaźnikiem antyfeminizmu. W rozdziale 6 zbadamy kilka typów skal.

Rozdział 7. Logika doboru próby

Podstawowym elementem badań społecznych, który odgrywa na ogół znacznie mniejszą rolę w naukach przyrodniczych, jest wybór niektórych obserwacji spośród nieskończonej liczby tych, które możemy przeprowadzić. Skoro możemy obserwować wszystko, jak wybrać przedmiot obserwacji?

Kiedy przyrodnik postanawia zbadać właściwości tlenu, nie jest w stanie zbadać *całego* tlenu, mimo że pragnie wyciągnąć wnioski co do tlenu jako takiego. Na tej samej zasadzie politolog, który chce sformułować wnioski dotyczące preferencji wyborczych głosujących, nie może przepytać *wszystkich* głosujących. Jak więc każdy z tych uczonych decyduje konkretnie, co będzie badać?

W przypadku tlenu każda cząsteczka jest równie dobra. Fizyk wysyła po prostu asystenta do sklepu, w którym się na ogół zaopatruje, żeby kupił litr tlenu. Nikogo nie obchodzi, *który* litr. Politolog badający wszystkich wyborców, których akurat ma pod ręką, będzie jednak w kłopocie, ponieważ nie wszyscy głosujący są tacy sami. Wyobraź sobie, że poszukujesz wyborców w hipermarkecie znanej sieci, na zebraniu związkowym albo na walnym zgromadzeniu akcjonariuszy. Wybór grupy wyborców ma tu duże znaczenie.

Na szczęście można wybrać stosunkowo niedużą próbę wyborców w taki sposób, by ich odpowiedzi były reprezentatywne dla zamiarów dziesiątek milionów wyborców w całym kraju. Jeden z głównych zarzutów wobec sondaży wyborczych dotyczy faktu, że „psują one całą zaba-

wę z wyborów", informując nas o wynikach jeszcze przed oddaniem głosów. Z tych samych powodów niektórzy obawiają się, że wyniki prognoz wyborczych mogą mieć wpływ na to, jak ludzie zagłosują i czy w ogóle pójdą do wyborów.

Rozdział 7 wyjaśnia szczegółowo najprostsze techniki prostego doboru losowego i doboru systematycznego. Omawia on także bardziej złożone formy wielostopniowego doboru warstwowego, stosowanego w dużych sondażach ogólnokrajowych. Nawet jeśli nigdy nie będziesz prowadził badań sondażowych, będziesz w stanie zrozumieć i ocenić opisy technik doboru próby i błędu statystycznego zamieszczane w artykułach z badań czy też w prasowych omówieniach sondaży.

Nawet jeśli nie możemy wyłonić próby losowej z danej populacji, możemy i tak posługiwać się logiką doboru próby. W badaniach terenowych dotyczących gangów młodzieżowych istotna byłaby obserwacja i objęcie wywiadami wielu różnych członków wielu różnych gangów w celu uniknięcia zawężonego i stronniczego obrazu całości. Gdybyśmy zaś chcieli dojść do pogłębionego oglądu życia zawodowych polityków, wyniki naszych badań byłyby silnie uzależnione od tego, których polityków byśmy uwzględnili.

Gdybyśmy nawet chcieli przedstawić pełny opis pojedynczego przypadku (na przykład jednej gminy), bez intencji rozszerzenia zakresu naszych wyników na liczniejszą klasę grup czy jednostek, musielibyśmy dobrać próbę możliwych obserwacji danego przypadku. Rozważania na temat logiki doboru próby zawarte w rozdziale 7 rozwijają kwestię możliwości przemyślanego wyboru przedmiotu obserwacji w celu otrzymania właściwego obrazu obserwowanej rzeczywistości.

Rozdział 8. Eksperyment

Zapewne w wyobrażeniu większości laików na temat badań pojawiają się eksperymenty, na ogół prowadzone przez uczonych w długich, białych kitlach, czemu może – lecz nie musi – towarzyszyć demoniczny, pogardliwy wyraz twarzy czy błysk w oku. *Eksperyment kontrolowany* jest jedną z wielu technik stosowanych w badaniach społecznych. Bywa on często przeprowadzany w warunkach laboratoryjnych, mimo że badacze społeczni rzadko noszą białe kitle (czy też przybierają demoniczny wyraz twarzy). Eksperyment kontrolowany jest nie tylko powszechnie stosowaną techniką badawczą, lecz także praktycznym modelem wyjaśniającym logikę badania związków przyczynowych w naukach społecznych.

Badacze śledzą często relacje przyczynowe między zmiennymi. W języku zmiennych mówi się, że zmiany *zmiennych niezależnych* powodują zmiany *zmiennych zależnych*. Załóżmy na przykład, że chcemy się dowiedzieć, czy czytanie tej książki przez osiem godzin w noc poprzedzającą egzamin śródsemestralny spowoduje poprawę twojej oceny na egzaminie. Jak moglibyśmy się do tego zabrać?

Moglibyśmy zaprojektować prosty eksperyment, w którym niektórzy studenci czytaliby przez osiem godzin w noc przed egzaminem, inni natomiast oglądaliby telewizję. Następnie porównalibyśmy oceny w obu grupach i sprawdzili, czy ci, którzy się uczyli, wypadli lepiej. W tym przykładzie *uczenie się* i *nieuczenie się* to zmienne niezależne, *oceny* zaś byłyby zmienną zależną. Wyobraź sobie, że wynik jest następujący:

	Średnia liczba punktów
Uczył się	80
Oglądał telewizję	70

Z pewnością byłby to argument za przydatnością całonocnej nauki przed egzaminem, mógłbyś jednak zakwestionować wartość tego wniosku na kilka sposobów. Być może przed egzaminem uczyli się lepsi studenci, gorsi zaś oglądali telewizję. Ta właśnie różnica, nie zaś sam fakt uczenia się, mogła wpłynąć na różnice w ocenach. Przekonasz się, że metoda eksperymentalna posługuje się metodami doboru próby i dopasowywania, aby uchronić się przed nietrafnymi wynikami będącymi konsekwencją wadliwej konstrukcji samego eksperymentu, czyli inaczej *nietrafnością wewnętrzną*.

Zastanawiając się nad tym hipotetycznym eksperymentem, mógłbyś też zainteresować się, czy w takim przypadku, gdy studenci

wiedzą, że są obserwowani, wyniki eksperymentu mogą faktycznie odzwierciedlać to, co się dzieje w normalnym życiu. Jest to dobre pytanie – dotyczy ono problemu *nietrafności zewnętrznej*, którym zajmiemy się w rozdziale 8, wraz z kilkoma innymi kwestiami związanymi z nietrafnością wewnętrzną i zewnętrzną.

Stosunkowo niewielkie rozbieżności między średnimi wynikami w naszym hipotetycznym przykładzie – tylko dziesięć punktów – stanowią kolejny problem. Czy całonocna nauka istotnie w niewielkim tylko stopniu wpływa na wynik egzaminu? Czy tak niewielka zaobserwowana różnica mogłaby być tylko kwestią przypadku? Jeśli tak, ponowne przeprowadzenie eksperymentu mogłoby dać dokładnie odwrotne wyniki, zgodnie z którymi grupa oglądająca telewizję dostałaby lepsze oceny, niż grupa kująca przed egzaminem.

W dalszej części książki przyjrzymy się, jak analiza statystyczna dostarcza informacji wpływających na interpretację przeprowadzonych obserwacji. Przekonasz się, że nie wszystkie różnice między obserwowanymi grupami są znaczące – uwaga, którą sam możesz zastosować do niektórych sloganów reklamowych.

Rozdział 9. Badania sondażowe

Jedna z najpopularniejszych metod badawczych w naukach społecznych polega na posłużeniu się kwestionariuszem – w formie wywiadu prowadzonego przez ankietera lub za pośrednictwem poczty – do przebadania próby respondentów. Sondaże wyborcze i spisy rządowe są przykładami tego typu badań.

Centralnym elementem badań sondażowych jest standaryzowany kwestionariusz. *Standaryzowany kwestionariusz* gwarantuje, że w stosunku do każdego respondenta zostanie zastosowana dokładnie taka sama technika badawcza. Rozdział 5 zajmuje się technikami formułowania pytań, a rozdział 9 zawiera wskazówki w zakresie sposobu ich zadawania i konstrukcji kwestionariusza. Na przykład w sondażu dotyczącym postaw wobec feminizmu powinniśmy zapewne rejestrować płeć badanych, ponieważ jest prawdopodobne, że poglądy kobiet i mężczyzn w tej kwestii będą się nieco różnić. W tym przypadku konceptualizacja nie jest szczególnie trudna, operacjonalizacja również jest dość prosta. W kwestionariuszu moglibyśmy zoperacjonalizować ten pomiar w sposób następujący:

Płeć

- ☐ Kobieta
- ☐ Mężczyzna

Dobrym sposobem pomiaru wieku w kwestionariuszu jest na przykład takie pytanie:

Wiek: _____

Moglibyśmy też poprosić o zaznaczenie przybliżonego wieku respondenta w ramach kategorii w następujący sposób:

Wiek

- ☐ Poniżej 21 lat
- ☐ 21–34 lata
- ☐ 35–49 lat
- ☐ 50–64 lata
- ☐ Powyżej 65 lat

Rozdział 9 wyjaśnia także możliwe sposoby dystrybucji kwestionariusza: wywiad osobisty, telefoniczny, ankieta pocztowa lub inne formy kwestionariusza do samodzielnego wypełniania. Na koniec zajmiemy się zastosowaniem komputerów do dystrybucji kwestionariuszy oraz mocnymi i słabymi stronami badań sondażowych.

Rozdział 10. Jakościowe badania terenowe

Chyba najbardziej naturalna technika prowadzenia badań społecznych przypomina reportaż prasowy: po prostu idziemy w miejsce, gdzie się coś dzieje, i obserwujemy. Metoda ta jest jednak dobrze ugruntowana w badaniach społecznych i nazywa się *badaniami terenowymi*. Jest to technika jakościowa, co oznacza, że nie wymaga zastosowania analizy statystycznej.

W jednym z wariantów tej techniki, zwanym *obserwacją uczestniczącą*, badacz faktycznie włącza się w obserwowane zdarzenia i analizu-

je je od wewnątrz. Załóżmy, że interesują cię dyskusje polityczne wśród punk-rockowców w Los Angeles. W takim przypadku mógłbyś odkopać swoją gitarę i zejść pod ziemię[1].

W innych przypadkach badacze terenowi ujawniają się jako badacze. Oprócz obserwacji mogą również przeprowadzić wywiady pogłębione z niektórymi uczestnikami. Przekonasz się, że oba te podejścia mają swoje wady i zalety.

Poza tymi i im podobnymi technikami, postrzeganie świata przez badacza także wpływa na sposób przeprowadzania badań terenowych. Badacze terenowi, posługujący się metodami jakościowymi, w starannie przemyślany sposób formułują wiele paradygmatów, wzbogacających obserwację życia społecznego. Omówimy niektóre z nich: naturalizm, etnometodologię, teorię ugruntowaną, studium przypadku i metodę rozszerzonych przypadków, etnografię instytucjonalną i badania uczestniczące.

Niezależnie od osobistego punktu widzenia wszyscy prowadzimy w życiu codziennym pewnego rodzaju badania terenowe: obserwujemy, notujemy, analizujemy obserwacje. W rozdziale 10 opisane zostały szczegółowe techniki wykonywania tego typu czynności, wypracowane przez badaczy terenowych. W rozdziale tym omawiam również kilka najciekawszych przykładów badań społecznych, przeprowadzonych tą metodą.

Rozdział 11. Badania niereaktywne

Większość badań społecznych wymaga jakiegoś rodzaju interakcji z badanymi ludźmi: zadawania im pytań w sondażu, dawania im zadań do wykonania w eksperymencie czy też spotykania się z nimi w badaniach terenowych. Tego typu ingerencji w bieg życia społecznego towarzyszy pewne stałe niebezpieczeństwo: sam fakt badania może wpływać na przedmiot badań. Pytając ludzi, na kogo mają zamiar głosować, prawdopodobnie wpływamy w pewnym stopniu na ich zamiary i być może wzmacniamy ich stanowisko poprzez wymuszenie głośnej deklaracji.

Wiele technik badawczych można określić jako *niereaktywne*, dlatego że nie wymagają wywierania wpływu na przedmiot obserwacji. W rozdziale 11 omawiam kilka tego typu technik. Badacze społeczni mogą na przykład badać wiele zjawisk poprzez analizę istniejących danych statystycznych. W swoim nowatorskim badaniu Émile Durkheim, wielki socjolog francuski, poszukiwał przyczyn samobójstw, analizując i porównując wskaźniki samobójstw w różnych krajach i różnych regionach. Dokumenty historyczne mogą również dostarczyć danych dla analizy, jak w klasycznym studium Maxa Webera na temat związku między ascetyczną filozofią kalwinizmu a powstaniem kapitalizmu. Inny przykład badań niereaktywnych stanowi *analiza treści*, która dotyczy badania zarejestrowanych przekazów – od książek, przemówień, wierszy po obrazy, piosenki i reklamy telewizyjne. Rozdział 11 po raz kolejny dowodzi, że możliwości badań społecznych są praktycznie nieskończone.

Rozdział 12. Badania ewaluacyjne

Znaczna część badań społecznych poświęcona jest ewaluacji programów interwencji społecznej. Załóżmy, że ustawodawca stanowy przeznacza kilka milionów dolarów na wprowadzenie nowego programu resocjalizacji skazanych w więzieniu stanowym. Miło byłoby wiedzieć, czy program ostatecznie osiągnął zamierzony skutek: czy ktoś się zresocjalizował? Ewaluacja jest dla badaczy społecznych jedną z istotnych możliwości działania na rzecz społeczeństwa.

Gdybyś miał dokonać ewaluacji więziennego programu resocjalizacyjnego, musiałbyś zacząć od sprecyzowania konkretnych celów programu. Co rozumie się tu pod pojęciem „resocjalizacja"? Jak określić, czy więzień się zresocjalizował, czy nie? Osiągnąwszy porozumienie w tej kwestii, zechciałbyś prawdopodobnie określić „normalny" wskaźnik resocjalizacji, by mieć do czego porównać oddziaływanie nowego programu. Następnie mógłbyś zaplanować

[1] W oryginale „underground" – określenie muzycznych środowisk alternatywnych (przyp. tłum.).

ewaluację programu resocjalizacji w dużej mierze zgodnie z logiką eksperymentu kontrolowanego. Jak to wyjaśniam w rozdziale 12, ewaluacja nie jest jednak nigdy tak nienaganna logicznie, jak mogłoby się wydawać. Osoby prowadzące badania ewaluacyjne muszą często radzić sobie z bardziej złożonymi kwestiami etycznymi, politycznymi i administracyjnymi, niż się to zdarza w innych typach badań.

Rozdział 13. Analiza danych jakościowych

W rozdziałach 13–17 tej książki zajmuję się *analizą danych*, czyli procesem objaśniania tego, co obserwujemy. Znaczna część tych rozważań poświęcona jest prostym analizom statystycznym, jednak rozdział 13 dowodzi, że nie zawsze potrzebujemy liczb, by odkryć i zrozumieć prawidłowości w życiu społecznym. Niekiedy możemy zastosować tak zwaną *analizę danych jakościowych*.

Analiza danych jakościowych okazała się owocna w badaniach interakcji bezpośrednich w skali mikro, na przykład procesów decyzyjnych w gangu młodzieżowym. Jest również skuteczna w odkrywaniu prawidłowości zmian społecznych w całym społeczeństwie, a nawet w badaniu relacji między społecznościami.

Znaczna część rozdziału 13 poświęcona jest logice i własnościom analizy danych jakościowych: jakiego typu prawidłowości szukamy, jakich używamy technik, skąd wiemy, że znaleźliśmy coś istotnego. Większość analiz danych jakościowych prowadzonych w naukach społecznych wymaga dość skromnego wyposażenia technicznego: często tylko maszyny do pisania, kopiarki i nożyczek. Pod koniec prezentacji tego tematu omówimy jednak dokładniej możliwości edytorów tekstu oraz niektórych programów komputerowych zaprojektowanych specjalnie do analizy danych jakościowych: NUD*IST, Atlas/ti i Ethnograph.

Badacze wykazali się wielką pomysłowością, projektując badania jakościowe wnikające w istotne aspekty życia społecznego. Myślę, że po przeczytaniu tego rozdziału nabierzesz większego szacunku dla wyobraźni socjologicznej.

Rozdział 14. Kwantyfikacja danych

Dotychczas zajmowaliśmy się tym, jak zabrać się do dokonywania obserwacji i pomiarów dla różnych celów. Niekiedy skutkiem gromadzenia faktów w drodze obserwacji jest stworzenie zbiorów danych nadających się do analizy ilościowej (czyli liczbowej). Współcześnie oznacza to na ogół tworzenie zbiorów danych nadających się do obróbki statystycznej za pomocą komputera. W rozdziale 14 zajmiemy się wpływem rozwoju komputerów na badania społeczne. Następnie zapoznamy się ze sposobem przekształcania danych na surowiec do analizy komputerowej.

Komputer – zwłaszcza komputer osobisty – stał się dla nauk społecznych tym, czym był mikroskop dla biologii, a teleskop dla astronomii. Komputer pozwala badaczowi analizować zbiór danych społecznych z wielu perspektyw, umożliwia mu koncentrację na poszczególnych podzbiorach, testowanie procesów przyczynowych itp. To jakby obracać w dłoni część społeczeństwa, by przyjrzeć się jej z różnych punktów widzenia.

Rozdział 15. Podstawy analizy ilościowej

Skoro dane zostały już ujęte w formie nadającej się do analizy ilościowej, można je zacząć analizować. Rozpoczniemy nasze omówienie analizy danych ilościowych od prezentacji prawidłowości opisowych i związków przyczynowych między zmiennymi. Sprowadza się to ostatecznie do prawidłowości w rozkładach wartości jednej zmiennej lub większej ich liczby.

W najprostszej formie analizy – *analizie jednozmiennowej* – badamy rozkład wartości jednej zmiennej. Na przykład w Generalnym Sondażu Społecznym (GSS) w 1998 roku zadano próbie dorosłych Amerykanów pytanie, czy uważają swoje poglądy polityczne za liberalne, konserwatywne czy umiarkowane. Oto odpowiedzi:

2%	Skrajnie liberalne
13%	Liberalne
13%	Raczej liberalne
35%	Umiarkowane
15%	Raczej konserwatywne
15%	Konserwatywne
3%	Skrajnie konserwatywne

Możemy jeszcze uprościć tę jednozmiennową analizę, jeśli zauważymy, że 28% uznało się za liberałów (suma pierwszych trzech kategorii), 33% zaś za konserwatystów. (Tego typu zestawienia procentowe nie zawsze sumują się do równych 100%, ze względu na zaokrąglenia). Możesz uznać te dane za interesujące same w sobie, być może ciekawsza byłaby jednak informacja, jak określają się różne typy ludzi. Oto przykład dla płci:

	Poglądy polityczne	
	mężczyźni	kobiety
Liberalne	27%	31%
Umiarkowane	35%	38%
Konserwatywne	38%	32%

Powyższe dane wskazują na ciekawe różnice między płciami. Badane kobiety mieściły się na ogół symetrycznie po obu stronach środka, natomiast mężczyźni częściej bywali konserwatywni. Oto przykład *analizy dwuzmiennowej*, w której staramy się odkryć zależności między dwiema zmiennymi – w tym przypadku *płcią* i *poglądami politycznymi.*

Jaka zmienna z największym prawdopodobieństwem jest powiązana z posiadaniem określonych poglądów politycznych? Może tu przyjść do głowy na przykład identyfikacja z partią polityczną. W ostatnich dziesięcioleciach Partia Republikańska była na ogół bardziej konserwatywna niż Partia Demokratyczna. Sprawdźmy, czy prawidłowość ta występuje w próbie GSS. Weźmiemy też pod uwagę osoby określające się jako „niezależne".

	Poglądy polityczne		
	demokraci	niezależni	republikanie
Liberalne	39%	19%	11%
Umiarkowane	39%	39%	21%
Konserwatywne	22%	42%	68%

Poglądy polityczne demokratów, republikanów i osób niezależnych odzwierciedlają ogólną prawidłowość, której wystąpienia oczekiwaliśmy, mimo że różnice nie są tak duże, jak można by się spodziewać. Widać zwłaszcza, że demokraci występują w całym spektrum politycznym, republikanie natomiast skupiają się raczej po stronie konserwatywnej.

Rozdział 15 zajmuje się dość szczegółowo analizą jedno- i dwuzmiennową. Wprowadza także do logiki *analizy wielozmiennowej*, w której badamy jednocześnie kilka zmiennych. Zagadnienie to rozwijane jest w rozdziałach 16 i 17.

Rozdział 16. Model analizy rozbudowanej

Analiza zależności wielozmiennowych w naukach społecznych posługuje się modelem logicznym nazywanym *modelem analizy rozbudowanej*. Model ten wymaga wprowadzenia zmiennej kontrolnej jako mechanizmu wyjaśniania związku stwierdzonego między dwiema innymi zmiennymi.

Model analizy rozbudowanej bywa niekiedy wykorzystywany do odkrywania mechanizmów wzajemnych powiązań między dwiema zmiennymi. Na przykład dzieci z rozbitych rodzin częściej stają się młodocianymi przestępcami, niż dzieci mieszkające z matką i ojcem. Model analizy rozbudowanej pozwala nam odkryć przyczynę tego stanu rzeczy: różnica w poziomie nadzoru ze strony osób dorosłych. Dzieci lepiej nadzorowane rzadziej stają się przestępcami, dzieci zaś z rozbitych domów częściej pozbawione są nadzoru. Jeśli jednak zapewni się im właściwy nadzór, nie zostają przestępcami częściej niż inne dzieci.

Model analizy rozbudowanej prowadzi niekiedy do wykrycia pozornych zależności, niekiedy pozwala też odkryć szczególne warunki, w których dana zależność występuje. A czasami jedynie potwierdza możliwość uogólniania zaobserwowanych zależności. Na przykład zależność między preferowaną partią polityczną a wyznawaną ideologią polityczną występuje u osób obu płci, jak wskazują następujące dane:

	Poglądy polityczne: mężczyźni		
	demokraci	niezależni	republikanie
Liberalne	38%	19%	12%
Umiarkowane	38%	34%	22%
Konserwatywne	23%	47%	66%

	Poglądy polityczne: kobiety		
	demokraci	niezależni	republikanie
Liberalne	40%	19%	10%
Umiarkowane	39%	44%	20%
Konserwatywne	21%	37%	71%

Prezentacja modelu analizy rozbudowanej uzmysłowi ci, pod jak wieloma względami badania społeczne przypominają pracę detektywa, gdyż usiłujemy w nich wyśledzić znaczące prawidłowości tkwiące w naszym życiu społecznym.

Rozdział 17. Statystyka w naukach społecznych

Zaprezentowane przeze mnie proste tabele procentowe są przykładem analizy statystycznej danych ilościowych. Mam dużą słabość do tego typu analiz, jako że większość ludzi jest w stanie zrozumieć znaczenie procentowych opisów i porównań. Niekiedy jednak zdarza się, że zależności między zmiennymi nie dają się w tak prosty sposób zobrazować. Przedstawiciele nauk społecznych stworzyli więc złożone techniki analityczne, służące badaniu i ujawnianiu tego typu zależności. Rozdział 17 zajmuje się logiką niektórych z tych technik analitycznych.

Zaawansowane techniki statystyczne wykorzystywane w naukach społecznych są na tyle złożone, że przed ich skutecznym zastosowaniem trzeba je dość dogłębnie studiować. Prezentacja podstawowych reguł zastosowania niektórych technik, na które możesz się natknąć w czasopismach naukowych, pozwoli ci jednak zrozumieć wyniki tego typu analiz, mimo że nie będziesz przeprowadzać ich osobiście. Gdybyś jednak chciał zająć się tymi technikami nieco dokładniej, rozdział 17 daje ci po temu dobre podstawy.

Rozdział 18. Etyka i polityka w badaniach społecznych

Książka przedstawia logikę i tajniki prowadzenia badań społecznych, zarówno z sposób wyidealizowany – czyli tak, jak gdyby badania były prowadzone w idealnych warunkach, jak i w kontekście przeszkód, z którymi musi sobie radzić badacz w świecie realnym. Badacze muszą rozumieć zasady logiczne, leżące u podstaw stosowania tych technik, zarazem jednak muszą wiedzieć, co zrobić, gdy nie ma listy, z której można dokonać losowania, gdy ludzie nie odsyłają kwestionariuszy itp.

Oprócz przeszkód natury *logistycznej*, na które natyka się badacz w świecie realnym, ma on do czynienia także z przeszkodami etycznymi i politycznymi. Ze względów etycznych nie można na przykład prowadzić badań, które mogą wyrządzić krzywdę badanym. Większość z nas zna przerażające przykłady lekceważenia tej reguły przez nazistowskich lekarzy podczas II wojny światowej. Badania społeczne, mimo że rzadko zagrażają życiu badanych, mogą wywołać u nich zażenowanie, zmartwienie, zwątpienie i inne formy dyskomfortu psychicznego. Co więcej, publiczne ujawnienie wyników ich dotyczących może narazić respondentów na nieprzyjemności. W rozdziale 18 omawiam wiele ścisłych reguł etyki badawczej, wypracowanych przez badaczy społecznych. Wyczerpujące rozważania na temat etyki w tym rozdziale są bardzo istotne, jako że omawiane zasady dotyczą każdego zagadnienia przedstawionego wcześniej – a będziesz już wówczas znał dogłębnie każdy temat.

Przeszkody polityczne także mają wpływ na praktykę badań społecznych. Niejednokrotnie zdarza się, że przeszkody te są w stanie powstrzymać badacza. Wyobraź sobie na przykład, że prowadzisz badania na temat uczciwości i efektywności urzędników państwowych w czasie kampanii wyborczej. Najprawdopodobniej nie uda ci się swobodnie działać wyłącznie na podstawie zasad naukowych. Niektórzy kandydaci polityczni mogliby odmówić udziału w badaniach czy wręcz podjąć próby ich zablokowania. Podobnie badania dotyczące drażliwych tematów, jak wychowanie seksualne, homoseksualizm, aborcja, AIDS są często prowadzone w kontekście sporu publicznego. Badania ewaluacyjne, mające na celu sprawdzenie skuteczności jakiegoś programu, muszą na ogół brać pod uwagę osobiste i zawodowe interesy osób odpowiedzialnych za ten program.

Mimo że także badania przyrodnicze muszą się liczyć z przeszkodami natury politycznej i etycznej, specyfika przedmiotu nauk społecznych sprawia, że przeszkody takie występują w nich częściej. Uwrażliwienie na etykę i polity-

kę w badaniach społecznych pomoże ci radzić sobie z tego typu realnymi czynnikami i, w przypadku etyki, ustrzec się przed nieświadomym łamaniem przyjętych zasad etycznego postępowania w twoich praktykach badawczych.

Rozdział 19. Zastosowania badań społecznych

Kończę tę książkę przypuszczeniem, że być może nie planujesz wcale kariery badacza społecznego, i dobrze. Gdyby każdy, kto przeszedł kurs metod badań społecznych, zamierzał być zawodowym badaczem, większość ludzi takich jak ja byłaby bezrobotna.

Książka ta może jednak zainteresować również ludzi nie zajmujących się zawodowo badaniami społecznymi. Po pierwsze przez całą resztę życia będziesz czytał i słyszał o badaniach społecznych – zrozumienie ich podstawowych zasad stanowi element Cywilizacyjnej Kamizelki Ratunkowej. Wielokrotnie w tej książce będziemy się zastanawiać, jak możesz wykorzystać to, czego się nauczyłeś, by stać się bardziej efektywnym konsumentem wyników badań społecznych.

Po drugie znaczna część tej książki stosuje się również do obszarów twojego życia nie mających związku z badaniami społecznymi. Jeśli nauczysz się na przykład budować kwestionariusze, możesz wykorzystać tę wiedzę równie dobrze przy opracowywaniu innego typu dokumentów. Umiejętność prowadzenia obserwacji uczestniczącej przyda ci się w zawodzie dziennikarza prasowego, przedsiębiorcy, detektywa czy szpiega (choć ten ostatni przypadek stwarza pewne problemy etyczne).

Rozdział 19 stanowi podsumowanie całej książki, na wypadek, gdyby kazano ci coś o niej powiedzieć na egzaminie (przy okazji: powodzenia i nie dziękuj!).

To koniec hologramu 4. Reszta książki jest kolejną holograficzną prezentacją badań społecznych. Mam nadzieję, że zajrzysz tu od czasu do czasu, aby utrzymać wszystkie szczegóły we właściwych proporcjach. Przede wszystkim jednak mam nadzieję, że nauka o badaniach społecznych będzie dla ciebie równie wielką radością jak dla mnie.

Część
1

Wprowadzenie do badań

1. Działalność badawcza człowieka a nauka

2. Paradygmaty, teoria i badania społeczne

3. Pojęcie przyczynowości w badaniach społecznych

Nauka jest słowem znanym każdemu i powszechnie używanym. Jednak jego rozumienie bywa bardzo różne. Dla jednych nauka to matematyka, dla innych – białe fartuchy i laboratoria. Często myli się ją z techniką lub utożsamia z seminariami uniwersyteckimi.

Nauka nie jest oczywiście żadną z tych rzeczy *per se*. Trudno jednak precyzyjnie wyjaśnić, czym jest. Tak naprawdę naukowcy nie są zgodni w kwestii jej właściwej definicji. Dla celów tej książki wystarczy, jeśli spojrzymy na naukę jako na metodę badawczą – sposób, w jaki poznajemy otaczający nas świat. Nauka ma pewne cechy charakterystyczne, które odróżniają ją od innych metod poznawania świata. Jest przedsięwzięciem świadomym, przemyślanym i kierującym się rygorystycznymi zasadami. Czasem, choć nie zawsze, korzysta z analiz statystycznych. W tych wstępnych rozdziałach przyjrzymy się tym i pozostałym jej cechom.

Dr Benjamin Spock, słynny autor i pediatra, zaczyna swoje książki od zapewnienia młodych rodziców, że wiedzą więcej o opiece nad dziećmi, niż im się wydaje. Chciałbym wyjść od podobnego stwierdzenia. Szybko uświadomisz sobie, że znasz już całkiem dobrze praktykę badań społecznych, że w istocie przez całe życie prowadziłeś takie badania. Z tej perspektywy, celem książki jest pomoc w doskonaleniu posiadanych przez ciebie umiejętności i pokazanie pewnych sztuczek, które nie przyszły ci może wcześniej do głowy.

Pierwsza część książki kładzie fundament pod pozostałe. Rozważam w niej, co odróżnia naukę od innych form poznawania świata: jej podstawowe cechy oraz główne zagadnienia, które podejmuje. W rozdziale 1 zaczniemy od omówienia dociekań naturalnych dla każdego człowieka, takich, które prowadziłeś przez całe życie. W toku rozważań pokazane będą niektóre fałszywe tropy, jakimi ludzie podążają próbując zrozumieć otaczający ich świat, w podsumowaniu zaś – zasadnicze cechy badania naukowego pozwalające ustrzec się przed popełnianiem tego rodzaju błędów.

Rozdział 2 traktuje o teoriach społecznych i powiązaniach między teorią a badaniami. Przyjrzymy się kilku paradygmatom teoretycznym, które kształtują naturę badań i w ogromnym stopniu określają ich cele oraz interpretację wyników.

Dążąc do uogólnionego rozumienia, naukowcy chcą odkryć prawidłowości we wzajemnych relacjach między zmiennymi. Relacje te często przybierają postać zależności przyczynowo-skutkowych. Rozdział 3 odnosi się do związków przyczynowych, ich natury i logiki w przypadku badań społecznych. Ten teoretyczny rozdział stanowi też podstawę dla rozdziałów poświęconych technikom analitycznym.

Ogólnym celem pierwszej części jest stworzenie tła dla ukazania specyfiki planowania i przeprowadzania badań. Po jej przeczytaniu będziesz przygotowany do rozważenia niektórych bardziej szczegółowych aspektów badań społecznych. ■

Rozdział 1

Działalność badawcza człowieka a nauka

Przegląd holograficzny

Każdy z nas próbuje rozumieć i przewidywać zjawiska świata społecznego. Zadaniem nauki – a badań społecznych w szczególności – jest omijać pułapki, w które zazwyczaj wpada człowiek w swych dociekaniach.

Wprowadzenie

W poszukiwaniu rzeczywistości
Zwykła działalność badawcza człowieka
Tradycja
Autorytet
Błędy w badaniach oraz jak ich uniknąć
Co rzeczywiście jest rzeczywiste?

Podstawy nauk społecznych
Teoria zamiast filozofii czy wiary
Prawidłowości społeczne
Zbiorowości zamiast jednostek
Język zmiennych

Trochę dialektyki badań społecznych
Wyjaśnianie idiograficzne i nomotetyczne
Teoria indukcyjna i dedukcyjna
Dane jakościowe i ilościowe

Etyka w badaniach społecznych
Dobrowolne uczestnictwo
Nie krzywdzić badanych

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Oto książka o poznawaniu – nie tyle o tym, *co* wiemy, lecz w jaki *sposób* się tego dowiadujemy. Zacznijmy od przyjrzenia się temu, co prawdopodobnie wiesz już teraz.

Wiesz, że Ziemia jest okrągła. Możliwe, że wiesz również, iż po ciemnej stronie Księżyca jest zimno, a ludzie w Chinach mówią po chińsku. Wiesz, że witamina C zapobiega przeziębieniom, a seks bez zabezpieczenia może prowadzić do AIDS.

Skąd o tym wiesz? Jeśli nie byłeś ostatnio na Księżycu po jego ciemnej stronie ani nie prowadziłeś badań eksperymentalnych nad zaletami witaminy C, to wiesz o tym, ponieważ ktoś ci powiedział, a ty uwierzyłeś w to, co ci powiedziano. Być może o tym, że w Chinach ludzie mówią po chińsku, przeczytałeś w „National Geographic" i nie podałeś tego w wątpliwość, ponieważ wydawało się sensowne. Może to nauczyciel fizyki czy astronomii powiedział ci, że po ciemnej stronie Księżyca jest zimno, a może dowiedziałeś się o tym z „Newsweeka".

Niektóre ze znanych ci rzeczy wydają się całkiem oczywiste. Gdyby spytał cię ktoś, skąd wiesz, że Ziemia jest okrągła, prawdopodobnie rzekłbyś: „wszyscy to wiedzą". Mnóstwo jest rzeczy, które wszyscy wiedzą. Naturalnie, kiedyś wszyscy „wiedzieli", że Ziemia jest płaska.

Większość z tego, co wiemy, jest kwestią umowy i wiary. Niewiele opiera się na osobistym doświadczeniu i własnych odkryciach. W istocie dużą część dorastania w każdym społeczeństwie zajmuje proces, w trakcie którego uczymy się akceptować to, o czym wszyscy wokół „wiedzą". Jeśli nie wie się tych samych rzeczy, nie można tak naprawdę przynależeć do grupy. Gdybyś zamierzał poważnie zakwestionować kulistość Ziemi, szybko okazałoby się, że odróżniasz się od innych. Mógłbyś nawet zostać umieszczony w szpitalu, gdzie przebywałbyś z ludźmi kwestionującymi takie rzeczy.

Ważne jest, aby zdać sobie sprawę, że większość naszej wiedzy zależy od wiary w to, co nam powiedziano. Pod tym względem jednak wszystko jest z nami w najlepszym porządku. W ten sposób po prostu są zbudowane społeczeństwa i jest to naprawdę użyteczne. Umowa leży u podstaw wiedzy. Ponieważ nie sposób wyłącznie dzięki własnym odkryciom i doświadczeniu poznać wszystkiego, co wiedzieć należy, świat jest tak ułożony, abyśmy mogli po prostu uwierzyć w to, co mówią inni. Pewne rzeczy wiemy dzięki tradycji, inne – dzięki „ekspertom".

Istnieją jednak też odmienne metody poznawcze. W przeciwieństwie do wiedzy pochodzącej z umowy poznawać można również dzięki bezpośredniemu doświadczeniu – poprzez obserwację. Jeśli zanurkujesz w lodowatym strumieniu płynącym w Górach Skalistych, nikt nie musi ci mówić, że jest zimny; sam to zauważasz. Gdy pierwszy raz nadepnąłeś na cierń, wiedziałeś, że zabolało, zanim ktokolwiek ci o tym powiedział.

Gdy nasze własne doświadczenie przeczy temu, co wszyscy wiedzą, jest całkiem możliwe, że zrezygnujemy z własnego doświadczenia na rzecz wspomnianej umowy.

Zilustrujemy to na przykładzie. Wyobraź sobie, że przyszedłeś na przyjęcie do mojego domu. Jest to spotkanie osób z wyższych sfer; drinki i jedzenie są wyśmienite. Szczególnie interesująca wydaje ci się jedna z przekąsek, które roznoszę na tacy: panierowana, smażona, wyjątkowo ciekawa w smaku. Zjadłeś kilka – są takie pyszne! Jesz więcej. Chwilę potem dyskretnie przemykasz po pokoju, by znaleźć się wszędzie tam, gdzie zjawiam się z tymi smakołykami na tacy.

W końcu nie możesz się już powstrzymać. „Co to jest?" – pytasz. „Skąd mogę wziąć przepis?" A ja dopuszczam cię do tajemnicy: „Jadłeś panierowane smażone dżdżownice!" Twoja reakcja jest gwałtowna: żołądek buntuje się i siłą powstrzymujesz się od zwymiotowania wszystkiego na dywan w pokoju gościnnym. Ohyda! Jak można podać takie paskudztwo gościom!

Sens tej historii jest taki, że oba doznania dotyczące przekąski były równie prawdziwe. Upodobanie, jakie początkowo w niej znajdowałeś, płynące z twych bezpośrednich wrażeń, z całą pewnością było prawdziwe, podobnie jak uczucie obrzydzenia, którego doznałeś dowiedziawszy się, że jadłeś dżdżownice. Jednak powinno być jasne, że odraza, którą poczułeś, była w całości produktem umowy, jaką zawarłeś

z otoczeniem, a zgodnie z którą dżdżownice nie nadają się do jedzenia. To umowa, na którą przystałeś po raz pierwszy, gdy rodzice znaleźli cię na kupie śmieci, a z ust zwisała ci wijąca się połowa dżdżownicy. Kiedy oni rozwierali ci buzię i sięgali w głąb gardła w poszukiwaniu drugiej połowy, pojąłeś, że dżdżownice nie są pożywieniem akceptowanym w naszym społeczeństwie.

Abstrahując od tej umowy, cóż złego jest w dżdżownicach? Prawdopodobnie zawierają dużo białka, a mało kalorii. Mają rozmiary jednego kęsa i dają się łatwo zapakować – są więc marzeniem dystrybutora. Są także przysmakiem w społeczeństwach, którym obca jest nasza umowa uznająca robaki za odrażające. Niektórzy ludzie mogą nawet uwielbiać dżdżownice, ale odrzuca ich myśl o smażeniu w panierce.

Można zastanowić się nad następującym pytaniem: „Czy dżdżownice są «rzeczywiście» dobre, czy też «rzeczywiście» nie nadają się do jedzenia?". A oto i pytanie ciekawsze: *„Skąd można wiedzieć*, co naprawdę jest rzeczywiste?". Oto książka, która dotyczy pytań tego drugiego typu.

W pozostałej części tego rozdziału rozważymy, skąd wiemy, co jest rzeczywiste. Zaczniemy od spojrzenia na badanie jako na naturalną działalność człowieka, coś, czym zajmujemy się w naszym życiu na co dzień. Przyjrzymy się źródłom wiedzy potocznej i kilku rodzajom błędów popełnianych przy takim zwyczajnym badaniu. Potem zastanowimy się, czym pod tym względem odróżnia się nauka – a szczególnie nauki społeczne. Zastanowiwszy się nad niektórymi podstawowymi zasadami badań społecznych, zakończymy wstępnym omówieniem problemów pojawiających się w tych badaniach.

W poszukiwaniu rzeczywistości

Rzeczywistość jest podstępna. Zacząłeś już prawdopodobnie podejrzewać, że to, co się „wie" o niektórych sprawach, może nie być prawdą, ale skąd masz wiedzieć, co rzeczywiście jest rzeczywiste? Ludzie borykają się z tym problemem od tysięcy lat.

Jedną z odpowiedzi, powstałą w wyniku tych zmagań, jest nauka, która oferuje zbliżenie się zarówno do rzeczywistości umowy, jak i do rzeczywistości doświadczenia. Uczeni dysponują pewnymi kryteriami, które muszą być spełnione, aby mogli zaakceptować realność tego, czego osobiście nie doświadczyli. Ogólnie rzecz biorąc, twierdzenie naukowe musi mieć oparcie jednocześnie w logice i w empirii: musi być sensowne i nie może przeczyć faktycznej obserwacji. Dlaczego ziemscy naukowcy uznają twierdzenie, że po ciemnej stronie Księżyca jest zimno? Po pierwsze, ma ono sens, ponieważ ciepło powierzchni Księżyca pochodzi od promieni słonecznych; ciemna strona Księżyca jest ciemna, ponieważ jest „odwrócona tyłem" do Słońca. Po drugie, naukowe pomiary wykonane po ciemnej stronie Księżyca potwierdzają te logicznie uzasadnione przewidywania. A zatem naukowcy akceptują rzeczywistość zjawisk, które nie są przedmiotem ich osobistego doświadczenia – uznają rzeczywistość umowy – ale służą im do tego odpowiednie kryteria.

Jednakże – i tu zbliżamy się do celu tej książki – nauka oferuje nam również szczególne podejście do odkrywania rzeczywistości poprzez osobiste doświadczenie. Innymi słowy, proponuje nam szczególne podejście do działalności badawczej. *Epistemologia* jest nauką o poznaniu; *metodologię* (jedną z dziedzin epistemologii) można by nazwać nauką o dowiadywaniu się. Oto książka, która rozpatruje i prezentuje metodologię nauk społecznych, czyli to, w jaki sposób badacze społeczni dowiadują się różnych rzeczy o życiu społecznym człowieka.

Dlaczego potrzebujemy nauk społecznych, by odkryć rzeczywistość życia społecznego? Aby się tego dowiedzieć, rozważmy najpierw, jak przebiega zwykłe, nienaukowe badanie.

Zwykła działalność badawcza człowieka

Praktycznie rzecz biorąc, wszyscy ludzie, podobnie jak inne zwierzęta, wykazują potrzebę przewidywania swego przyszłego położenia. Ludzie wydają się predysponowani do tego, by podejmując się takiego zadania stosować rozu-

mowanie przyczynowo-skutkowe i probabilistyczne. Po pierwsze, na ogół uznajemy fakt, że przyszłe wydarzenia są w jakiś sposób spowodowane czy uwarunkowane przez obecne. Dowiadujemy się, że zdobycie wykształcenia wpłynie na ilość pieniędzy, które kiedyś będziemy zarabiać, a wypływanie poza rafę może doprowadzić do przykrego spotkania z rekinem. Rekinom z kolei zapewne wiadomo – niezależnie od tego, czy wnikliwie przeanalizowały tę kwestię, czy też nie – że jeśli będą szwendać się wokół rafy, może im się poszczęścić i spotkają nieszczęsnego pływaka.

Po drugie, ludzie, a najwyraźniej i pozostałe zwierzęta, uczą się, że takie przyczynowo-skutkowe modele mają charakter probabilistyczny: skutki pojawiają się częściej, gdy wystąpią przyczyny, a rzadziej, gdy tych przyczyn zabraknie. A zatem studenci uczą się, że ich trud zaowocuje lepszymi stopniami w większości przypadków, lecz nie za każdym razem. Uznajemy, że niebezpiecznie jest pływać poza rafą, choć jednocześnie nie wierzymy, że każda taka wyprawa musi się skończyć fatalnie. Jak zobaczymy w dalszych częściach książki, nauka czyni pojęcia przyczynowości i prawdopodobieństwa bardziej jednoznacznymi oraz dostarcza technik pozwalających radzić sobie z nimi w sposób bardziej rygorystyczny, niż czyni to na co dzień zwykły człowiek. Wyostrza narzędzia, którymi już wcześniej dysponowaliśmy, sprawiając, że stajemy się bardziej świadomi, zdyscyplinowani i kategoryczni w swoich badaniach.

Analizując zwykłą działalność badawczą człowieka, należy odróżnić przewidywanie od rozumienia. Często możemy przewidywać nie rozumiejąc – niektórzy przewidują deszcz na podstawie bólu w kolanie. I często, nawet jeżeli nie rozumiemy dlaczego, jesteśmy skłonni działać na podstawie wykazanych zdolności przewidywania. Miłośnik wyścigów, który zauważy, że w trzeciej gonitwie zawsze wygrywa koń na trzeciej pozycji, będzie przypuszczalnie stawiał na tego konia nie rozumiejąc, dlaczego tak się dzieje, ani nie dbając o to. Rzecz jasna, wada przewidywań nie opartych na rozumieniu stanie się oczywista, gdy wygra jeden z pozostałych koni, a nasz entuzjasta straci tygodniową wypłatę.

Niezależnie od tego, jakie prymitywne popędy czy instynkty kierują istotami ludzkimi i innymi zwierzętami, zaspokojenie ich zależy w ogromnym stopniu od zdolności przewidywania przyszłych wydarzeń. Jednak w przypadku ludzi próba przewidywania jest często umieszczona w kontekście wiedzy i rozumienia. Gdy można zrozumieć, dlaczego różne zjawiska są wzajemnie powiązane, dlaczego pojawiają się pewne prawidłowości, wówczas predykcja staje się dokładniejsza niż wtedy, gdy się te prawidłowości po prostu obserwuje i zapamiętuje. Ludzkie poszukiwania zatem mają na celu udzielenie odpowiedzi zarówno na pytanie „co", jak i „dlaczego"; dążymy do tego celu obserwując i starając się rozumieć.

Jak sugerowałem już w tym rozdziale, nasze próby poznania świata tylko częściowo opierają się na bezpośrednim, osobistym doświadczeniu czy badaniach. Inna, znacznie większa część, pochodzi z uzgodnionej wiedzy, którą otrzymujemy od innych; są to właśnie rzeczy, które „wszyscy wiedzą". Ta rzeczywistość umowy jednocześnie pomaga nam i przeszkadza w próbach poznania czegoś na własną rękę. Aby zobaczyć, jak to się dzieje, rozważmy dwa istotne źródła naszej wiedzy „z drugiej ręki" – tradycję i autorytet.

Tradycja

Każdy z nas dziedziczy jakąś kulturę stworzoną częściowo z powszechnie akceptowanej wiedzy o tym, jak funkcjonuje świat. Możemy nauczyć się od innych, że uprawa kukurydzy wiosną spotka się z największym wsparciem ze strony bogów, że jedzenie zbyt dużej ilości słodyczy spowoduje psucie się zębów, że obwód koła to, w przybliżeniu, dwadzieścia dwie siódme jego średnicy, albo że masturbacja prowadzi do ślepoty. Możemy na własnej skórze sprawdzać niektóre z tych „prawd", lecz większość z nich po prostu akceptujemy. Są to rzeczy, które „wszyscy wiedzą".

W tym sensie tradycja przynosi ludzkim poszukiwaniom wyraźne korzyści. Jeśli akceptujemy to, co wszyscy wiedzą, w naszym dążeniu do odkrywania prawidłowości i zrozumienia

oszczędzone nam zostaje przytłaczające zadanie rozpoczynania wszystkiego od zera. Wiedza ma charakter kumulatywny, a odziedziczony zbiór informacji i wyjaśnień jest trampoliną dla dalszego jej rozwoju. Często mówi się, że „stoimy na barkach olbrzymów", czyli poprzednich generacji.

Jednocześnie tradycja może być przeszkodą dla ludzkich poszukiwań. Jeśli szuka się świeżego podejścia do czegoś, co wszyscy już rozumieją i co zawsze rozumieli, można za swoje wysiłki zostać napiętnowanym i uznanym za głupca. Jednak bliższe rzeczywistości jest stwierdzenie, że większości z nas rzadko się zdarza poszukiwać odmiennego sposobu pojmowania tego, co w powszechnym odczuciu jest prawdą.

Autorytet

Pomimo siły tradycji nowa wiedza pojawia się każdego dnia. Niezależnie od naszych osobistych poszukiwań, korzystamy przez całe życie z nowych odkryć i wyjaśnień będących dziełem innych ludzi. Często akceptacja tych nowych dokonań zależy od statusu odkrywcy. Chętniej uwierzysz epidemiologowi, który oznajmia na przykład, że zwykłym przeziębieniem można zarazić się przy całowaniu, mniej chętnie – wujkowi Piotrowi.

Podobnie jak tradycja, autorytet może jednocześnie pomagać i przeszkadzać w działalności badawczej człowieka. Mądrze jest zaufać, szczególnie w obliczu kontrowersji, twierdzeniu osoby, która w danej dziedzinie jest odpowiednio wyszkolona, ma osiągnięcia i jest ekspertem. Jednocześnie jednak uznane autorytety mogą w znacznym stopniu utrudnić badanie, popełniając błędy w dziedzinie należącej do obszaru ich kompetencji. A przecież biologowie popełniają błędy w dziedzinie biologii. Poza tym wiedza biologiczna zmienia się w czasie.

Badania są także utrudnione, gdy uzależnieni jesteśmy od autorytetu ekspertów, którzy wypowiadają się na tematy nie należące do ich specjalizacji. Weźmy przykład politycznego czy religijnego przywódcy, który oznajmia, nie dysponując żadną wiedzą medyczną czy biochemiczną, że marihuana może usmażyć mózg. Przemysł reklamowy umiejętnie posługuje się tym nadużywaniem autorytetu, prezentując na przykład popularnych sportowców omawiających wartość odżywczą płatków śniadaniowych albo aktorów filmowych oceniających samochody.

Zarówno tradycja, jak i autorytet, gdy używamy ich zdobywając wiedzę o świecie, są jak obosieczny miecz. Ujmując to prościej, dają nam punkt wyjścia dla własnych badań, lecz mogą sprawić, że zaczniemy w złym miejscu, mogą też popchnąć nas w złym kierunku.

Błędy w badaniach oraz jak ich uniknąć

Całkiem niezależnie od potencjalnych zagrożeń związanych z tradycją i autorytetem potknięcia i upadki zdarzają się nam często, gdy zaczynamy uczyć się na własną rękę. Przyjrzyjmy się kilku powszechnym błędom, które popełniamy w naszych zwykłych badaniach, oraz sposobom, dzięki którym nauka zabezpiecza się przed tymi błędami.

Błędne obserwacje

Nierzadko w swoich obserwacjach popełniamy błędy. Na przykład, jak był ubrany twój wykładowca metodologii pierwszego dnia zajęć? Jeśli musisz zgadywać, wynika to stąd, że większość naszych codziennych obserwacji jest powierzchowna i półświadoma. Dlatego często nie zgadzamy się między sobą co do rzeczywistego przebiegu zdarzeń.

W przeciwieństwie do okazjonalnej działalności badawczej człowieka obserwacja naukowa jest działaniem świadomym. Już samo uczynienie obserwacji zajęciem bardziej przemyślanym pomaga ograniczyć błędy. Jeśli miałbyś zgadywać, w co był ubrany twój wykładowca pierwszego dnia zajęć, prawdopodobnie popełniłbyś błąd. Gdybyś jednak poszedł na te pierwsze zajęcia świadomie planując obserwację i zapamiętanie ubrania owego wykładowcy, miałbyś dużo większe szanse na bezbłędność.

W wielu przypadkach zarówno proste, jak i skomplikowane zalecenia dotyczące pomiaru

pozwalają ustrzec się przed błędnymi obserwacjami. Co więcej, dzięki nim stopień precyzji jest dużo wyższy, niż pozwalają na to niczym nie wspomagane ludzkie zmysły. Powiedzmy na przykład, że owego pierwszego dnia zajęć zrobiłbyś swojemu wykładowcy kolorowe zdjęcia.

Nieuprawnione uogólnianie

Gdy obserwując poszczególne zjawiska zachodzące wokół nas, poszukujemy w nich prawidłowości, często zakładamy, że kilka podobnych zdarzeń stanowi dowód istnienia ogólnego wzorca. Oznacza to, że pozwalamy sobie na nieuprawnione uogólnianie na podstawie ograniczonych obserwacji. (Przypomnijmy naszego całkiem już spłukanego miłośnika wyścigów konnych).

Skłonność do nieuprawnionych generalizacji jest prawdopodobnie większa przy pracy pod dużą presją, kiedy wymaga się od nas ogólnych wniosków. Lecz pojawia się również wtedy, gdy nie ma takiej presji. W każdym przypadku nieuprawnione uogólnianie może nadawać badaniom niewłaściwy kierunek lub je hamować.

Wyobraź sobie, że jesteś reporterem, który pisze o demonstracji w obronie praw zwierząt. Polecono ci dostarczyć artykuł już po dwóch godzinach, a musisz dowiedzieć się, dlaczego ludzie biorą udział w proteście. Spieszysz na miejsce wydarzeń, zaczynasz rozmawiać, pytasz demonstrantów o ich motywy. Jeśli pierwszych trzech, z którymi rozmawiasz, poda ci właściwie tę samą przyczynę, możesz najzwyczajniej założyć, że pozostałe 3000 przybyło tu z tego samego powodu. Niestety, gdy artykuł się ukazuje, twój wydawca otrzymuje mnóstwo listów od uczestników demonstracji, którzy brali w niej udział z zupełnie innych powodów.

Naukowcy zabezpieczają się przed ryzykiem nieuprawnionej generalizacji z odpowiednim wyprzedzeniem, zobowiązując się do zbadania odpowiednio dużej i reprezentatywnej próby. Innym sposobem jest **replikacja**, czyli powtarzanie badań. W gruncie rzeczy replikacja służy sprawdzeniu, czy każde kolejne badanie da te same wyniki. Potem, w ramach dalszego testowania, badanie może zostać powtórzone w nieco zmienionych warunkach.

Wybiórcza obserwacja

Jednym z niebezpieczeństw nieuprawnionych uogólnień jest to, że mogą prowadzić one do wybiórczej obserwacji. Gdy tylko dojdziemy do wniosku, że istnieje określona prawidłowość, oraz gdy opracujemy ogólne wyjaśnienie jej przyczyn, będziemy skłonni koncentrować się w przyszłości na tych wydarzeniach i sytuacjach, które pasują do naszego wzorca, a ignorować te, które nie pasują. Uprzedzenia rasowe i etniczne trwałość swą zawdzięczają w znacznej mierze wybiórczej obserwacji.

Czasami plan badań będzie z góry określał liczbę i rodzaj obserwacji, jakie mają być przeprowadzone, aby dały podstawę do sformułowania wniosków. Chcąc dowiedzieć się, czy kobiety w większym stopniu niż mężczyźni są gotowe popierać wolny wybór w sprawie aborcji, trzeba by zadeklarować określoną liczbę obserwacji dotyczących tego zagadnienia w danym planie badawczym. Można wyselekcjonować tysiąc starannie dobranych osób, z którymi przeprowadzone zostaną wywiady na ten temat. Albo, gdy chodzi o bezpośrednią obserwację jakiegoś wydarzenia, na przykład demonstracji obrońców praw zwierząt, badacze społeczni zadają sobie dodatkowy trud znalezienia „nietypowych przypadków" – takich, które nie pasują do ogólnego wzorca. Wnioskowanie, że jakiś młodzieniec został przestępcą głównie z powodu braku pozytywnych modeli ról osób dorosłych, zwraca uwagę na znaczenie, jakie mają modele ról społecznych dla kształtowania moralnie właściwych zachowań młodzieży.

Nielogiczne rozumowanie

W życiu codziennym istnieją jeszcze inne sposoby radzenia sobie z obserwacjami, które przeczą naszemu pojmowaniu tego, jak się rzeczy mają. Z pewnością jednym z najbardziej godnych uwagi wytworów ludzkiego umysłu jest „wyjątek, który potwierdza regułę". Myśl ta nie ma najmniejszego sensu. Wyjątek może pozwolić nam zwrócić uwagę na jakąś regułę czy na to, co wydaje się regułą, lecz w żadnym systemie logiki nie może on potwierdzić reguły, której zaprzecza. A jednak często korzystamy z tego

zwięzłego powiedzonka, gdy chcemy zmieść sprzeczności zadając prosty nielogiczny cios.

To, co statystycy nazywają *błędem hazardzisty*, jest kolejnym przykładem nielogicznego rozumowania w życiu codziennym. Często zakładamy, że przedłużająca się dobra lub zła passa zapowiada swoje przeciwieństwo. Wieczór pokerowych porażek może skłonić do myśli, że od zwycięskiej karty dzieli nas już tylko kilka chwil. Z powodu tego błędnego przekonania wielu pokerzystów pozostawało przy stole zdecydowanie za długo. Bywa też odwrotnie, przedłużający się okres dobrej pogody martwi nas, gdyż sądzimy, że w trakcie weekendowego pikniku na pewno już będzie padać.

Choć każdemu z nas czasem zdarza się zawstydzająco nielogiczne rozumowanie, naukowcy starają się uniknąć tej pułapki, stosując systemy logiczne w sposób świadomy i jednoznaczny. Logikę nauki rozważymy dokładniej w rozdziale 2. Na razie wystarczy stwierdzić, że logiczne rozumowanie jest dla naukowca świadomym działaniem, a stała aktywność innych uczonych zmusza do uczciwości.

A zatem nauka stara się chronić swe badania przed pułapkami zwykłej działalności badawczej. Dokładna obserwacja i rozumienie rzeczywistości nie są sprawą oczywistą ani trywialną. Są to w istocie problemy bardziej skomplikowane, niż dotąd sugerowałem.

Co rzeczywiście jest rzeczywiste?

Filozofowie, by opisać sposób, w jaki większość z nas funkcjonuje na co dzień, używają czasami określenia „naiwny realizm". Gdy zasiadasz przy stole do pisania, prawdopodobnie nie spędzasz wiele czasu zastanawiając się, czy stół naprawdę składa się z atomów, które to z kolei są na ogół pustą przestrzenią. Gdy wchodzisz na ulicę i widzisz pędzący na ciebie autobus, nie jest to najlepszy moment, by zastanawiać się nad metodami sprawdzania, czy autobus ten naprawdę istnieje. Wszyscy żyjemy w przekonaniu o oczywistości tego, co realnie istnieje – pogląd ten zazwyczaj pozwala nam radzić sobie na co dzień.

Nie chcę, by ta książka osłabiła czyjąś zdolność radzenia sobie z codziennym życiem. Mam jednak nadzieję, że powyższe argumenty wykazały, iż natura „rzeczywistości" jest prawdopodobnie bardziej złożona, niż jesteśmy skłonni zakładać w naszym codziennym funkcjonowaniu. Oto trzy ujęcia natury rzeczywistości, które zapewnią filozoficzne tło i poprzedzą dalsze rozważania o nauce. Są one czasem nazywane *premodernistycznym*, *modernistycznym* i *postmodernistycznym* ujęciem rzeczywistości (W. Anderson 1990).

Ujęcie premodernistyczne

Ten pogląd na rzeczywistość był przewodnikiem ludzkości w prawie całej jej historii. Nasi najdawniejsi przodkowie przyjmowali, że widzą rzeczy takimi, jakimi one w rzeczywistości są. Właściwie założenie to było tak fundamentalne, że nie uważano go nawet za założenie. Żadna jaskiniowa mama nie mówiła do swego jaskiniowego dziecka: „Nasze plemię wychodzi z założenia, że w Starym Pokręconym Drzewie przebywają złe duchy". Przeciwnie, mówiła: „TRZYMAJ SIĘ Z DALA OD TEGO DRZEWA, BO INACZEJ ZAMIENISZ SIĘ W ROPUCHĘ!".

Ludzie w toku ewolucji stawali się świadomi swej różnorodności i zaczęli dostrzegać, że inni nie zawsze podzielają ich poglądy. Może odkryli, że członkowie innego plemienia nie uwierzyli w historię o złym drzewie; właściwie to drugie plemię myślało, że duchy w drzewie są święte i pożyteczne. Odkrycie tej różnicy doprowadziło członków pierwszego plemienia do wniosku, że: „niektóre plemiona, których nazw nie podamy, są bardzo głupie". Dla nich bowiem drzewo wciąż było złe i oczekiwali, że niektórzy pomyleńcy wkrótce przeniosą się do Ropuszego Miasta.

Ujęcie modernistyczne

Ujęcie nazywane przez filozofów *modernistycznym* uznaje taką różnorodność za uprawnioną; jest to filozoficzna wersja powiedzenia „co kraj, to obyczaj". Osoba myśląca w duchu modernistycznym rzekłaby: „Uznaję duchy z tego drzewa za złe, ale wiem, że inni uznają je za dobre. Nikt z nas się nie myli ani nikt nie ma

racji. W drzewie po prostu są duchy. Nie są ani dobre, ani złe, a tylko różni ludzie mają o nich różne wyobrażenie".

Wielu z nas łatwo jest zaakceptować ten modernistyczny pogląd. Niektórzy mogą uznać mlecz za piękny kwiat, podczas gdy inni widzą w nim tylko denerwujące zielsko. Dla osób wyznających pogląd premodernistyczny mlecz musiał być jednym albo drugim. Jeśli myślisz, że to chwast, to rzeczywiście jest to chwast, choć możesz przyznać, że niektórzy ludzie mają wypaczone pojęcie piękna. W ujęciu modernistycznym mlecz to po prostu mlecz. To roślina z żółtymi płatkami i zielonymi liśćmi. Pojęcia „piękny kwiat" i „denerwujące zielsko" to subiektywne punkty widzenia narzucane tej roślinie przez różnych ludzi. Żaden z nich nie jest własnością samej rośliny. Podobnie „dobry" i „zły" były pojęciami narzucanymi duchom w drzewie.

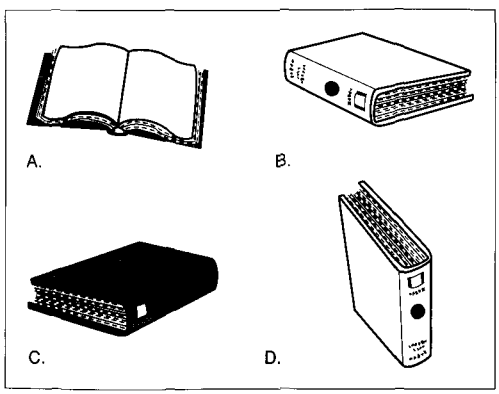
Ujęcie postmodernistyczne

Filozofowie coraz częściej mówią o *postmodernistycznym* ujmowaniu rzeczywistości. Zgodnie z nim duchy nie istnieją. Ani mlecz. „Rzeczywiste" są tylko wyobrażenia, które uzyskujemy dzięki naszym punktom widzenia. Innymi słowy, nic nie istnieje tam, na zewnątrz, wszystko jest tu, w środku. Jak Gertruda Stein powiedziała o mieście Oakland: „Nie ma tam żadnego tam".

Niezależnie od tego, jak dziwaczne może się na pierwszy rzut oka wydawać ujęcie postmodernistyczne, to jest ono w pewien paradoksalny sposób nieuniknione. Zastanów się przez chwilę nad książką, którą czytasz; w szczególności zwróć uwagę, jak ona wygląda. Skoro czytasz te słowa, książka prawdopodobnie wygląda jak rysunek 1.1A.

Rysunek 1.1

Książka



Ale czy rysunek 1.1A dobrze przedstawia, jak twoja książka „rzeczywiście" wygląda? A może jedynie oddaje to, jak książka wygląda z twojej obecnej perspektywy. Z pewnością ilustracje 1.1B, C oraz D przedstawiają ją równie trafnie. A przecież te spojrzenia na książkę tak się różnią między sobą. Które z nich jest „rzeczywistością"?

Jak pokazuje ten przykład, nie ma odpowiedzi na pytanie: „jak ta książka wygląda w rzeczywistości?". Wszystko, co możemy tu zaproponować, to różne wersje wyglądu książki, widzianej z różnych perspektyw. A zatem, zgodnie z ujęciem postmodernistycznym, nie istnieje „książka" jako taka, a jedynie różne jej wizerunki uchwycone z różnych stron. Wszystkie te wizerunki są tak samo „prawdziwe".

Zastosujemy teraz tę logikę do sytuacji społecznej. Wyobraźmy sobie spór między mężem a żoną. Gdy żona spogląda na kłócącego się męża, widzi to, co przedstawia rysunek 1.2. Zastanów się przez chwilę, co byś czuł i myślał będąc kobietą z tego rysunku. Jak wyjaśniłbyś później, co zaszło, swojej najlepszej przyjaciółce? Jaki sposób rozwiązania konfliktu wydawałby ci się odpowiedni, gdybyś był tą kobietą?

Oczywiście, jak pokazuje rysunek 1.3, zupełnie inną sprawą jest, co widzi mąż tej pani. Wyobraź sobie przez chwilę, że przeżywasz tę sytuację będąc na jego miejscu. Co byś czuł i o czym myślał? Jak opowiedziałbyś o tym zdarzeniu swemu najlepszemu przyjacielowi? Jakie rozwiązania wydawałyby ci się stosowne, by rozwiązać konflikt?

Rozważmy teraz trzeci punkt widzenia. Przypuśćmy, że jesteś obserwatorem i z zewnątrz przyglądasz się tej interakcji między mężem a żoną. Jak by to teraz wyglądało? Niestety, niełatwo jest przedstawić ten trzeci punkt widzenia, nie wiedząc nic o osobistych uczuciach, przekonaniach, wcześniejszych doświadczeniach itd., które wniósłby do danej sytuacji „zewnętrzny" obserwator. (Choć nazywam cię

Rysunek 1.2

Punkt widzenia żony



obserwatorem „zewnętrznym", obserwujesz oczywiście z wewnątrz twego własnego umysłu).

Biorąc najbardziej skrajny przykład, gdybyś był zdeklarowanym męskim szowinistą, widziałbyś prawdopodobnie tę walkę w sposób bardzo podobny do tego, jak widział ją mąż. Z kolei jeślibyś zgadzał się z poglądem, że mężczyźni są na ogół głupimi bumelantami, ujrzałbyś tę sprawę tak, jak żona widziała ją na wcześniejszym obrazku.

Wyobraź sobie jednak, że zamiast tego widzisz dwoje zdenerwowanych ludzi bezmyślnie się ze sobą kłócących. Czy uznałbyś ich za nieodpowiedzialnych głupków, w takim samym stopniu odpowiedzialnych za zaistniały konflikt? A może ujrzałbyś w nich dwoje ludzi, którzy znalazłszy się w trudnej dla człowieka sytuacji, próbują robić, co mogą, by jej zaradzić? Wyobraź sobie, że odczuwasz współczucie w stosunku do każdego z nich i że dostrzegasz, jak bardzo oboje starają się położyć kres wrogości, choć ze względu na wagę problemu nadal się kłócą.

Zauważmy, jak tych kilka ujęć różni się od siebie. Które z nich „prawdziwie" oddaje to, co dzieje się między mężem a żoną? Sęk w tym, by dostrzec, że osobisty punkt widzenia, który wnosimy jako obserwator, ponownie zabarwi nasze postrzeganie zdarzenia.

Ujęcie postmodernistyczne dobrze oddaje zasadniczy dylemat naukowców. Choć ich zadanie polega na obserwowaniu i rozumieniu tego, co się „rzeczywiście" dzieje, to przecież są wszyscy ludźmi i jako tacy wnoszą do badania swe osobiste przekonania, które zabarwiają obserwowane i wyjaśniane zdarzenia. Ostatecznie nie istnieje żaden sposób, w jaki ludzie mogliby całkowicie wykroczyć poza swoje człowieczeństwo, aby ujrzeć i zrozumieć świat takim, jakim on „rzeczywiście" jest – tzn. niezależnie od wszystkich ludzkich perspektyw.

O ile modernistyczne ujęcie uznaje nieuchronność ludzkiej subiektywności, spojrzenie postmodernistyczne sugeruje, że przede wszystkim nie ma tak naprawdę żadnej „obiektywnej" rzeczywistości, którą można by obserwować. Istnieje tylko kilka naszych subiektywnych perspektyw.

Możesz chcieć zastanowić się przez chwilę samodzielnie nad tymi trzema ujęciami rzeczywistości. Powrócimy do nich w rozdziale 2, kiedy skoncentrujemy się na bardziej konkretnych naukowych paradygmatach. W końcu dojdziemy do dwóch wniosków. Po pierwsze, ustalone procedury naukowe pozwalają nam czasami skutecznie radzić sobie z powyższym dylematem – tzn. możemy badać ludzkie zachowania i pomagać ludziom rozwiązywać problemy, jednocześnie nie będąc w stanie ujrzeć bezpośrednio „rzeczywistości". Po drugie, różne stanowiska filozoficzne dają liczne możliwości ustrukturyzowania naszych badań.

Rysunek 1.3

Punkt widzenia męża



Przejdźmy teraz od ogólnych idei filozoficznych do podstaw różnych szczegółowych sposobów wyjaśniania w naukach społecznych. Rozważenie tych fundamentów badań społecznych przygotuje nas do poznania poszczególnych technik badawczych.

Podstawy nauk społecznych

Nauka bywa czasem charakteryzowana jako logiczno-empiryczna. Ten niezgrabny termin niesie ze sobą istotny przekaz. Jak zauważyliśmy wcześniej, dwa filary nauki to logika i obserwacja. Znaczy to, że naukowe rozumienie świata musi jednocześnie być sensowne i zgodne z obserwacją. Oba te elementy mają zasadnicze znaczenie dla nauki i odnoszą się do trzech głównych aspektów działalności nauk społecznych: teorii, gromadzenia danych i ich analizy.

Upraszczając nieco tę kwestię, można powiedzieć, że **teoria** naukowa zajmuje się logiką, podczas gdy gromadzenie danych – obserwacją jako aspektem nauki. W trakcie analizy danych poszukuje się prawidłowości w obserwacjach, a tam gdzie to stosowne, porównuje się to, co zaobserwowane, z tym, co logicznie daje się przewidzieć. Choć książka ta przede wszystkim dotyczy gromadzenia danych i ich analizy – tzn. tego, jak przeprowadzać badania społeczne – pozostałe fragmenty 1 części poświęcone są teoretycznemu kontekstowi badań. Części 2 i 3 skupiają się na gromadzeniu danych, a w części 4 proponuje się wstęp do analizy danych.

U podstaw pojęć przedstawionych w tej książce leżą pewne fundamentalne idee, które odróżniają nauki społeczne – teorię, gromadzenie danych i ich analizę – od innych ujęć zjawisk społecznych. Zastanówmy się nad tymi ideami.

Teoria zamiast filozofii czy wiary

Teoria społeczna zajmuje się w dzisiejszych czasach tym, co jest, a nie tym, co być powinno, choć przez wiele setek lat nie rozróżniano tych dwóch orientacji. Filozofowie społeczni z dużą swobodą łączyli swoje obserwacje dotyczące tego, co wokół nich się działo, ze swoimi spekulacjami dotyczącymi przyczyn wydarzeń i swoimi pomysłami dotyczącymi tego, jak świat powinien wyglądać. Chociaż współcześni badacze społeczni mogą od czasu do czasu czynić to samo, jednak jako naukowcy skupiają się na tym, co faktycznie się dzieje i dlaczego.

Oznacza to, że teoria naukowa – a szerzej, także sama nauka – nie może rozstrzygać sporów o wartości. Nauka nie może określić, czy kapitalizm jest lepszy, czy gorszy od socjalizmu. Może natomiast określić, jak oba te systemy funkcjonują w kategoriach pewnego zestawu uzgodnionych kryteriów. Na przykład możemy naukowo określić, czy kapitalizm, czy socjalizm bardziej sprzyja ludzkiej godności i wolności tylko wtedy, gdy wcześniej uzgodnimy jakieś mierzalne definicje wolności i godności. Nasze wnioski będą zatem ograniczone do znaczeń określonych w naszych definicjach. Poza nimi nie będą miały żadnego ogólnego znaczenia.

Podobnie, jeśli zgodzimy się na to, że wskaźnik samobójstw lub, powiedzmy, datki na cele dobroczynne są dobrymi miernikami jakości religii, wówczas będziemy mogli naukowo określić, czy buddyzm, czy chrześcijaństwo jest lepszą religią. I ponownie nasze wnioski będą nierozerwalnie związane z wybranymi przez nas kryteriami. W praktyce ludzie rzadko uzgadniają precyzyjne kryteria przy rozwiązywaniu problemów dotyczących wartości, a zatem nauka rzadko przydaje się przy rozstrzyganiu takich debat.

Właściwie pytania tego typu są do tego stopnia sprawą opinii i wiary, że badania naukowe często postrzegane są jako zagrożenie dla tego, co „już wiadomo".

Rozważymy ten problem bardziej szczegółowo w rozdziale 12, gdy będziemy się przyglądać badaniom ewaluacyjnym. Jak zobaczymy, badacze w coraz większym stopniu zajmują się analizowaniem programów społecznych odzwierciedlających perspektywy ideologiczne, takich jak wyrównywanie szans grup dyskryminowanych czy reforma państwa opiekuńczego. Jeden z największych problemów, z jakimi się stykają, to uzgodnienie kryteriów sukcesu czy porażki danego programu. Kryteria takie mają

jednak zasadnicze znaczenie, jeżeli badania społeczne mają nam powiedzieć cokolwiek przydatnego w kwestiach dotyczących wartości. Analogicznie, stoper nie może powiedzieć nam, czy jeden sprinter jest lepszy od drugiego, jeżeli wcześniej nie umówimy się, że prędkość jest kryterium rozstrzygającym.

Nauki społeczne mogą nam zatem pomóc dowiedzieć się jedynie, co istnieje i dlaczego. Możemy je wykorzystać, by określić, jak powinno być, tylko jeśli ludzie wspólnie uzgodnią pewne kryteria pozwalające zdecydować, które wyniki są lepsze od innych – zgoda taka pojawia się rzadko.

Jak wskazałem wcześniej, nawet poznanie tego „co i dlaczego" nie jest zadaniem łatwym. Przejdźmy teraz do kilku zasadniczych idei, które leżą u podstaw wysiłków podejmowanych przez nauki społeczne, mających na celu opis i zrozumienie rzeczywistości społecznej.

Prawidłowości społeczne

Badania społeczne zmierzają w dużej części do odkrywania prawidłowości w życiu społecznym. Choć znalezienie prawidłowości jest dążeniem każdej nauki, cel ten niekiedy stanowi przeszkodę dla ludzi, którzy po raz pierwszy stykają się z naukami społecznymi.

Oczywiście na pierwszy rzut oka wydaje się, że zjawiska badane przez nauki przyrodnicze w większym stopniu podlegają prawom niż zjawiska społeczne. Ciężki przedmiot spada na ziemię zawsze, gdy go upuścimy, ale człowiek może w jednych wyborach zagłosować na jakiegoś kandydata, a w następnych przeciwko temu samemu kandydatowi. Podobnie lód zawsze topi się, gdy go wystarczająco podgrzejemy, natomiast ludzie zazwyczaj uczciwi czasem kradną. Pomimo tych przykładów zjawiska społeczne wykazują duży stopień regularności, którą można odkryć za pomocą badania i wyjaśnić za pomocą teorii.

Zacznijmy od tego, że ogromna ilość formalnych norm w społeczeństwie tworzy znaczące regularności. Na przykład zasady ruchu drogowego skłaniają ogromną większość ludzi do prowadzenia samochodów po prawej, a nie po lewej stronie drogi. Wymóg rejestracji głosujących prowadzi do powstania przewidywalnych wzorców, według których pewne grupy ludzi biorą udział w wyborach powszechnych. Prawo pracy wprowadza wysoki stopień uniformizacji, określając minimalny wiek pracownika oraz jego minimalną płacę. Takie zalecenia formalne regulują zachowania społeczne.

Obok zaleceń formalnych można obserwować inne normy społeczne, które tworzą kolejne prawidłowości. W grupie zarejestrowanych głosujących republikanie są w większym stopniu skłonni głosować na republikanów niż na demokratów. Profesorowie uniwersytetu zazwyczaj zarabiają więcej niż robotnicy niewykwalifikowani. Mężczyźni zazwyczaj zarabiają więcej niż kobiety.

W odniesieniu do takich prawidłowości społecznych stawiane są czasem trzy zarzuty. Po pierwsze, niektóre z nich mogą się wydawać trywialne. Na przykład wszyscy wiedzą, że republikanie głosują na republikanów. Po drugie, można podać przykłady, które przeczą danej „prawidłowości", a więc nie ma ona absolutnego charakteru. I po trzecie, można argumentować, że w przeciwieństwie do ciężkich przedmiotów, które, upuszczone, nie mogą zadecydować, że *nie* upadną, ludzie, których dana prawidłowość dotyczy, jeśli zechcą, są w stanie wszystko popsuć.

Zajmijmy się teraz każdym z tych zarzutów z osobna.

Zarzut trywialności

Podczas II wojny światowej Samuel Stouffer, jeden z największych badaczy społecznych, zorganizował w Armii Stanów Zjednoczonych zespół badawczy, którego prace miały na celu wesprzeć wojenny trud (Stouffer i in. 1949–1950). Wiele z przeprowadzonych analiz dotyczyło nastrojów panujących wśród żołnierzy. Stouffer i jego współpracownicy odkryli istnienie rozbudowanej „wiedzy potocznej" dotyczącej podstaw żołnierskiego morale. Duża część ich badań poświęcona była testowaniu tych „oczywistych" prawd.

Na przykład już dawno ludzie zorientowali się, że możliwość awansu wpływa na żołnierskiego ducha. Nastroje poprawiają się, gdy

żołnierze awansują i gdy system awansowania wydaje się sprawiedliwy. Co więcej, logiczne jest, że ludzie, którzy awansowali, będą skłonni myśleć, że system jest sprawiedliwy, ci zaś, których pominięto, będą raczej myśleć, że sprawiedliwy nie jest. Rozwijając ten tok rozumowania, wydaje się rozsądne, że żołnierze w jednostkach, w których awansuje się wolniej, będą skłonni sądzić, że system jest niesprawiedliwy, a ci z jednostek, w których awans przychodzi szybko, uznają system za sprawiedliwy. Ale czy właśnie w ten sposób myśleli o tym żołnierze?

Stouffer i jego współpracownicy skupili się na dwóch jednostkach: na Żandarmerii Wojskowej (ŻW), w której w porównaniu z całą armią awansowano najwolniej, oraz na Wojskowym Korpusie Powietrznym (poprzedniku Sił Powietrznych USA), gdzie awansowano najszybciej. Było jasne, że żołnierze z ŻW powinni twierdzić, że system awansów był niesprawiedliwy, a ci z wojsk powietrznych, że sprawiedliwy. Badania wykazały jednak rzecz całkiem przeciwną.

Zwróćmy uwagę na dylemat, przed jakim stoi badacz w podobnej sytuacji. Z jednej strony, obserwacje wydają się bezsensowne. Z drugiej strony, wyjaśnienie, które ma oczywisty sens, nie zostało poparte faktami.

Osoba nie tak wybitna jak Stouffer odłożyłaby problem na bok „do dalszego zbadania". On jednak szukał wyjaśnienia dla swoich obserwacji i w końcu je znalazł. Robert Merton i kilku innych socjologów z Columbia University zaczęło myśleć i pisać o czymś, co nazywali *teorią grup odniesienia*. Teoria ta mówi, że ludzie w mniejszym stopniu oceniają swój los odnosząc się do obiektywnych warunków, a w większym – poprzez porównywanie się z ludźmi wokół siebie – ze swoją grupą odniesienia. Na przykład gdybyś mieszkał wśród ubogich, pensja w wysokości 50 000 USD rocznie sprawiłaby, że czułbyś się jak milioner. Lecz z tą samą pensją 50 000 USD, żyjąc wśród ludzi, którzy zarabiają 500 000 USD rocznie, czułbyś się jak biedak.

Stouffer zastosował ten tok rozumowania do badanych żołnierzy. Nawet jeśli jakiś żołnierz z ŻW przez długi czas nie dostawał awansu, było mało prawdopodobne, by znał kogoś, kto mniej na awans zasługując, dostał go wcześniej. Nikt nie dostawał awansu w ŻW. Gdyby zaś należał do Korpusu Powietrznego – nawet jeżeli dostał kilkakrotnie awans w krótkich odstępach czasu – byłby prawdopodobnie w stanie wskazać kogoś, kto, będąc mniej zasłużonym, awansował jeszcze szybciej. Grupą odniesienia dla żołnierza z ŻW byli inni żołnierze z ŻW, a żołnierze z Korpusu Powietrznego porównywali się z kolegami z tegoż korpusu. Wreszcie więc Stouffer zrozumiał stosunek żołnierzy do systemu awansów, a wyjaśnienie (1) miało sens oraz (2) zgadzało się z faktami.

Historia ta pokazuje, że dokumentowanie tego, co oczywiste, jest cenną funkcją każdej nauki, zarówno przyrodniczej, jak i społecznej. Karol Darwin ukuł termin „eksperyment głupca", by opisać większość swych własnych badań – badań, w których sprawdzał to, o czym wszyscy inni „już dawno wiedzieli". Darwin rozumiał, że zdecydowanie zbyt często oczywistości okazują się nieprawdziwe. A zatem pozorna trywialność nie jest uprawnionym zarzutem przeciwko jakiemukolwiek naukowemu przedsięwzięciu.

A wyjątki?

Zarzut, że od każdej prawidłowości społecznej zawsze istnieją wyjątki, nie oznacza, że sama ta prawidłowość nie istnieje albo że jest bez znaczenia. Jakaś konkretna kobieta może z powodzeniem zarabiać więcej pieniędzy niż większość mężczyzn, lecz będzie to małe pocieszenie dla tej większości kobiet, które zarabiają mniej. Reguła wciąż działa. Innymi słowy, społeczne prawa to modele probabilistyczne, które nie stają się mniej rzeczywiste tylko dlatego, że jakieś przypadki nie pasują do ogólnego wzorca.

Uwaga ta stosuje się zarówno do nauk przyrodniczych, jak i społecznych. Fizyka kwantowa, na przykład, jest nauką o prawdopodobieństwach. W genetyce zaślubiny osoby o oczach niebieskich z osobą o oczach brązowych dadzą prawdopodobnie efekt w postaci potomstwa z brązowymi oczami. Narodziny dziecka z niebieskimi oczami nie unieważniają zaobserwowanej prawidłowości, ponieważ genetycy twierdzą jedynie, że potomstwo o brązowych oczach

jest bardziej prawdopodobne oraz że pojawi się ono w określonym procencie wszystkich takich przypadków. Podobnie badacze społeczni przewidują, opierając się na prawdopodobieństwie, że generalnie rzecz biorąc kobiety będą raczej zarabiać mniej od mężczyzn. Gdy reguła taka zostanie już zaobserwowana, badacze społeczni mają podstawy, by pytać, dlaczego ona istnieje.

Wpływ ludzi

Wreszcie zarzut, że zaobserwowane prawidłowości społeczne mogą być zakłócone dzięki świadomej woli aktorów, nie jest poważnym wyzwaniem dla nauk społecznych, choć nie ma dla niego odpowiednika w naukach przyrodniczych. (Przypuszczalnie obiekty fizyczne nie mogą pogwałcić praw fizyki, natomiast probabilistyczna natura fizyki kwantowej pozwoliła pewnym obserwatorom postulować, że elektrony mają wolną wolę). Nie ma wątpliwości, że prawicowy religijny fanatyk mógłby pójść na wybory i zagłosować na lewicowego afroamerykańskiego agnostyka, jeśli zechciałby zmartwić badaczy nauk politycznych zajmujących się tymi wyborami. Wszyscy głosujący mogliby nagle przerzucić swe głosy na stronę dotąd przegrywającą tylko po to, by sfrustrować osoby przeprowadzające sondaże. Podobnie pracownicy mogliby pójść do pracy wcześniej lub w ogóle zostać w domu i tym samym zapobiec spodziewanym godzinom szczytu na drogach. Rzeczy takie nie pojawiają się jednak wystarczająco często, aby poważnie zagrozić obserwacji społecznych prawidłowości.

A zatem społeczne prawidłowości naprawdę istnieją, a badacze społeczni mogą je wyśledzić i obserwować ich skutki. Gdy prawidłowości te zmieniają się w czasie, badacze mogą obserwować i wyjaśniać zmiany.

Zbiorowości zamiast jednostek

Prawidłowości życia społecznego, które studiują badacze społeczni, odzwierciedlają, ogólnie rzecz biorąc, zbiorowe zachowanie wielu jednostek. Choć badacze społeczni często zajmują się motywacjami jednostek, to jednak jednostka jako taka rzadko bywa przedmiotem zainteresowania nauk społecznych. Przeciwnie, badacze społeczni tworzą raczej teorie odnoszące się do natury życia grupowego, a nie jednostkowego. Również przedmiotem ich badań są raczej zbiorowości czy grupy, a nie jednostki.

Czasami prawidłowości zachodzące w zbiorowości są zadziwiające. Rozważmy na przykład wskaźnik urodzeń. Ludzie mają dzieci z niezliczonych osobistych powodów. Niektórzy dlatego, że ich rodzice chcą mieć wnuki. Niektórzy czują, że w ten sposób osiągną pełnię kobiecości czy męskości. Inni chcą utrzymać swe małżeństwo, doświadczać radości wychowywania dzieci, uwiecznić nazwisko rodowe czy osiągnąć pewien rodzaj nieśmiertelności. Jeszcze inni mają dzieci przez przypadek.

Jeśli jesteś ojcem lub jeśli urodziłaś dziecko, możesz prawdopodobnie opowiedzieć jeszcze bardziej szczegółową, pełną osobistych przeżyć historię. Dlaczego dziecko pojawiło się właśnie wtedy, a nie wcześniej czy później? Może straciłeś pracę i musiałeś poczekać jeszcze rok, zanim stać cię było na dziecko? Może po prostu poczułaś się ponaglona faktem, że ktoś z twych bliskich miał dziecko? Każdy, kto zeszłego roku zdecydował się na dziecko, miał po temu swoje powody. Ale pomimo tej ogromnej różnorodności, i pomimo swoistości motywów każdego człowieka, ogólny wskaźnik urodzeń w społeczeństwie – liczba żywych urodzeń na 1000 – jest z roku na rok zadziwiająco stały. W tabeli 1.1 przedstawiono wskaźniki urodzeń w USA z dwudziestu kolejnych lat.

Jeśli wskaźnik urodzeń w USA wyniósłby 15,9, 35,6, 7,8, 28,9 i 16,2 w ciągu kolejnych pięciu lat, demografowie zaczęliby padać jak muchy. Jak jednak widać, życie społeczne jest znacznie bardziej uporządkowane. Co więcej, prawidłowość ta pojawia się bez żadnych ogólnych rozporządzeń. Nikt nie planuje, jak wiele dzieci się ma urodzić, ani nie określa, kto je urodzi. Nie trzeba pozwolenia, aby mieć dziecko; tak naprawdę wiele dzieci zostało poczętych niespodziewanie, a niektóre rodzą się niechciane.

Tabela 1.1

Wskaźnik urodzeń; USA 1977–1996

1977	15,1	1987	15,7
1978	15,0	1988	16,0
1979	15,6	1989	16,4
1980	15,9	1990	16,7
1981	15,8	1991	16,3
1982	15,9	1992	15,9
1983	15,6	1993	15,5
1984	15,6	1994	15,2
1985	15,8	1995	14,8
1986	15,6	1996	14,7

Źródło: Centers for Disease Control and Prevention, National Center for Health Statistics (1998), „Monthly Vital Statistics Report" 46 (11, Suppl.), s. 29.

Teorie nauk społecznych zajmują się więc na ogół zbiorowymi, a nie indywidualnymi zachowaniami. Ich celem jest wyjaśnienie, dlaczego zagregowane wzorce zachowań wykazują taką regularność, nawet gdy uczestniczące w nich osoby mogą zmieniać się w czasie. Można by rzec, że nauki społeczne nawet nie starają się rozumieć ludzi. Starają się pojąć system, w którym ludzie funkcjonują, system, który wyjaśnia, dlaczego ludzie robią to, co robią. Elementami takiego systemu nie są ludzie, lecz zmienne.

Język zmiennych

Nasze najbardziej naturalne próby rozumienia zazwyczaj podejmowane są na poziomie tego, co konkretne i specyficzne. W taki sposób po prostu myślimy.

Wyobraź sobie, że ktoś mówi ci: „kobiety powinny wrócić do kuchni, tam, gdzie ich miejsce". Najpewniej usłyszysz tę uwagę i zinterpretujesz ją w zależności od tego, co wiesz o wypowiadającej ją osobie. Jeśli to twój stary wujek Henio, który, jak sobie przypominasz, jest również zdecydowanie przeciwny oszczędzaniu światła dziennego dzięki zmianie czasu, systemowi kodów pocztowych i komputerom osobistym, prawdopodobnie pomyślisz, że jego ostatnie oświadczenie dobrze wpisuje się w ten cały, raczej przedawniony światopogląd. Jeśli jednak stwierdzenie to wymamrotał jakiś prominentny polityk, który śledził kroki swej przeciwniczki w walce wyborczej, prawdopodobnie tłumaczyłbyś jego wypowiedź w zupełnie inny sposób.

W obu przypadkach starasz się zrozumieć zachowanie danych jednostek. Badania *społeczne* dążą do zrozumienia klas czy typów jednostek. Badacze społeczni chcieliby dowiedzieć się czegoś o tego rodzaju ludziach, którzy podzielają podobny pogląd na temat „właściwej" roli kobiet. Czy ludzie ci mają inne wspólne cechy, które mogą pomóc wyjaśnić ich poglądy?

Nawet jeżeli badacze skupiają swą uwagę na studium pojedynczego przypadku – takiego jak jakaś społeczność lub gang młodzieżowy – ich celem jest uzyskanie wyjaśnień, które pomogłyby zrozumieć inne społeczności lub inne gangi młodzieżowe. Podobnie jak próba pełnego zrozumienia jakiejś osoby pociąga za sobą cel szerszy – rozumienie ludzi w ogóle lub typów ludzkich.

Kiedy kończy się taki projekt służący zrozumieniu i wyjaśnieniu, badacze społeczni są zdolni zrozumieć więcej niż jedną osobę. Wyjaśniając, co sprawia, że pewna grupa osób jest wrogo nastawiona do kobiet aktywnych poza domem, rozumieją wszystkie jednostki o podobnych cechach. Jest to możliwe, ponieważ badali nie tyle antyfeministów, co antyfeminizm. Może się wówczas okazać, że wujek Henio i tamten polityk mają więcej ze sobą wspólnego, niż wydawało się na początku.

O antyfeminizmie mówi się jako o **zmiennej**, ponieważ się zmienia. Niektórzy ludzie wykazują tę postawę w większym stopniu niż inni. Badacze społeczni zainteresowani są zrozumieniem systemu zmiennych, który sprawia, że określona postawa jest silniejsza w jednym przypadku, a słabsza w innym.

Pomysł systemu stworzonego ze zmiennych może się wydać dość dziwny, przypatrzmy się więc jakiejś analogii. Przedmiotem uwagi lekarza jest pacjent. Jeśli pacjent jest chory, celem lekarza jest pomóc mu wydobrzeć. Natomiast przedmiot badań medycznych jest inny: na przykład zmienne, które są przyczyną choroby. Badacz w naukach medycznych może zająć się naszym pacjentem, lecz ma on dla nie-

go znaczenie tylko jako przykład przebiegu choroby.

Nie ma to znaczyć, że badaczy zajmujących się medycyną nie obchodzą prawdziwi ludzie. Z całą pewnością ich obchodzą. Ostatecznym celem studiowania chorób jest ochronienie przed nimi ludzi. Lecz w swych badaniach bardziej interesują się oni regułami rządzącymi pojawianiem się chorób niż indywidualnymi pacjentami. W rzeczywistości, jeśli mogą sensownie studiować chorobę nie angażując w to konkretnych pacjentów, tak właśnie czynią.

A zatem badania społeczne wiążą się ze studiowaniem zmiennych i związków między nimi. Teorie społeczne są sformułowane w języku zmiennych, a ludzie uwikłani są w nie tylko jako „nosiciele" tych zmiennych.

Zmienne z kolei mają coś, co badacze społeczni nazywają **wartościami** lub atrybutami. Wartości to własności lub cechy, które opisują przedmiot – w tym przypadku osobę. Na przykład: kobieta, Azjatka, wyalienowana, konserwatystka, nieuczciwa, inteligentna i pracująca w rolnictwie. Wszystko, co mógłbyś powiedzieć, by opisać siebie lub kogoś innego, wiąże się z wartościami zmiennych.

Zmienne zaś to logiczne pogrupowania wartości. A więc, na przykład, mężczyzna i kobieta to wartości, a płeć to utworzona z nich zmienna. Zmienna *zawód* jest utworzona z wartości takich jak: rolnik, profesor i kierowca ciężarówki. *Klasa społeczna* jest zmienną utworzoną z następujących wartości: klasa wyższa, klasa średnia, klasa niższa. Czasem warto pomyśleć o wartościach jako o „kategoriach", które konstruują zmienną. (Rysunek 1.4 przedstawia schematyczny przegląd tego, co badacze społeczni rozumieją przez zmienne i ich wartości).

Związek między zmiennymi a wartościami to samo sedno zarówno opisu, jak i wyjaśniania w nauce. Na przykład możemy opisać dany rocznik studentów w kategoriach zmiennej *płeć*, odnotowując zaobserwowane częstości występowania wartości mężczyzna i kobieta: „Na roku jest 60% mężczyzn i 40% kobiet". O wskaźniku bezrobocia można myśleć jako o opisie zmiennej *status zatrudnienia siły roboczej*, stosując kategorie: zatrudniony i bezrobotny. Nawet raport dotyczący *dochodu gospodarstw domowych w danym mieście* to podsumowanie wartości składających się na tę zmienną: 3124, 10 980, 35 000 USD itd.

Rysunek 1.4

Zmienne i wartości. W badaniach społecznych i w teoriach zarówno zmienne, jak i ich wartości są odpowiednikami pojęć społecznych. Zmienne to zestawy związanych ze sobą wartości czy atrybutów (cech)

Kilka popularnych pojęć społecznych	
kobieta	wiek
wyższa klasa	Afroamerykanin
młody	zawód
klasa społeczna	płeć
rasa/grupa etniczna	hydraulik

⬇

Zmienna	Wartość
wiek	młody, w średnim wieku, stary
płeć	kobieta, mężczyzna
zawód	hydraulik, prawnik, osoba wprowadzająca dane
rasa/grupa etniczna	Afroamerykanin, Azjata, mieszkaniec Kaukazu, mieszkaniec Ameryki Płd.
klasa społeczna	wyższa, średnia, niższa

Czasem znaczenie pojęć w naukach społecznych jest od razu jasne. W innych przypadkach tak nie jest. Taka sytuacja została omówiona w ramce „Najpoważniejsze straty poniosło...".

Związek między zmiennymi a ich wartościami jest bardziej skomplikowany w przypadku wyjaśniania, a dotyczy on sedna języka zmiennych używanego w teoriach naukowych. Oto prosty przykład z zastosowaniem dwóch zmiennych: wykształcenia i uprzedzeń. Dla ułatwienia przyjmijmy, że zmienna *wykształcenie* przyjmuje tylko dwie wartości: wykształcony i niewykształcony. Analogicznie zmiennej *uprzedzenia* przypiszmy również dwie wartości: uprzedzony i nieuprzedzony.

Przypuśćmy teraz, że 90% niewykształconych jest uprzedzonych, a pozostałe 10% jest nieuprzedzonych. Przyjmijmy również, że 30% ludzi wykształconych ma uprzedzenia, a pozostałe 70% ich nie ma. Graficznie przedstawione jest to na rysunku 1.5A.

Rysunek 1.5A przedstawia relację czy związek między zmiennymi *wykształcenie* a *uprzedzenia*. Relację tę można postrzegać w kategoriach połączenia w pary wartości obu zmiennych. Oto dwa dominujące połączenia: (1) ci, którzy są wykształceni i nieuprzedzeni, oraz (2) ci, którzy są niewykształceni i uprzedzeni. A oto dwa inne pożyteczne sposoby ujęcia tej relacji.

Przypuśćmy najpierw, że wspólnie gramy w grę, zakładając się o to, czy potrafisz zgadnąć, która osoba jest uprzedzona, a która nie. Wybieram za każdym razem jedną osobę (nie mówiąc ci, kogo wybrałem), a ty musisz zgadnąć, czy osoba ta ma uprzedzenia. Zrobimy to z wszystkimi 20 osobami z rysunku 1.5A. Najlepszą dla ciebie strategią byłoby za każdym razem mówić, że ta osoba jest uprzedzona, ponieważ 12 osób z tych 20 jest zaklasyfikowanych w ten sposób. Trafiłbyś wówczas 12 razy na właściwą odpowiedź, a 8 razy na złą – czysty zysk w postaci przewagi 4 trafień.

A teraz przypuśćmy, że wybrawszy którąś osobę z rysunku, muszę powiedzieć ci, czy jest ona wykształcona, czy nie. Tym razem najlepszą strategią jest zgadywać, że osoba jest uprzedzona w przypadku każdej niewykształconej osoby, a w przypadku każdej wykształconej, że uprzedzona nie jest. Jeśli zastosujesz taką stra-

Najpoważniejsze straty poniosło...

Z początkiem 1982 roku straszliwy sztorm spustoszył region Zatoki San Francisco, pozostawiając po sobie śmierć, okaleczenia i zniszczone posiadłości. Środki masowego przekazu, chcąc zwrócić uwagę na najbardziej tragiczne skutki sztormu, raz koncentrowały się na kilkorgu ludziach pogrzebanych żywcem w lawinie błotnej w Santa Cruz, innymi zaś razy opisywały trudne położenie 2900 osób z Marin County, które zostały bez dachu nad głową.

Jak można się domyślać, wszyscy chcieli wiedzieć, gdzie miały miejsce największe zniszczenia, ale odpowiedzieć na to pytanie nie było łatwo. Oto kilka danych opisujących efekty sztormu w dwóch hrabstwach: Marin i Santa Cruz. Przeanalizuj przedstawione porównanie i sprawdź, czy potrafisz określić, które hrabstwo poniosło „najpoważniejsze straty".

	Marin	Santa Cruz
Zniszczone przedsiębiorstwa	15,0 mln USD	56,5 mln USD
Zabici	5	22
Ranni	379	50
Wysiedleni	370	400
Zburzone domy	28	135
Zniszczone domy	2900	300
Zburzone przedsiębiorstwa	25	10
Zniszczone przedsiębiorstwa	800	35
Straty mienia prywatnego	65,1 mln USD	50,0 mln USD
Straty mienia publicznego	15,0 mln USD	56,5 mln USD

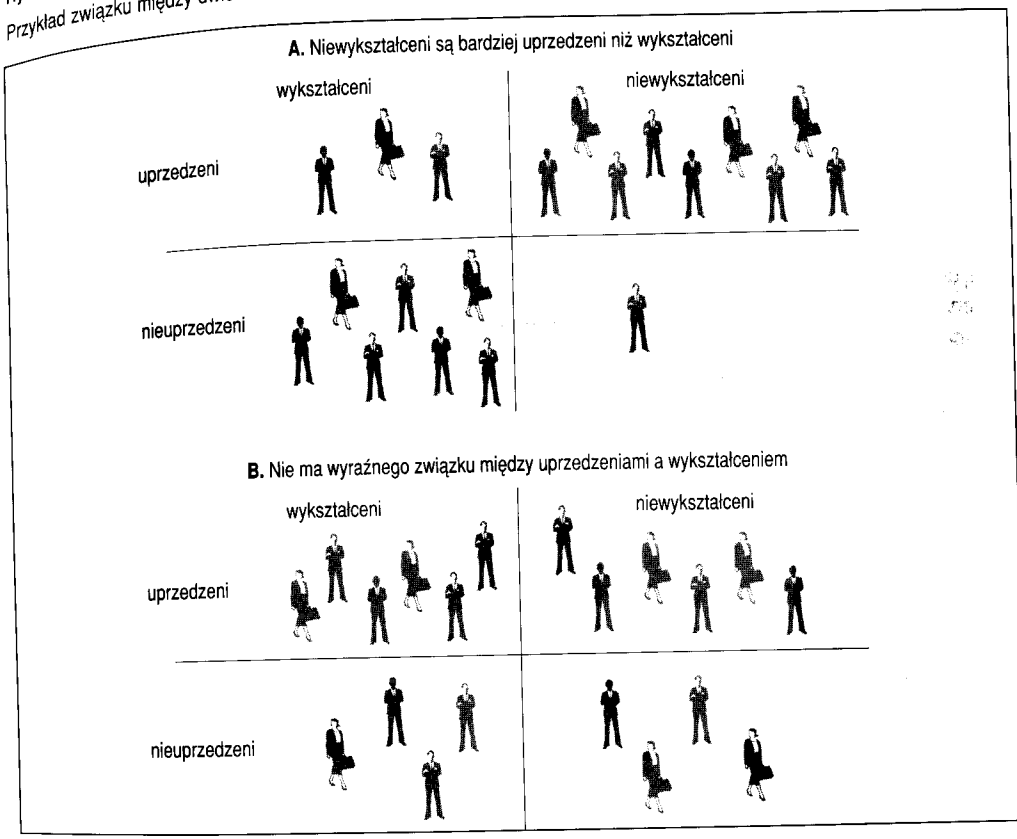
Źródło: „San Francisco Chronicle", 13 stycznia 1982, s.16.

Oczywiście, jeśli mówimy o stratach w ludziach, to z tych dwóch hrabstw Santa Cruz poniosło straty poważniejsze. Lecz w Marin zostało rannych ponad siedem razy więcej osób niż w Santa Cruz; z pewnością pod tym względem to Marin ucierpiało bardziej. Albo, jeśli rozważymy liczbę zburzonych domów, to gorzej było w Santa Cruz, a jeśli tych zniszczonych, to gorzej było w Marin: wszystko zależy od tego, na czym się skupimy. Dylemat ten jest tak samo istotny przy ocenianiu wartości powstałych strat: czy większą uwagę powinniśmy zwrócić na straty mienia prywatnego czy publicznego?

A zatem które hrabstwo poniosło „najpoważniejsze straty"? Ostatecznie na tak zadane pytanie nie ma odpowiedzi. Choć wszyscy mamy wyobrażenie o tym, jak wygląda społeczność, która została „zdruzgotana", oraz jak wygląda ta jedynie „lekko draśnięta", to jednak wizerunki te nie są wystarczająco precyzyjne, aby pozwalały na ścisłe pomiary.

Odpowiedzieć na to pytanie można tylko wtedy, jeśli uściślimy, co rozumiemy przez „najpoważniejsze straty". Jeśli będziemy je mierzyć liczbą ofiar śmiertelnych, wówczas Santa Cruz ucierpiało bardziej. Jeśli zdecydujemy się zdefiniować tę zmienną w kategoriach liczby rannych i/lub wysiedlonych, wtedy w Marin katastrofa była większa. Po prostu nie możemy odpowiedzieć na to pytanie nie wyszczególniając, co dokładnie mamy na myśli mówiąc o „najpoważniejszych stratach". Jest to wymóg podstawowy, który będzie się wciąż na nowo pojawiał przy próbach pomiaru zmiennych w naukach społecznych.

Rysunek 1.5
Przykład związku między dwiema zmiennymi (dwie możliwości)



tegię, otrzymasz 16 dobrych odpowiedzi i 4 złe. Fakt, że dzięki wiedzy o wykształceniu danej osoby poprawiłeś swoje wyniki w odgadywaniu uprzedzeń, dobrze ilustruje, co to znaczy twierdzić, że istnieje związek między zmiennymi.

Teraz dla odmiany rozważmy, jak zaklasyfikowani byliby ci ludzie, gdyby wykształcenie i uprzedzenia nie były ze sobą związane. Przedstawia to rysunek 1.5B. Zauważ, że połowa ludzi jest wykształcona, a połowa nie jest. Zauważ też, że 12 z 20 (60%) jest uprzedzonych. Jeśli w każdej grupie 6 osób na 10 byłoby uprzedzonych, wnioskowalibyśmy, że obie te zmienne nie są ze sobą związane. Wówczas wiedza o tym, że ktoś jest wykształcony, nie miałaby dla nas żadnej wartości przy odgadywaniu, czy osoba ta jest uprzedzona.

Naturze związków między zmiennymi przyjrzymy się dokładniej w części 4. W szczególności prześledzimy kilka sposobów odkrywania i interpretowania tych związków w ramach analiz badawczych. Teraz jednak ważne jest ich ogólne zrozumienie, pozwalające docenić logikę teorii tworzonych w naukach społecznych.

Teorie opisują takie związki między zmiennymi, których z logicznego punktu widzenia można się spodziewać. Oczekiwanie takie często związane jest z ideą przyczynowości. Znaczy to, że spodziewamy się, iż pewne cechy osoby w ramach jednej zmiennej powodują, wywołują czy wspomagają pojawienie się określonych wartości innej zmiennej. Na przedstawionym powyżej przykładzie możemy budować teorię, że wykształcenie bądź jego brak powoduje, że dana osoba będzie z mniejszym lub większym prawdopodobieństwem żywić uprzedzenia.

Jak omówię to ze szczegółami w dalszej części książki, wykształcenie i uprzedzenia

z tego przykładu mogłyby być traktowane odpowiednio jako **zmienna niezależna** i **zmienna zależna**. Przyczynowe, czy inaczej deterministyczne modele zawierają oba te pojęcia. W naszym przykładzie przyjmujemy, że prawdopodobieństwo bycia osobą uprzedzoną jest przez coś zdeterminowane czy spowodowane. Innymi słowy, *uprzedzenia* zależą od czegoś i dlatego nazywamy je zmienną zależną. To, od czego zmienna zależna zależy, jest zmienną niezależną. W tym przypadku jest nią *wykształcenie*. Dla celów tego badania *wykształcenie* jest zmienną niezależną, ponieważ jest niezależne od *uprzedzeń* (tzn. poziom wykształcenia ludzi nie wynika z tego, czy są, czy nie są oni uprzedzeni).

Oczywiście zróżnicowanie poziomu wykształcenia może z kolei samo od czegoś zależeć. Ludzie, których rodzice mają wyższe wkształcenie, będą z dużym prawdopodobieństwem mieć lepsze wykształcenie niż ci, których rodzice byli słabo wykształceni. W tym związku zmiennych wykształcenie osoby jest zmienną zależną, a wykształcenie rodziców – niezależną. Możemy powiedzieć, że zmienna niezależna jest przyczyną, a zmienna zależna – skutkiem.

Powracając do naszego przykładu, omówienie rysunku 1.5 wymagało interpretacji danych. Patrzyliśmy na rozmieszczenie 20 osób pod względem dwóch zmiennych. Konstruując teorię w naukach społecznych, wyprowadzalibyśmy nasze oczekiwania co do związku między dwiema zmiennymi na podstawie tego, co wiemy o każdej z nich. Wiemy na przykład, że wykształcenie pozwala ludziom zetknąć się z szeroką gamą kulturowych zróżnicowań i rozlicznymi punktami widzenia – jednym słowem, poszerza ich horyzonty. Z kolei uprzedzenia świadczą o węższych horyzontach. A więc, logicznie rzecz biorąc, możemy spodziewać się, że wykształcenie i uprzedzenia będą w jakimś stopniu nie do pogodzenia. Moglibyśmy zatem zacząć oczekiwać, że wzrastający poziom wykształcenia będzie ograniczał pojawianie się uprzedzeń, które to oczekiwanie zostałoby poparte przez nasze obserwacje.

Skoro rysunek 1.5 przedstawiał dwie możliwości – że wykształcenie redukuje prawdopodobieństwo zaistnienia uprzedzeń oraz że nie ma na nie wpływu – możesz być ciekaw, co dzieje się faktycznie. W Generalnym Sondażu Społecznym z 1996 roku zapytano ogólnokrajową próbę dorosłych mieszkańców USA, stosując jeden ze wskaźników uprzedzeń, co sądzą o opinii: „Biali, jeśli chcą, mają prawo nie dopuszczać czarnych do zamieszkania w sąsiedztwie, a czarni powinni to prawo respektować". Tylko 6% próby zdecydowanie zgodziło się z tym twierdzeniem, a kolejne 5% zadeklarowało, że raczej się zgadza. Większość – 71% – zdecydowanie się nie zgodziła.

W tabeli 1.2 przedstawiona jest analiza tych danych, grupująca respondentów zgodnie z poziomem zdobytego przez nich wykształcenia. Najprościej jest czytać tę tabelę skupiając się na procentach z ostatniego wiersza: na osobach, które zdecydowanie nie zgadzają się z tą opinią. Silne przeciwstawianie się segregacji stopniowo wzrasta od 62% wśród osób z wykształceniem średnim (lub niższym niż średnie) do 85% wśród absolwentów szkół wyższych. To odkrycie wyraźnie potwierdza pogląd, że wykształcenie zmniejsza uprzedzenia, gdyż to one były przedmiotem tego pomiaru.

Tabela 1.2

Wykształcenie a poparcie dla segregacji

	Poziom wykształcenia respondentów			
	mniej niż średnie	średnie	niepełne wyższe	wyższe
Zdecydowanie się zgadza	10%	7%	6%	1%
Raczej się zgadza	8%	5%	5%	4%
Raczej się nie zgadza	19%	26%	18%	10%
Zdecydowanie się nie zgadza	62%	62%	70%	85%
	100% = (98)	(189)	(190)	(193)

Zauważmy, że teoria ta odnosi się do zmiennych *wykształcenie* i *uprzedzenia*, a nie do ludzi jako takich. Ludzie są nośnikami tych zmiennych, a więc relacja między zmiennymi może zostać dostrzeżona tylko wtedy, gdy obserwujemy ludzi. Ostatecznie jednak teoria stosuje język zmiennych. Opisuje związki między poszczególnymi wartościami różnych zmiennych, których istnienia moglibyśmy się logicznie spodziewać.

Trochę dialektyki badań społecznych

Nie istnieje jeden sposób przeprowadzania badań społecznych. (Gdyby istniał, książka ta byłaby znacznie krótsza). W istocie wiele siły i potencjału badania społeczne czerpią właśnie z wielości uprawnionych podejść.

Jednak u podstaw rozmaitości orientacji badawczych leżą trzy szeroko rozumiane i wzajemnie powiązane rozróżnienia. Choć rozróżnienia te mogą być postrzegane jako wybory konkurencyjne, to dobry badacz społeczny uczy się każdego podejścia. To właśnie rozumiem przez „dialektykę" badania społecznego: między komplementarnymi koncepcjami, które zamierzam opisać, występuje owocne napięcie.

Wyjaśnianie idiograficzne i nomotetyczne

Wszyscy idziemy przez życie, wyjaśniając różne rzeczy. Robimy to każdego dnia. Wyjaśniasz, dlaczego źle lub dobrze poszło ci na egzaminie, dlaczego twoja ulubiona drużyna wygrywa lub przegrywa, dlaczego możesz mieć problemy z umawianiem się na fajne randki czy ze zdobyciem porządnej pracy. W naszym codziennym wyjaśnianiu stosujemy dwie odmienne formy rozumowania przyczynowego, choć zazwyczaj nie rozróżniamy ich.

Czasem próbujemy wyjaśnić pojedynczą sytuację wyczerpująco. A więc: źle poszło ci na egzaminie, ponieważ (1) zapomniałeś, że tego dnia jest egzamin, (2) z tego przedmiotu jesteś najsłabszy, (3) spóźniłeś się, bo na drodze były korki, (4) w noc przed egzaminem nie spałeś, ponieważ twój współlokator słuchał głośnej muzyki, (5) policja zajęła cię aż do świtu, domagając się wyjaśnień, co zrobiłeś z zestawem stereo twego współlokatora, a także, co zrobiłeś z samym współlokatorem, no i (6) stado dzikich kojotów zjadło twój podręcznik. W takich okolicznościach nic dziwnego, że źle ci poszło.

Taki rodzaj przyczynowego rozumowania nazywany jest wyjaśnianiem **idiograficznym**. *Idio-* w tym kontekście oznacza wyjątkowy, wyodrębniony, szczególny lub odmienny, podobnie jak w słowie: idiosynkrazja. Gdy skończyliśmy dawać wyjaśnienie idiograficzne, mamy poczucie, że w pełni rozumiemy przyczyny tego, co stało się w tym konkretnym przypadku. Jednocześnie wyłącznie do tego konkretnego przypadku ograniczony jest zasięg naszego wyjaśnienia. Choć fragmenty wyjaśnienia idiograficznego mogą mieć zastosowanie do innych sytuacji, naszą intencją jest w pełni wyjaśnić jeden przypadek.

Rozważmy teraz inny rodzaj wyjaśniania. (1) Za każdym razem, gdy uczysz się w grupie, idzie ci na egzaminie lepiej, niż wtedy, gdy uczysz się sam. (2) Twoja ulubiona drużyna lepiej radzi sobie na własnym niż na cudzym boisku. (3) Członkowie stowarzyszeń studenckich częściej umawiają się na randki niż członkowie kółka biologicznego. Ten typ wyjaśniania – określany jako **nomotetyczny** – szuka wyjaśnień raczej dla pewnej klasy sytuacji czy zdarzeń niż dla pojedynczych przypadków. Co więcej, szuka wyjaśnień w sposób „ekonomiczny", używając tylko jednego lub kilku czynników wyjaśniających. Ostatecznie zadowala się raczej cząstkowym, a nie kompletnym wyjaśnieniem.

W każdym z tych przypadków mógłbyś osłabić swoje twierdzenia dotyczące przyczyn za pomocą takich słów czy terminów, jak: „w sumie", „zazwyczaj" czy „w takich samych warunkach". A więc zazwyczaj lepiej ci idzie na egzaminach, do których uczyłeś się w grupie, ale nie jest tak zawsze. Podobnie twoja drużyna wygrała kilka meczów poza własnym boiskiem, a przegrała kilka na własnym. Atrakcyjny przewodniczący koła biologicznego może ciągle chodzić na fajne randki, a brzydcy członkowie stowarzyszeń studenckich wiele sobotnich nocy spędzają samotnie, rozwiązując krzyżówki.

Obecność takich wyjątków to cena, jaką płacimy za ogólne wyjaśnienie mające szersze zastosowanie. Jak zauważyliśmy wcześniej, prawidłowości są realne i znaczące, nawet jeśli nie są doskonałe.

Zarówno idiograficzne, jak i nomotetyczne podejście do wyjaśniania może być użyteczne w życiu codziennym. Prawidłowości nomotetyczne, które odkryjesz, mogą być dobrym przewodnikiem na przykład przy planowaniu zwyczajów uczenia się, podczas gdy idiograficzne wyjaśnienie będzie bardziej przekonujące dla twego promotora.

Podobnie zarówno idiograficzne, jak i nomotetyczne rozumowanie jest skutecznym narzędziem w badaniach społecznych. Zastanówmy się najpierw nad rozumowaniem idiograficznym. Badacz, który dąży do wyczerpującego wyjaśnienia funkcjonowania wewnętrznej struktury jakiegoś konkretnego gangu młodzieżowego czy zasad zespołowego zarządzania jakąś międzynarodową korporacją, zajmuje się badaniem idiograficznym: próbuje zrozumieć tę konkretną grupę najpełniej, jak to możliwe.

Badacz David Wellman (1995) podjął się na przykład pogłębionej analizy świadomości klasowej członków 10 Oddziału Międzynarodowego Związku Robotników Portowych i Magazynierów w San Francisco (*International Longshoremen's and Warehousemen's Union – ILWU*). Wellman zauważył, że ten konkretny związek nie był typowym przykładem organizacji pracowników w USA. Choć zainteresowany był poznaniem związków zawodowych jako takich oraz natury kapitalizmu, jego bezpośrednim celem było pełne zrozumienie historii 10 Oddziału.

Często celem badań jest jednak bardziej ogólne wyjaśnienie dotyczące klasy zdarzeń – innymi słowy wyjaśnienie nomotetyczne – choć poziom zrozumienia jest tu z konieczności mniej pełny w odniesieniu do każdego pojedynczego przypadku. Na przykład badacze, którzy chcą odkryć główne czynniki prowadzące do przestępczości wśród młodzieży, przeprowadzają badania nomotetyczne. Mogą odkryć, że dzieci z rozbitych rodzin częściej popełniają przestępstwa niż te z rodzin pełnych. Wyjaśnienie to wykraczałoby znacznie poza dowolny przypadek pojedynczego dziecka, ale uogólnienie to odbywa się kosztem pełnego wyjaśnienia, dlaczego dane konkretne dziecko wchodzi na drogę przestępstwa.

W przeciwieństwie do badania Wellmana poświęconego 10 Oddziałowi, Susan Tiano (1994) próbowała wyjaśnić ogólny wpływ uprzemysłowienia w Trzecim Świecie na sytuację kobiet. Czy wkroczenie do przemysłu kobiet jako siły roboczej oznaczało wyzwolenie czy ucisk? Jej badania robotnic przemysłowych w Meksyku ilustrują podejście nomotetyczne do wyjaśniania.

A zatem badacze społeczni mają dostęp do dwóch różnych rodzajów wyjaśniania. Tak jak fizycy czasem traktują światło jak cząsteczkę, a czasem jak falę, tak i badacze społeczni jednego dnia mogą poszukiwać związków na szeroką skalę, a innego zgłębiać to, co specyficzne i jedyne w swoim rodzaju. Oba te podejścia są prawdziwą nauką, oba dają satysfakcję i oba mogą sprawiać przyjemność.

Teoria indukcyjna i dedukcyjna

Podobnie jak w przypadku wyjaśniania idiograficznego i nomotetycznego, zarówno myślenie indukcyjne, jak i dedukcyjne odgrywa dużą rolę w naszym życiu codziennym. Jest to kolejne ważne rozróżnienie w badaniach społecznych.

Są dwie metody, by dojść do wniosku, że lepiej ci idzie na egzaminach, gdy uczysz się do nich z innymi. Z jednej strony, mniej więcej na półmetku twych studiów możesz zastanawiać się, dlaczego raz na egzaminach idzie ci dobrze, a innym razem słabo. Możesz zrobić listę wszystkich egzaminów, jakie zdawałeś, odnotowując, jak poradziłeś sobie na każdym z nich. Potem możesz spróbować przypomnieć sobie wszelkie okoliczności, jakie towarzyszyły wszystkim egzaminom dobrze zdanym i wszystkim tym, na których poszło ci gorzej. Lepiej szło ci na testach czy przy pisaniu esejów? W godzinach rannych czy popołudniowych? Na egzaminach z nauk przyrodniczych, humanistycznych, czy z nauk społecznych? Wtedy, gdy uczyłeś się sam czy... BINGO! Okazuje się, że prawie zawsze najlepiej szło ci na

egzaminach, do których uczyłeś się z innymi. Taki tryb badawczy jest nazywany *indukcją*.

Rozumowanie indukcyjne, czy inaczej **indukcja**, przechodzi od tego, co konkretne, do tego, co ogólne, od zestawu szczegółowych obserwacji do odkrycia prawidłowości, która odzwierciedla pewien stopień uporządkowania wszystkich danych zdarzeń. Zauważ przy okazji, że twoje odkrycie mówi ci jedynie, że określony wzór istnieje, a niekoniecznie: *dlaczego*.

Istnieje drugi, całkiem odmienny sposób dojścia do tego samego wniosku dotyczącego uczenia się do egzaminów. Wyobraź sobie, że podchodzisz do pierwszej sesji egzaminacyjnej na studiach. Zastanawiasz się, jakie są najlepsze metody uczenia się – czy bardziej powinieneś się skupić na notatkach z zajęć, czy raczej przejrzeć lektury. Dowiadujesz się, że niektórzy studenci przygotowują się, przepisując w systematyczny sposób swoje notatki. Potem zastanawiasz się, czy powinieneś uczyć się w umiarkowanym tempie, czy raczej zarwać noc tuż przed egzaminem. Namyślając się tak, możesz zadać sobie pytanie, czy powinieneś spotkać się z innymi studentami z twojej grupy, czy raczej uczyć się sam. Możesz oszacować wszystkie za i przeciw każdej opcji.

Studiowanie z innymi bywa nie tak efektywne, ponieważ wiele czasu spędza się wówczas nad rzeczami, które już się rozumie. Jednak możesz zrozumieć coś jeszcze lepiej, jeżeli wytłumaczysz to komuś innemu. Inni studenci mogą rozumieć te części programu, których ty jeszcze nie rozumiesz. Kilka umysłów może odsłonić perspektywy, które tobie mogłyby umknąć. A zobowiązanie się do nauki z innymi uczyni bardziej prawdopodobnym to, że będziesz się uczył, a nie oglądał powtórkę popularnego serialu.

Możesz w podobnym stylu dodać jeszcze kilka za i przeciw i logicznie wywnioskować, że z korzyścią dla ciebie będzie uczyć się z innymi. Wydaje ci się to rozsądne, tak jak rozsądne wydaje ci się, że lepiej jest uczyć się, niż się nie uczyć. Mówi się czasami, że takie rzeczy są prawdą „w teorii". Aby uzupełnić proces badawczy, sprawdzamy, czy są prawdziwe w praktyce. Aby test był dokładny, możesz do połowy egzaminów przygotowywać się z innymi, a do drugiej połowy sam. Ta procedura sprawdzi logikę twojego rozumowania.

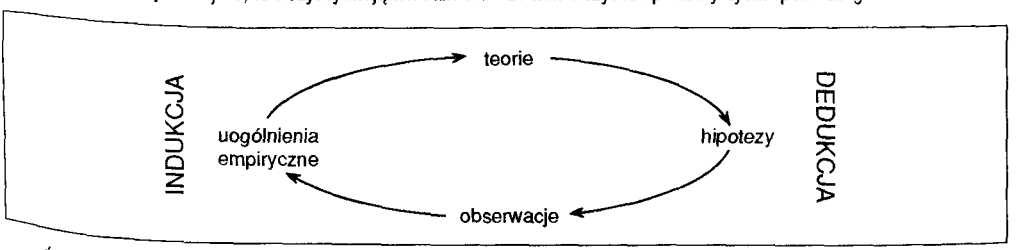
Ten drugi tryb badawczy, znany jako rozumowanie dedukcyjne lub **dedukcja**, przechodzi od tego, co ogólne, do tego, co szczegółowe. Przechodzi od (1) wzorca, którego możemy oczekiwać z logicznego lub teoretycznego punktu widzenia, do (2) obserwacji, które sprawdzają, czy oczekiwany wzorzec rzeczywiście się pojawia. Zauważ, że dedukcja zaczyna od „dlaczego" i przechodzi do „czy", podczas gdy indukcja to ruch w przeciwnym kierunku.

Oba te tak różne podejścia są prawomocnymi drogami w nauce. Co więcej, często współgrając, dostarczają jeszcze bardziej znaczących i kompletnych wyjaśnień, jak przedstawiono na rysunku 1.6.

Przy okazji zauważ, że rozróżnienie dedukcyjnego i indukcyjnego rozumowania niekoniecznie wiąże się z rozróżnieniem nomotetycznego i idiograficznego trybu wyjaśniania. Te

Rysunek 1.6

Koło nauki. Cykl teorii i badań może być porównany do sztafety; choć nie wszyscy uczestnicy muszą rozpoczynać czy kończyć w tym samym miejscu, to wszyscy mają ten sam cel – zbadać wszystkie poziomy życia społecznego



Źródło: Walter Wallace, *The Logic of Science in Sociology*, Aldine deGruyter, New York 1971. Copyright © 1971 by Walter Wallace. Dostosowano i wykorzystano za pozwoleniem.

cztery charakterystyki reprezentują cztery możliwości, dające się wykorzystać zarówno w codziennym życiu, jak i w badaniach społecznych.

Na przykład stosując dedukcję i wyjaśnianie idiograficzne, możesz się przygotować do jakiejś randki, biorąc pod uwagę wszystko, co wiesz o osobie, z którą masz się spotkać, i próbując logicznie antycypować, jaki rodzaj ubrania, zachowania, uczesania, higieny ustnej itd. sprzyjają udanej randce. Albo, idiograficznie i za pomocą indukcji, możesz spróbować ocenić, co właściwie sprawiło, że dziewczyna, z którą się spotkałeś, zadzwoniła pod numer 997.

Nomotetyczne, dedukcyjne podejście pojawia się, gdy udzielasz innym korepetycji z „podstaw udanej randki", gdy mądrze im objaśniasz, dlaczego można chłopakowi zaimponować przedstawiając mu niebezpieczeństwa satanistycznego przesłania ukrytego w tekstach rockandrollowych piosenek. Gdy potem przyjrzysz się swemu życiu i zastanowisz, dlaczego nie miałaś więcej randek z muzykami, możesz przejść na drogę nomotetycznej indukcji.

Do indukcji i dedukcji powrócimy w rozdziale 2. A teraz zwróćmy się ku trzeciemu ogólnemu rozróżnieniu, które tworzy wiele odmian badań społecznych.

Dane jakościowe i ilościowe

Rozróżnienie danych jakościowych i ilościowych w naukach społecznych zasadniczo pokrywa się z rozróżnieniem danych liczbowych i nieliczbowych. Gdy nazywamy kogoś inteligentnym, stwierdzenie to ma charakter jakościowy. Analogiczne stwierdzenie, dotyczące kogoś mniej szczęśliwie wyposażonego, brzmiałoby „nieinteligentny". Kiedy psychologowie i inni naukowcy mierzą inteligencję na skali IQ, próbują takie stwierdzenia jakościowe skwantyfikować. Na przykład psychologowie mogą powiedzieć, że IQ jakiejś osoby wynosi 120.

Każda obserwacja ma na początku charakter jakościowy, bez względu na to, czy jest to doświadczanie czyjejś inteligencji, czy umiejscowienie wskazówki na skali pomiaru, czy odpowiedź zaznaczona w kwestionariuszu. Natura żadnej z tych rzeczy nie jest liczbowa czy ilościowa, czasem jednak korzystnie jest nadać im tę numeryczną formę.

Kwantyfikacja często sprawia, że nasze obserwacje stają się bardziej wyraziste. Może też ułatwić agregację, porównywanie i podsumowywanie danych. Poza tym otwiera możliwość zastosowania analiz statystycznych, począwszy od prostych średnich, aż do złożonych wzorów i modeli matematycznych.

Dane ilościowe mają te same zalety, co liczby użyte jako mierniki jakiejś cechy w stosunku do słów. Ale mają też te same co liczby wady, włączając w to potencjalną utratę bogactwa znaczeniowego. Na przykład badacz społeczny może chcieć dowiedzieć się, czy studenci w wieku od 18 do 22 lat skłonni są raczej umawiać się z młodszymi czy ze starszymi od siebie. Udzielenie ilościowej odpowiedzi na to pytanie wydaje się zadaniem prostym. Badacz pyta pewną liczbę studentów o wiek każdej z osób, z którymi się spotykali, oblicza średnią i porównuje ją z wiekiem osoby pytanej. Sprawa zamknięta.

Ale czy rzeczywiście? Wprawdzie tutaj wiek równa się liczbie przeżytych lat, ale czasem ludzie używają tego terminu inaczej; może dla niektórych „wiek" tak naprawdę oznacza „dojrzałość". Choć osoby, z którymi się umawiasz, mogą być na ogół młodsze od ciebie, to jednocześnie mogą zachowywać się dojrzalej, a zatem reprezentować tę samą grupę „wieku". Ktoś inny znów może przez „wiek" rozumieć to, jak się wygląda, młodo czy staro, albo poziom zróżnicowania czyichś życiowych doświadczeń i jego obycia. Tracimy te znaczenia obliczając średnią. Krótko mówiąc, jakościowe dane mogą być bogatsze znaczeniowo niż dane skwantyfikowane. To właśnie milcząco zakłada formułka „jest starszy od swojej metryki". Utracilibyśmy sens tego wyrażenia, próbując uszczegółowić „o ile starszy".

Z kolei jakościowe dane mają te same wady co czysto słowne opisy. Na przykład bogactwo znaczeń, o jakim wspomniałem, częściowo jest funkcją wieloznaczności. Jeśli wyrażenie „jest starszy od swojej metryki" coś dla ciebie oznaczało, gdy je czytałeś, to znaczenie owo wyrasta z twych własnych doświadczeń albo doświadczeń ludzi, których znałeś i do których wyrażenie to pasowało, czy też z tych sytuacji, w któ-

rych ktoś go w twojej obecności używał. Dwie rzeczy są pewne: (1) ty i ja zapewne nie mamy na myśli dokładnie tego samego oraz (2) nie wiesz dokładnie, co ja mam na myśli i *vice versa*.

Mam młodego przyjaciela, Raya Zhanga, który, gdy w 1989 roku demonstrowano w imię wolności na placu Tiananmen w Pekinie, był tam odpowiedzialny za łączność. Po interwencji wojska Ray zbiegł na południe, został aresztowany, a potem zwolniony z poleceniem powrotu do Pekinu. Zamiast tego uciekł z Chin i przedostał się do Paryża. W końcu przybył do Stanów Zjednoczonych, gdzie skończył studia, które wcześniej zmuszony był przerwać, uchodząc ze swej ojczyzny. Widziałem, jak radził sobie z trudnościami – musiał dostać się do szkoły, nie mając żadnych papierów z Chin, studiować w obcym języku, zaspokajać swe potrzeby finansowe – wszystko to na własną rękę, tysiące kilometrów od swej rodziny. Ray wciąż mówi o dniu, kiedy wróci do Chin budować tam system demokratyczny.

Uderza mnie w Rayu to, że „jest starszy od swojej metryki". Prawdopodobnie się z tym zgodzisz. Ten dodatkowy szczegół mojego jakościowego opisu, choć uszczegóławia znaczenie tego sformułowania, wciąż nie pozwala nam powiedzieć, o ile ktoś jest starszy, ani nawet porównać pod tym względem dwóch ludzi bez ryzyka nieporozumień co do tego, który jest bardziej przez życie doświadczony.

Niemniej jednak możliwa jest kwantyfikacja tego pojęcia. Moglibyśmy, na przykład, stworzyć listę doświadczeń, które składałyby się na to, co rozumiemy przez doświadczenie życiowe:

Małżeństwo
Rozwód
Śmierć ojca lub matki
Bycie świadkiem morderstwa
Bycie aresztowanym
Bycie wygnańcem
Utrata pracy
Ucieczka z cyrkiem

Moglibyśmy skwantyfikować ludzkie doświadczenie życiowe jako liczbę tego typu doświadczeń: o tych, którzy mieli ich więcej, powiedzielibyśmy, że są bardziej doświadczeni. Gdybyśmy uznali niektóre z tych doświadczeń za ważniejsze, moglibyśmy przypisać im więcej punktów. Gdy już ustalilibyśmy listę i system punktacji, przyznawanie ludziom punktów i umieszczanie ich na skali liczbowej byłoby proste. Nie mielibyśmy żadnych trudności ze zgodnym ustalaniem, kto od kogo ma więcej punktów.

A zatem, aby skwantyfikować takie nieliczbowe pojęcie jak doświadczenie życiowe, musimy jednoznacznie określić, co to pojęcie znaczy. Jednak skupiając się szczególnie na tym, co włączymy do naszej miary tego pojęcia, jednocześnie wykluczamy każde inne znaczenie. Nieuchronnie dokonujemy tu pewnego wyboru: każdy jednoznaczny, ilościowy pomiar będzie uboższy w znaczenie niż odpowiadający mu opis jakościowy.

Co za dylemat! Które podejście powinniśmy wybrać? Które jest lepsze? Które jest bardziej odpowiednie dla badań społecznych?

Dobra wiadomość. Nie musimy wybierać. Właściwie nawet nie powinniśmy. Zarówno jakościowe, jak i ilościowe metody są użyteczne i uprawnione w badaniach społecznych. Niektóre zagadnienia i sytuacje badawcze lepiej poddają się metodom ilościowym, inne – jakościowym.

Choć badacze społeczni mogą używać obu tych podejść, każde wymaga innych umiejętności i procedur. W rezultacie możesz stwierdzić, że lepiej się czujesz – i uzyskujesz większą biegłość – w jednym z nich. Aczkolwiek im skuteczniej będziesz umiał korzystać z obu, tym lepszym badaczem się staniesz. Na pewno wszyscy badacze, niezależnie od swych osobistych skłonności, powinni uznawać obie te metody za uprawnione.

Zauważyłeś może, że podejście jakościowe wydaje się bardziej związane z wyjaśnianiem idiograficznym, natomiast wyjaśnianie nomotetyczne jest łatwiej osiągalne poprzez kwantyfikację. Jest to prawda, choć związki te nie są bezwarunkowe. Co więcej, w ramach obu tych perspektyw istnieje znacząca ilość obszarów trudnych do zaklasyfikowania. Uznanie podziału na badania jakościowe i ilościowe nie oznacza ścisłego związania własnej działalności badawczej tylko z jednym ich rodzajem, z wykluczeniem drugiego. Pełne

zrozumienie jakiegoś zagadnienia często wymaga zastosowania obu technik.

Etyka w badaniach społecznych

Większość stron tej książki jest poświęcona logice i umiejętności przeprowadzania badań społecznych, różnym technikom preferowanym przez badaczy i powodom, dla których badacze je cenią. Istnieje jednak kilka istotnych, nienaukowych kwestii, które kształtują działalność badaczy społecznych. Kwestią kluczową jest sprawa etyki w badaniach.

Rozdział 18 tej książki traktuje obszernie o etyce badawczej, a gdy będzie to zasadne, odniesienia do zagadnień etycznych będą się pojawiać także w innych rozdziałach. Na razie chciałbym wprowadzić dwie podstawowe kwestie, o których należy pamiętać, czytając dalej tę książkę.

Dobrowolne uczestnictwo

Podstawowa zasada etyki badań społecznych głosi, że udział w nich powinien być dobrowolny. W zasadzie wydaje się to warunek dość łatwy do spełnienia. Badacz, który zmusiłby ludzi do udziału w eksperymencie, zostałby powszechnie skrytykowany. Podobnie zmuszanie ludzi do wypełniania kwestionariusza poprzez zamknięcie ich w pokoju czy grożenie im strasznymi konsekwencjami, jeżeli odmówią (na przykład niezaliczeniem seminarium), byłoby wyraźnie nieetyczne.

Jednak sprawy te nie zawsze są tak oczywiste. Gdy zgodnie z zasadami obserwujemy demonstrację w studenckim campusie, nie pytamy wszystkich uczestników o zgodę. Gdy badacz w celu przeprowadzenia badań udaje, że dołączył do jakiejś grupy wyznaniowej, ci, których obserwuje, właściwie nie zgłosili się dobrowolnie do udziału w projekcie badawczym. Badacze społeczni często dyskutują o tym, czy plan konkretnych badań nie gwałci ustalonych zasad etyki badawczej. Gdy przyglądasz się już przeprowadzonym badaniom albo myślisz o swych własnych planach badawczych, zadaj sobie pytanie, czy wybrana technika respektuje zasadę dobrowolnego uczestnictwa.

Nie krzywdzić badanych

Fundamentalna norma etyczna badań społecznych zabrania w jakikolwiek sposób krzywdzić badanych. Zasadniczo każdy z pewnością zgodzi się z tą regułą. Zobaczymy jednak, że czasem trudno jest w pełni się jej podporządkować.

Wyobraźmy sobie na przykład, że niektórzy ludzie pytani przez badaczy o poglądy religijne uświadamiają sobie po raz pierwszy, że mają wątpliwości dotyczące religii. Albo wyobraźmy sobie badania poświęcone traktowaniu kobiet w społeczeństwie, które sprawią, że część kobiet przestaje być zadowolona ze swych małżeństw czy pracy zawodowej. Kiedy badania mogą przynieść szkodę badanym poprzez wywieranie na nich wpływu?

Jak zobaczymy, trzymanie się tej na pierwszy rzut oka prostej zasady etycznej – „nie krzywdzić" – wymaga czujności ze strony badaczy. Planując własne badania, upewnij się, że zadałeś sobie pytanie, czy mogłyby one przynieść szkodę ludziom, których chcesz badać. Ponieważ wszystko, co w życiu robimy, potencjalnie rzecz biorąc może kogoś skrzywdzić, każdy badacz musi wyważyć między względnym ryzykiem a tym, jak istotne są dane badania i jakie przyniosą one korzyści.

Badacze społeczni mają wiele sposobów, by ustrzec się przed krzywdzeniem ludzi. Na przykład uważamy, by respektować prawo do prywatności badanych. Badania często wymagają poznania szczegółów życia prywatnego ludzi, więc przyjmujemy na siebie zobowiązanie, że zachowamy w tajemnicy to, czego się dowiemy. Niejednokrotnie przechowujemy dane anonimowo, a więc nie ma sposobu na ustalenie tożsamości człowieka, od którego pochodzą dobrowolnie przekazane informacje; zabezpieczamy się w ten sposób nawet przed przypadkowym wydostaniem się informacji na zewnątrz.

Przekonasz się, że choć wprowadzanie w błąd jest często konieczne, by przeprowadzić projekty badawcze pewnego rodzaju, badacze zobowiązują się go unikać, chyba że jest ono

niezbędne. (Na przykład jeśli przedstawisz sondaż czy eksperyment słowami: „Chcemy się dowiedzieć, w jakim stopniu jest pan uprzedzony", badani raczej tak dostosują swoje słowa i czyny, aby wyglądać na pozbawionych uprzedzeń). Gdy jednak wprowadzanie w błąd ludzi uznamy za konieczne dla naszych celów badawczych, musimy spytać, czy ewentualna wartość tych badań usprawiedliwia takie działanie.

Powaga problemów etycznych jest wyraźnie widoczna w kodeksach etycznych tworzonych i publikowanych przez stowarzyszenia zawodowe, których członkowie zajmują się badaniami społecznymi. W kodeksach tych rozpatruje się kwestie dobrowolnego uczestnictwa, niewyrządzania krzywdy oraz wiele innych, które poznamy w rozdziale 18.

Oto więc kilka zagadnień leżących u podstaw badań społecznych. Mam nadzieję, że omówienie to pomogło wykazać, że z badaniami społecznymi wiąże się wszystko oprócz rutyny i nudy. W swym najlepszym wydaniu są one działalnością tętniącą życiem, ekscytującą i doniosłą. Wszystko, czego nam trzeba, to umysł otwarty i spragniony przygód.

Główne zagadnienia

- Tematem tej książki jest to, w jaki sposób poznajemy rzeczywistość społeczną.
- Badania są naturalną działalnością człowieka. Duża część zwykłej działalności badawczej ma na celu wyjaśnianie i przewidywanie wydarzeń.
- Gdy rozumiemy coś dzięki bezpośredniemu doświadczeniu, to czynimy obserwacje i szukamy prawidłowości w tym, co obserwujemy.
- Większość naszej wiedzy bierze się z umowy, a nie z doświadczenia. Zwłaszcza dwa źródła uzgodnionej wiedzy są znaczące: tradycja i autorytet. Jednak, choć przydatne, oba mogą nas sprowadzić na manowce.
- Nauka szuka zabezpieczeń przed błędami, które popełniamy w zwykłych badaniach.
- Często nasze obserwacje są niepoprawne, więc badacze starają się uniknąć takich błędów, prowadząc obserwacje w sposób staranny i przemyślany.
- Czasem pochopnie wyciągamy ogólne wnioski na podstawie zaledwie kilku obserwacji, zatem naukowcy, chcąc uniknąć nieuprawnionego uogólniania, zobowiązują się do przeprowadzenia odpowiedniej liczby obserwacji i powtarzają badania.
- W życiu codziennym często rozumujemy nielogicznie. Badacze, chcąc uniknąć nielogicznego rozumowania, starają się o taką samą ostrożność i rozwagę w swoim rozumowaniu, co w swoich obserwacjach. Ponadto publiczny charakter nauki oznacza, że zawsze znajdą się inni badacze, którzy zakwestionują błędne rozumowanie.
- Ujęcie premodernistyczne, modernistyczne i postmodernistyczne to trzy spojrzenia na „rzeczywistość". Zgodnie z ujęciem postmodernistycznym, nie istnieje „obiektywna" rzeczywistość, niezależna od naszych subiektywnych doświadczeń. Różne filozoficzne perspektywy oferują badaniom społecznym wiele możliwych podejść.
- Teoria społeczna próbuje rozważać i wyjaśniać to, co jest, a nie to, co powinno być. Teorii nie należy mylić z filozofią czy wiarą.
- Nauki społeczne szukają prawidłowości w życiu społecznym.
- Badacze społeczni interesują się zbiorowościami, a nie jednostkami.
- Teorie formułuje się w języku zmiennych.
- Zmienna jest logicznym zbiorem wartości. Wartość to cecha, na przykład: mężczyzna i kobieta. *Płeć* jest zmienną utworzoną z tych wartości.
- W wyjaśnianiu przyczynowym domniemana przyczyna jest zmienną niezależną, podczas gdy zmienna, na którą oddziałuje, jest zmienną zależną.
- Wyjaśnienia idiograficzne dążą do pełnego zrozumienia poszczególnych przypadków, nomotetyczne zaś mają na celu uogólnione rozumienie wielu przypadków.
- W teoriach indukcyjnych rozumowanie przebiega od poszczególnych obserwacji do ogólnych reguł. Teorie dedukcyjne zaczynają od stwierdzeń ogólnych i przewidują konkretne obserwacje.
- Dane ilościowe są liczbowe; jakościowe

nie są. Oba rodzaje danych są użyteczne dla różnych celów badawczych.

- Etyka jest kluczową kwestią przy planowaniu badań społecznych. Stwierdzenia, że uczestnictwo w badaniu powinno być dobrowolne oraz że nie należy w żaden sposób krzywdzić badanych, to dwie podstawowe wskazówki etyczne.

Kluczowe pojęcia

Oto terminy, które w tym rozdziale zostały zdefiniowane w odpowiednich kontekstach, i które można też znaleźć w słowniczku na końcu książki:

replikacja	zmienna zależna
teoria	idiograficzny
zmienna	nomotetyczny
wartość	indukcja
zmienna niezależna	dedukcja

Pytania i ćwiczenia

1. Przypomnij, jakie błędy powszechnie popełniane przez ludzi w ich badaniach omówione są w tym rozdziale. Znajdź artykuł w gazecie czy czasopiśmie albo list do redakcji, który ilustrowałby jeden z tych błędów. Opowiedz, w jaki sposób naukowiec uniknąłby tego błędu.
2. Wymień pięć zmiennych społecznych i określ wartości, które się na nie składają.
3. Wejdź na jedną z następujących stron w Internecie i znajdź przykłady danych zarówno jakościowych, jak i ilościowych.
   a. Wysoki Komisarz Narodów Zjednoczonych ds. Uchodźców (*UN High Commissioner for Refugees*):
   http://www.unhcr.ch/
   b. Amerykańskie Centra Kontroli Medycznej i Zapobiegania Chorobom (*U.S. Centers for Disease Control and Prevention*):
   http://www.cdc.gov/
   c. Biblioteka Narodowa Australii (*National Library of Australia*):
   http://www.nla.gov.au/

Dodatkowe lektury

Babbie Earl, *The Sociological Spirit*, Wadsworth, Belmont, Cal. 1994. Jest to elementarz poświęcony niektórym socjologicznym punktom widzenia. Wprowadza wiele pojęć powszechnie używanych w naukach społecznych.

Babbie Earl, *Observing Ourselves: Essays in Social Research*, Waveland Press, Prospect Heights, Ill. 1998. Zbiór esejów rozwijających kilka zagadnień filozoficznych, na które natkniesz się w kolejnych rozdziałach, m.in. obiektywność, paradygmaty, determinizm, pojęcia, rzeczywistość, przyczynowość oraz wartości.

Becker Howard S., *Tricks of the Trade: How to Think about Your Research While You Are Doing It*, University of Chicago, Chicago 1997. Ta bardzo przystępna książka sprawia, że ma się wielką ochotę na prowadzenie badań społecznych, czy to jakościowych, czy ilościowych. Pełna jest anegdot dotyczących badań, które pokazują badania społeczne jako śmiałe wyzwanie.

Cole Stephen, *Making Science: Between Nature and Society*, Harvard University Press, Cambridge, Mass. 1992. Jeśli interesuje cię pogłębienie wiedzy o nauce jako przedsięwzięciu społecznym, prawdopodobnie uznasz tę książkę za fascynującą analizę.

Gallup George Jr. i in., *Polls that Made the Difference*, „The Public Perspective" maj–czerwiec 1990, s. 17–21. Kilku badaczy opinii publicznej dyskutuje tu o sondażach, które miały znaczący wpływ na codzienne życie.

Hoover Kenneth R., *The Elements of Social Scientific Thinking*, St. Martin's Press, New York 1992. Hoover przedstawia wspaniały przegląd kluczowych elementów analizy w naukach społecznych.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

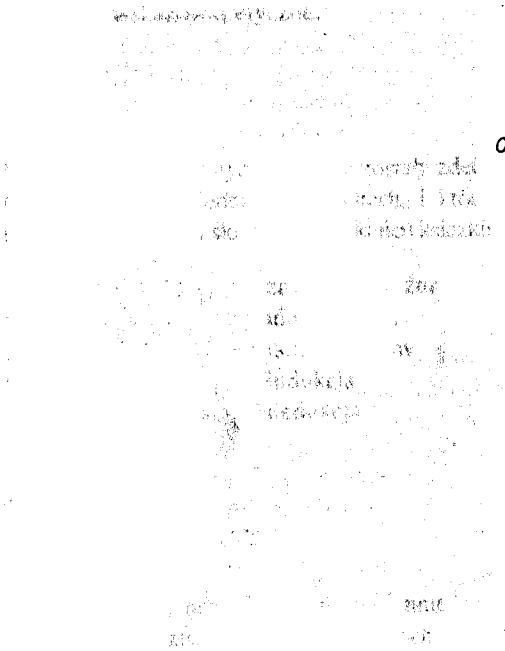
Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział
2

Paradygmaty, teoria i badania społeczne



Przegląd holograficzny

Praktyka nauk społecznych to współgranie teorii i badań, logiki i obserwacji, indukcji i dedukcji – oraz podstawowych układów odniesienia znanych jako paradygmaty.

Wprowadzenie

Kilka paradygmatów nauk społecznych
Makroteoria i mikroteoria
Wczesny pozytywizm
Darwinizm społeczny
Paradygmat konfliktu
Symboliczny interakcjonizm
Etnometodologia
Funkcjonalizm strukturalny
Paradygmaty feministyczne
Racjonalna obiektywność raz jeszcze

Elementy teorii społecznej

Powrót do dwóch systemów logicznych
Tradycyjny model nauki
Rozumowanie dedukcyjne i indukcyjne – zilustrowane na przykładzie
Kontrast ujęty graficznie

Budowa teorii dedukcyjnej
Zaczynamy
Tworzenie własnej teorii
Przykład teorii dedukcyjnej: sprawiedliwość dystrybutywna

Budowa teorii indukcyjnej
Przykład teorii indukcyjnej: dlaczego ludzie palą marihuanę?

Związki między teorią a badaniami

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

W Stanach Zjednoczonych są restauracje, które z upodobaniem przeprowadzają sondaże wśród swoich gości, gdy tylko zbliżają się wybory. Niektóre z nich traktują to bardzo poważnie, ponieważ wielokrotnie w tajemniczy sposób udało im się przewidzieć zwycięzców. Są kina, które osiągnęły podobny sukces, proponując popcorn w torebkach z rysunkiem osła albo słonia[1]. Przed wielu laty w spichlerzach środkowego Zachodu dano farmerom sposobność wskazania swoich preferencji politycznych. W tym celu mieli oni wybierać określone worki ze zbożem.

Wszystkie takie interesujące, choć dość specyficzne sposoby określania trendów spotyka po pewnym czasie ten sam los: sprawdzają się przez chwilę, potem zawodzą. Co gorsza, nie możemy przewidzieć, kiedy i dlaczego zawiodą.

Te niecodzienne techniki sondażowe uwydatniają znaczącą wadę „badań", których „wyniki" oparte są wyłącznie na obserwacji prawidłowości. Dopóki nie jesteśmy w stanie zaproponować logicznego wyjaśnienia tych prawidłowości, dopóty obserwowany przez nas wzorzec może być tylko wynikiem szczęśliwego trafu czy zbiegu okoliczności. Jeśli wystarczająco długo będziesz rzucał monetą, otrzymasz wreszcie orzełka dziesięć razy pod rząd. Naukowcy, aby opisać tę sytuację, mogliby zaadaptować nieco rynsztokowe wyrażenie „Każdemu się zdarza wdepnąć w prawidłowość"[2].

Logiczne wyjaśnienia są właśnie tym, czego dostarczyć mają teorie. W badaniach teorie pełnią trojaką funkcję. Po pierwsze, zapobiegają sytuacji, w której damy się nabrać szczęśliwym trafom. Jeśli nie potrafimy wyjaśnić, dlaczego jedna z restauracji osiągnęła takie sukcesy w przewidywaniu wyników wyborów, podejmujemy ryzyko, że staniemy się rzecznikami zbiegu okoliczności. Jeśli wiemy, dlaczego tak się stało, możemy przewidywać, czy zadziała to także w przyszłości.

Po drugie, teorie wyjaśniają zaobserwowane wzorce w sposób dopuszczający także inne możliwości. Jeśli rozumiemy, dlaczego przestępczość jest zjawiskiem częstszym wśród młodzieży z rozbitych rodzin niż wśród młodzieży z rodzin pełnych – na przykład: ze względu na brak nadzoru – możemy podjąć skuteczne działania, takie jak programy organizowania czasu młodzieży po szkole.

Po trzecie, teorie kształtują i ukierunkowują wysiłki badawcze, pokazując, jakich odkryć można by dokonać w wyniku obserwacji empirycznych. Szukając swoich zgubionych kluczy na ciemnej ulicy, mógłbyś światłem latarki jak batem chłostać na oślep okolicę, licząc, że na nie trafisz – albo mógłbyś przypomnieć sobie, gdzie konkretnie byłeś, aby ograniczyć rejon poszukiwań do najbardziej prawdopodobnych miejsc. Analogicznie, teorie kierują latarkę badacza tam, gdzie jest największa szansa zaobserwowania ciekawych prawidłowości życia społecznego.

Nie znaczy to, że wszystkie badania w naukach społecznych ściśle splecione są z teorią społeczną. Czasami badacze społeczni podejmują badania tylko po to, by ujawnić istniejący stan rzeczy – na przykład przeprowadzają badania ewaluacyjne, aby ocenić, czy innowacyjny program społeczny sprawdza się, albo sondaż, aby określić, który kandydat wygrywa w politycznym wyścigu. Podobnie etnografia opisowa, jak na przykład antropologiczne opisy społeczeństw przedpiśmiennych, dostarcza cennych informacji i daje wgląd w badane kwestie. A jednak nawet takie badania często wznoszą się ponad czysty opis, by spytać: *„dlaczego?"*. Teoria jest bezpośrednio związana z pytaniami „dlaczego".

W rozdziale tym zastanowimy się nad kilkoma specyficznymi sposobami, z jakich teoria i badania korzystają, idąc ramię w ramię na wyprawę zwaną zgłębianiem życia społecznego. Zaczniemy od przyjrzenia się pewnym zasadniczym układom odniesienia, zwanym *paradygmatami*, które leżą u podstaw teorii i badań społecznych.

---

[1] W USA słoń jest emblematem Partii Republikańskiej, osioł zaś żartobliwym emblematem Partii Demokratycznej (wg: Władysław Kopaliński, *Słownik symboli*, Wiedza Powszechna, Warszawa 1991, wyd. 2) (przyp. red. nauk.).

[2] W oryginale „patterns happen" (przyp. tłum.).

Kilka paradygmatów nauk społecznych

Zazwyczaj istnieje więcej niż jeden sposób rozumienia różnych spraw. W codziennym życiu na przykład liberałowie i konserwatyści często zupełnie inaczej wyjaśniają to samo zjawisko – powiedzmy, używanie broni przez nastolatków w szkołach. Podobnie zresztą rodzice i sami uczniowie. Lecz u podstaw tych różnych wyjaśnień czy teorii leżą **paradygmaty** – fundamentalne modele czy układy odniesienia, których używamy, by uporządkować nasze obserwacje i rozumowanie.

Czasem trudno jest rozpoznać paradygmaty, ponieważ są tak głęboko ukryte, milcząco zakładane, traktowane jako oczywiste. Przypominają bardziej stwierdzenie „tak już po prostu jest" niż jeden z wielu możliwych punktów widzenia. Oto przykład tego, co mam na myśli.

Jakie jest twoje stanowisko w kwestii praw człowieka? Czy uważasz indywidualne istoty ludzkie za świętość? Czy są one „obdarzone przez Stwórcę pewnymi niezbywalnymi prawami"[3], jak głosi Deklaracja Niepodległości Stanów Zjednoczonych? Czy są rzeczy, których żaden rząd nie powinien robić swoim obywatelom?

Przejdźmy do konkretów. W czasie wojny obywatele są czasem używani jako żywe tarcze, które mają ochronić obiekty wojskowe. Czasem siłą zmusza się ich do niewolniczej pracy lub wykorzystuje jako ruchome banki krwi dla szpitali wojskowych. Nie wspominając już o zorganizowanych, planowych gwałtach i morderstwach w służbie „czystek etnicznych".

Tym z nas, których takie praktyki przerażają i doprowadzają do wściekłości, trudno pewnie będzie uznać własny indywidualistyczny paradygmat za jeden z wielu możliwych punktów widzenia. A jednak zachodnie (a szczególnie wypracowane w USA) przywiązanie do świętości jednostki jest w dzisiejszym świecie w wielu innych kulturach uznawane za dziwaczne. Z historycznego punktu widzenia jest to zdecydowanie pogląd mniejszości.

Obecnie wiele krajów azjatyckich podpisuje się na przykład pod stwierdzeniem, że jednostki mają pewne „prawa". Jednak przeciwstawia się je „prawom" rodziny, organizacji oraz społeczeństwa jako całości. Krytykowani za gwałcenie praw człowieka azjatyccy przywódcy często wskazują na wysoki wskaźnik przestępczości i społecznej dezorganizacji w społeczeństwach Zachodu jako na cenę tego, co uznają za nasz skrajny „kult jednostki".

Nie będę próbował zmienić twojego stanowiska w sprawie jednostkowej godności człowieka, ale też nie rezygnuję ze swojego. Pożytecznie jest natomiast uznawać nasze poglądy i uczucia w tej sprawie za efekt działania paradygmatu, w ramach którego nas socjalizowano: nie są one obiektywnym zjawiskiem przyrodniczym. Wszyscy funkcjonujemy w ramach wielu takich paradygmatów. Przykładem głęboko zakorzenionego paradygmatu może być tradycyjne na Zachodzie postrzeganie obiektywnej rzeczywistości jako czegoś różnego od naszych indywidualnych doświadczeń.

Dwie korzyści płyną z uzmysłowienia sobie, że funkcjonujemy w ramach paradygmatu. Po pierwsze, jesteśmy w stanie łatwiej zrozumieć na pozór dziwne poglądy i działania tych, którzy funkcjonują w ramach innego paradygmatu. Po drugie, czasem przydaje się wychodzenie poza własny paradygmat. Nagle możemy ujrzeć nowe sposoby rozumienia i wyjaśniania rzeczy. Nie uda nam się to, póki mylnie bierzemy własny paradygmat za rzeczywistość.

Paradygmaty odgrywają w nauce, podobnie jak w codziennym życiu, zasadniczą rolę. Thomas Kuhn (1970) zwrócił uwagę na rolę paradygmatów w historii nauk przyrodniczych. Główne paradygmaty naukowe zawierają tak fundamentalne punkty widzenia jak kopernikańska koncepcja Ziemi poruszającej się wokół Słońca (zamiast na odwrót), teoria ewolucji Darwina, mechanika newtonowska i teoria względności Einsteina. Sensowność teorii naukowych zależy od tego, za jakim paradygmatem opowiadają się naukowcy.

Choć czasem myślimy, że nauka rozwija się stopniowo w czasie, naznaczana ważnymi od-

[3] *Deklaracja Niepodległości*, tłum. W. C. Borucki, w: W. C. Borucki, *Historia Stanów Zjednoczonych Ameryki Północnej*, Liberty Book Publishing Co., Detroit 1953 (przyp. tłum.).

kryciami i wynalazkami, Kuhn twierdzi, że na ogół paradygmaty naukowe obwarowują się, stawiając opór każdej znaczącej zmianie. A zatem zarówno teorie, jak i badania przyjmują pewien zasadniczy kierunek. Ostatecznie jednak, gdy wady danego paradygmatu stają się oczywiste, pojawia się nowy i zajmuje miejsce tamtego. Na przykład naturalny z pozoru pogląd, że wokół Ziemi obraca się reszta wszechświata, zmusił astronomów do obmyślenia znacznie bardziej skomplikowanych ruchów ciał niebieskich niż te, które w rzeczywistości obserwowali. W końcu paradygmat ten wyparty został przez pogląd, że Ziemia i inne planety obracają się wokół Słońca. Była to ni mniej, ni więcej tylko rewolucyjna zmiana perspektywy, która skierowała teorię i badania w zupełnie odmiennym kierunku. Klasyczna książka Kuhna dotycząca tego zagadnienia zatytułowana jest, całkiem odpowiednio, *Struktura rewolucji naukowych*.

Badacze społeczni stworzyli kilka paradygmatów rozumienia zachowań społecznych. W naukach społecznych jednak los wypieranych paradygmatów różnił się od tego, co Kuhn zaobserwował w naukach przyrodniczych. Przyrodoznawcy generalnie rzecz biorąc wierzą, że zastępowanie jednego paradygmatu przez inny odzwierciedla postęp w przechodzeniu od poglądów fałszywych ku prawdziwym. Na przykład żaden współczesny astronom nie wierzy, że Słońce obraca się wokół Ziemi.

Natomiast w naukach społecznych teoretyczne paradygmaty mogą zyskiwać lub tracić popularność, lecz rzadko zostają odrzucone w całości. Paradygmaty nauk społecznych oferują różne perspektywy. Każda z nich uwzględnia pewne aspekty społecznego życia nieobecne w pozostałych, a jednocześnie pomija wątki ujawniane gdzie indziej.

Ostatecznie, paradygmaty nie są ani prawdziwe, ani fałszywe; jako sposoby patrzenia są tylko mniej lub bardziej użyteczne. Każdy z paradygmatów, które zaraz omówimy, oferuje odmienny sposób patrzenia na życie społeczne człowieka. Każdy czyni pewne założenia co do natury rzeczywistości społecznej. Jak zobaczymy, każdy może otworzyć przed nami nowe możliwości rozumienia, zasugerować inny rodzaj teorii, być inspiracją dla innych rodzajów badań.

Makroteoria i mikroteoria

Zacznijmy od pewnego zróżnicowania dotyczącego obszarów zainteresowania teorii. Różnicę tę znajdziemy w obrębie wielu paradygmatów. Niektórzy teoretycy nauk społecznych skupiają swą uwagę na społeczeństwie jako całości albo przynajmniej na dużej jego części. Do zagadnień badawczych takiej **makroteorii** należy walka klas ekonomicznych w społeczeństwie, stosunki międzynarodowe czy wzajemne relacje między głównymi instytucjami w społeczeństwie, takimi jak władza, religia i rodzina. Makroteoria zajmuje się wielkimi, zagregowanymi bytami społecznymi albo nawet całymi społeczeństwami.

Inni znów uczeni patrzą na życie społeczne w sposób bardziej osobisty. **Mikroteoria** zajmuje się problemami życia społecznego na poziomie jednostek i małych grup. Tematy lepiej pasujące do perspektywy mikroteoretycznej to umawianie się na randki, obrady ławy przysięgłych, stosunki między studentami a wykładowcami. Badania takie często zbliżają się do dziedziny psychologii, ale jeśli psychologowie na ogół skupiają się na tym, co dzieje się wewnątrz człowieka, to socjologowie studiują to, co dzieje się między ludźmi.

Oś rozróżnienia makroteorii i mikroteorii przecina też różne paradygmaty, które przeanalizujemy. Niektóre z nich, takie jak symboliczny interakcjonizm czy etnometodologia, częściej ograniczają się do poziomu mikro. Inne, jak paradygmat konfliktu, mogą być stosowane zarówno w skali mikro, jak i makro.

Wczesny pozytywizm

Gdy francuski filozof Auguste Comte (1798–1857) w 1822 roku ukuł termin *socjologia*, rozpoczął intelektualną przygodę, która trwa do dziś. Co najważniejsze, Comte uznał społeczeństwo za zjawisko, które można badać naukowo. (Początkowo chciał nazwać swoje przedsięwzięcie „fizyką społeczną", ale termin ten przyjął już wcześniej inny uczony).

Przed czasami Comte'a społeczeństwo po prostu było. Ludzie rozpoznawali różne rodzaje

społeczeństw czy zmiany społeczeństw w czasie, ale na ogół dominującą rolę w wyjaśnianiu takich różnic odgrywały paradygmaty religijne. Stan spraw społecznych często był postrzegany jako odbicie Bożej woli. Bywało też, że ludzie czuli się wezwani do stworzenia „Miasta Boga" na ziemi, by zastąpić grzech i bezbożność.

Comte oddzielił swoje badania od religii. Czuł, że wiara religijna może zostać zastąpiona badaniami naukowymi i obiektywnością. Jego „filozofia pozytywna" postulowała trzy stadia rozwoju historii. Stadium „teologiczne" dominowało na całym świecie aż do, w przybliżeniu, 1300 roku. W ciągu kolejnych pięciuset lat stadium „metafizyczne" zastąpiło Boga ideami filozoficznymi, takimi jak „natura" i „prawo naturalne".

Comte myślał, że rozpoczyna trzecie stadium w historii, w którym nauka zastąpi religię i metafizykę, opierając wiedzę raczej na obserwacji za pomocą pięciu zmysłów niż na wierze czy na czystej logice. Czuł, że zjawiska społeczne mogą być obserwowane i wyjaśniane logicznie i racjonalnie oraz że socjologia może być tak naukowa, jak biologia czy fizyka.

W pewnym sensie wszystkie badania społeczne wywodzą się od Comte'a. Jego pogląd, że społeczeństwo można badać w sposób naukowy, dał podstawy do dalszego rozwoju nauk społecznych. Widząc optymistycznie przyszłość, Comte ukuł termin *pozytywizm*, aby opisać to naukowe podejście i przeciwstawić je temu, co uważał za negatywne strony Oświecenia. Jak zauważymy w dalszej dyskusji, dopiero w ostatnich dekadach ta idea pozytywizmu została poważnie zakwestionowana.

Darwinizm społeczny

Główne dzieło Comte'a poświęcone jego pozytywistycznej filozofii zostało wydane między 1830 a 1842 rokiem. Rok po opublikowaniu pierwszego tomu w tej serii pewien młody brytyjski przyrodnik wypłynął na okręcie Jego Królewskiej Mości nazwanym Psem Gończym (*Beagle*) w rejs, który miał zasadniczo zmienić nasz sposób myślenia o sobie i o naszym miejscu w świecie.

W 1858 roku Karol Darwin opublikow
O powstawaniu gatunków, gdzie przedstaw
ideę ewolucji poprzez proces doboru natura
nego. Upraszczając, teoria ta głosi, że gdy ga
tunek zmaga się ze swym środowiskiem natu
ralnym, jednostki najlepiej przystosowan
mają największe szanse przetrwać na tyl
długo, by się rozmnożyć. Jednostki gorze
przystosowane giną. Z czasem cechy tych, któ
re przetrwają, zaczynają dominować w danym
gatunku. Jak ujmowali to późniejsi darwiniści
gatunki w toku ewolucji osiągnęły różne form
dzięki „przetrwaniu najlepiej przystosowa
nych".

Skoro uczeni zaczęli badać społeczeństwo
w sposób analityczny, zastosowanie przez nich
idei Darwina do zmian w strukturze relacji mię
dzyludzkich było prawdopodobnie nieuchronne.
Drogę od prostych plemion zbieracko-łowiec
kich do wielkich cywilizacji przemysłowych
z łatwością postrzegano jako ewolucję coraz le
piej „przystosowanych" form społecznych.

Między innymi Herbert Spencer (1820–1903)
wywnioskował, że społeczeństwo staje się co
raz lepsze. Jego ojczyzna Anglia rzeczywiście
ogromnie zyskała na rozwoju kapitalizmu prze
mysłowego, a on popierał system wolnej kon
kurencji, która, jak sądził, zapewniała ciągły po
stęp i doskonalenie. Być może to nawet Spencer
ukuł termin „przetrwanie najlepiej przystoso
wanych". W każdym razie wierzył, że zasada ta
była pierwotną siłą kształtującą naturę społe
czeństwa. Społeczny darwinizm czy społeczny
ewolucjonizm był w czasach Spencera poglą
dem popularnym, choć nie uzyskał powszechnej
akceptacji.

Poniżej przedstawiony fragment z pod
ręcznika metod nauk społecznych, wydanego
w 1950 roku, ilustruje długotrwałą popular
ność koncepcji mówiącej, że jest coraz lepiej.

> „Wykorzystanie energii atomowej jako ma
> teriału wybuchowego stwarza bardzo interesu
> jące perspektywy zarówno w dziedzinie cywil
> nej, jak i wojskowej. Takie materiały wybuchowe
> mogą być wykorzystywane do przekształcania
> krajobrazu. Można ich używać do wysadzania
> i tworzenia wielkich dziur czy rowów, które na
> stępnie można zamienić w jeziora lub kanały

W ten sposób stanie się prawdopodobnie możliwe utworzenie jeziora w samym środku pustyni, a tym samym przemiana niektórych najgorszych miejsc na świecie w oazy i krainy urodzaju. Dzięki dostarczeniu stałego i nieograniczonego źródła ciepła będzie też prawdopodobnie możliwe uczynienie miejscem przyjemnym rejonów arktycznych. Biegun Północny mógłby zostać przekształcony w wakacyjny ośrodek wypoczynkowy".

(Gee 1950, s. 339–340)

Pominąwszy powszechne rozczarowanie energią jądrową, współczesne zaniepokojenie globalnym ociepleniem i zagrożeniem podnoszącym się poziomem mórz odzwierciedla rosnącą świadomość, że często „postęp", jak kij, ma dwa końce. Najwyraźniej dziś większość z nas funkcjonuje w ramach innego paradygmatu.

Paradygmat konfliktu

Jeden z autorów współczesnych Spencerowi przyjął drastycznie odmienny pogląd na ewolucję kapitalizmu. Karol Marks (1818–1883) zasugerował, że zachowania społeczne mogą być postrzegane jako proces, w którym stale dochodzi do konfliktu – prób zdominowania innych i uniknięcia cudzej dominacji. Marks skupił się przede wszystkim na walce między klasami ekonomicznymi. W szczególności gruntownie badał sposób, w jaki kapitalizm prowadzi do ucisku robotników przez właścicieli przedsiębiorstw przemysłowych. Zainteresowanie Marksa tym zagadnieniem nie kończyło się na analizie. Był on także ideologicznie zaangażowany w przekształcanie związków ekonomicznych, które położyłoby kres obserwowanemu przezeń uciskowi.

Kontrast między poglądami przedstawianymi przez Spencera i Marksa pokazuje, jaki jest wpływ paradygmatów na badania. Te podstawowe punkty widzenia nadają kształt typom obserwacji, które podejmujemy, rodzajom faktów, które próbujemy odkryć, i wnioskom, jakie z nich wyciągamy. Paradygmaty pomagają również określić, które pojęcia uznajemy za ważne i adekwatne. Na przykład w analizie Marksa zasadniczym pojęciem była klasa ekonomiczna, natomiast Spencer bardziej interesował się związkiem między jednostkami a społeczeństwem – a szczególnie tym, ile wolności jednostki muszą poświęcić, by społeczeństwo mogło funkcjonować.

Paradygmat konfliktu okazał się twórczy poza dziedziną czysto ekonomicznych analiz. W przeciwieństwie do walki klas, która zajmowała Marksa, Georga Simmla (1858–1918) szczególnie zainteresowały konflikty na małą skalę. Simmel zauważył na przykład, że wśród członków grup o silnych więzach konflikty miewają na ogół większe natężenie niż wśród ludzi, którzy nie podzielają uczucia przynależności i bliskości.

Paradygmat konfliktu zastosowany został niedawno w przeprowadzonej przez Michaela Chossudovsky'ego (1997) analizie Międzynarodowego Funduszu Walutowego (MFW) i Banku Światowego. Zasugerował on, że te dwie międzynarodowe organizacje zamiast zwalczać światowe ubóstwo raczej je zwiększają, i skierował swą uwagę na rywalizujące w tym procesie interesy. Teoretycznie, głównym interesem, którym miały się kierować te instytucje, był interes ludzi ubogich na całym świecie albo biednych narodów Trzeciego Świata. Badania jednak wykazały, że istnieje wiele innych zainteresowanych stron, które odnoszą korzyści: na przykład komercyjne instytucje kredytowe udzielające pożyczek w porozumieniu z MFW i Bankiem Światowym oraz ponadnarodowe korporacje szukające taniej siły roboczej i rynków zbytu dla swych dóbr. Chossudovsky w swojej analizie dochodzi do wniosku, że interesy banków i korporacji zaczęły przeważać nad interesami ubogich, którzy w zamierzeniu mieli być tu beneficjentami. Ponadto odkrył, że wiele programów osłabiało gospodarkę narodową w krajach Trzeciego Świata oraz podważało demokratyczne rządy.

Choć paradygmat konfliktu często koncentruje się na walce klasowej, walce etnicznej czy walce płci, właściwe byłoby jego stosowanie zawsze, gdy różne grupy mają przeciwstawne interesy. Na przykład można by go w sposób twórczy zastosować do wyjaśniania związków między różnymi wydziałami jakiejś organizacji albo

między studentami, wykładowcami i administracją, czy też do analizy napięć między kobiecymi a męskimi stowarzyszeniami studenckimi – by wymienić tylko kilka możliwości.

Symboliczny interakcjonizm

W swych ogólnych zainteresowaniach Georg Simmel różnił się zarówno od Spencera, jak i od Marksa. Tamci skupiali się głównie na problemach makroteoretycznych – wielkich instytucjach i całych społeczeństwach, ich ewolucji w toku historii, Simmel zaś zajął się raczej tym, jak jednostki wchodzą we wzajemne interakcje. Innymi słowy, jego myśl i badania zwróciły się ku skali „mikro". Skupił się na tych aspektach rzeczywistości społecznej, których nie dostrzeżemy u Marksa czy Spencera. Na przykład zaczął od badania diad (grup dwuosobowych) i triad (grup trzyosobowych). Pisał też o „sieci wewnątrzgrupowych powiązań".

Simmel był jednym z pierwszych socjologów europejskich, którzy oddziaływali na rozwój socjologii amerykańskiej. Jego koncentrowanie się na naturze interakcji miało szczególny wpływ na George'a Herberta Meada (1863–1931), Charlesa Hortona Cooleya (1864–1929) oraz innych, którzy podjęli tę problematykę i przekształcili ją we wpływowy paradygmat dla badań.

Cooley na przykład wprowadził pojęcie „grupy pierwotnej" dotyczące członków tych bliskich nam grup, z którymi dzielimy poczucie przynależności; chodzi tu o naszą rodzinę, przyjaciół itd. Pisał także o „jaźni odzwierciedlonej", którą kształtujemy, patrząc na reakcje ludzi wokół nas. Jeśli wszyscy traktują nas na przykład jako urodziwych, wnioskujemy, że tacy jesteśmy. Zauważmy, jak zasadniczo różnią się pojęcia i teoria zainspirowane tym paradygmatem od zainteresowań Spencera i Marksa, dotyczących poziomu społeczeństw.

Mead podkreślał wagę ludzkiej zdolności do „przyjmowania postaw innych" (*taking the role of the other*)[4], wyobrażania sobie, co czują inni i jak mogliby się zachować w pewnych okolicznościach. Gdy dowiadujemy się, jak ludzie ogólnie postrzegają sprawy, rozwijamy poczucie tego, co Mead nazwał „uogólnionym innym" (*the generalized other*)[5].

Mead wykazywał też wyjątkowe zainteresowanie rolą komunikowania się w relacjach międzyludzkich. Sądził, że większość interakcji obejmuje proces, w którym jednostki osiągają wzajemne porozumienie poprzez użycie języka i innych podobnych systemów, stąd termin *symboliczny interakcjonizm*.

Paradygmat ten pozwala nam zrozumieć naturę interakcji w zwyczajnym życiu społecznym, ale również niezwykłe formy interakcji, co pokazuje następujący przykład. Emerson, Ferris i Gardner (1998) zamierzali zrozumieć naturę „molestowania". Dzięki wywiadom z licznymi ofiarami molestowania, udało im się zidentyfikować różne motywacje molestujących, stadia w rozwijaniu się scenariusza molestowania, odpowiedzieć na pytanie, jak ludzie orientują się, że są molestowani i co mogą w związku z tym zrobić.

Etnometodologia

Nie wszystkie paradygmaty nauk społecznych podkreślają oddziaływanie struktury społecznej na zachowanie człowieka – tzn. efekt działania norm, wartości, czynników kontroli itd. Harold Garfinkel, współczesny socjolog, twierdzi, że ludzie ciągle tworzą strukturę społeczną poprzez swoje działania i interakcje, że w istocie tworzą swoje rzeczywistości.

Zatem kiedy spotykasz się ze swoim wykładowcą, aby omówić pracę zaliczeniową, nawet jeżeli istnieje niezliczone mnóstwo reguł dotyczących tego, jak powinniście się zachowywać, to wasza rozmowa będzie się trochę różnić od poprzedniej, a to, jak się zachowacie, w pewien sposób zmodyfikuje wasze oczekiwania w przyszłości. Znaczy to, że rozmowa na temat twojej pracy zaliczeniowej będzie oddziaływać na wasze przyszłe interakcje z innymi profesorami i studentami.

[4] G. H. Mead, *Umysł, osobowość i społeczeństwo*, tłum. Z. Wolińska, Warszawa 1975, s. 193–194 (przyp. tłum.).

[5] G. H. Mead, *Umysł*..., s. 214–216 (przyp. tłum.).

Uwzględniając tymczasowość rzeczywistości w tym ujęciu, Garfinkel sugeruje, że ludzie nieprzerwanie próbują zrozumieć życie, którego doświadczają. W pewnym sensie sugeruje, że każdy działa jak badacz społeczny, stąd też termin *etnometodologia*, czyli inaczej „metodologia ludowa".

W jaki sposób zabrałbyś się do poznawania ludzkich oczekiwań i tego, jak ludzie próbują zrozumieć swój świat? Jedną z technik stosowanych przez etnometodologów jest łamanie reguł, gwałcenie ludzkich oczekiwań. A więc jeżeli próbujesz rozmawiać ze mną o pracy zaliczeniowej, a ja wciąż mówię o piłce nożnej, to sytuacja ta może ujawnić, czego oczekujesz od moich zachowań. Moglibyśmy też zobaczyć, w jaki sposób nadajesz im sens. („Pewnie posługuje się piłką nożną jako analogią dla wyjaśnienia teorii systemów społecznych").

Weźmy inny przykład ujęcia etnometodologicznego. Johen Heritage i David Greatbatch (1992) zbadali rolę aplauzu w przemówieniach politycznych w Wielkiej Brytanii. Jak przemawiający wywoływali aplauz i jaką funkcję on pełnił (czy na przykład przesądzał sprawę)? Badania w ramach paradygmatu etnometodologicznego często skupiają się na komunikacji.

Możliwości wypróbowania paradygmatu etnometodologicznego jest nieskończenie wiele. Na przykład następnym razem, gdy wsiądziesz do windy, spędź całą drogę patrząc na jej tylną ścianę. Nie patrz ani na przód windy, ani na przeskakujące numery (to jest normalne, oczekiwane zachowanie). Po prostu stój cicho obrócony twarzą do tylnej ściany. Zobacz, jak inni zareagują na twoje zachowanie. I, co równie ważne, zauważ, jak sam się będziesz wtedy czuł. Jeśli wykonasz ten eksperyment kilkakrotnie, powinieneś zacząć rozwijać w sobie szczególną wrażliwość, charakterystyczną dla etnometodologicznego paradygmatu*.

Powrócimy do etnometodologii w rozdziale 10, w którym omówimy badania terenowe.

*Jestem wdzięczny mojemu koledze, Bernardowi McGrane, za ten eksperyment. Barney kazał też swym studentom jeść obiad palcami, oglądać telewizję nie włączając telewizora i angażować się w inne zadania w przedziwny sposób pouczające (McGrane 1994).

Na razie zwróćmy się ku zupełnie innemu paradygmatowi.

Funkcjonalizm strukturalny

Funkcjonalizm strukturalny, znany niekiedy także jako „teoria systemów społecznych", wyrasta z pojęcia wprowadzonego przez Comte'a i Spencera: byt społeczny, taki jak organizacja albo całe społeczeństwo, może być postrzegany jako organizm. Podobnie jak inne organizmy, system społeczny składa się z części, z których każda przyczynia się do funkcjonowania całości.

Zastanów się nad analogią do ludzkiego ciała. Każda jego część – taka jak serce, płuca, nerki, skóra i mózg – ma do wykonania określone zadanie. Ciało jako całość nie może przetrwać, jeżeli każda z tych części swego zadania nie wykona, a żadna z części nie może przetrwać inaczej niż jako część całego ciała. Albo zastanów się nad samochodem. Składa się z kół, kierownicy, baku, świec zapłonowych itd. Każda z części pełni jakąś funkcję w służbie całości; połączone w system, umożliwiają nam przejażdżkę po mieście. Żadna jednak sama w sobie nie byłaby dla nas zbyt użyteczna.

Spojrzenie na społeczeństwo jako na system społeczny szuka zatem „funkcji", jakie pełnią jego różne składniki. Badacze społeczni, stosujący paradygmat funkcjonalizmu strukturalnego, mogą zauważyć, że na przykład funkcją policji jest sprawowanie społecznej kontroli – zachęcanie ludzi do tego, by trzymali się norm społecznych oraz doprowadzanie tych, którzy tego nie robią, do organów wymiaru sprawiedliwości. Zauważ jednak, że równie dobrze można by spytać o to, jaką funkcję pełnią w społeczeństwie przestępcy. W ramach paradygmatu funkcjonalistycznego moglibyśmy rzec, że kryminaliści służą jako zapewnienie pracy policji. Nawiązując do tego, Émile Durkheim (1858–1917) zasugerował, że przestępstwa i ich karanie dostarczają okazji, by potwierdzić społeczne wartości. Chwytając i karząc złodziei, potwierdzamy nasz zbiorowy szacunek dla własności prywatnej.

Aby zrozumieć paradygmat strukturalno-funkcjonalistyczny, przyjmijmy, że interesuje

cię, jak działa twoja uczelnia. Mógłbyś przekartkować katalog tej instytucji i sporządzić listę pracowników administracji oraz obsługi (czyli rektora, dziekanów, sekretarza uniwersytetu, ochrony campusu, pracowników zajmujących się wyżywieniem). Mógłbyś potem oszacować, co każdy z nich robi, oraz powiązać ich role i czynności z głównymi funkcjami twojej uczelni, takimi jak nauczanie czy przeprowadzanie badań. Ten sposób patrzenia na instytucję edukacyjną wyraźnie sugeruje inny rodzaj badań niż, powiedzmy, paradygmat konfliktu, który podkreśliłby zderzenie interesów ludzi mających w tej instytucji władzę i tych, którzy jej nie mają.

Ludzie często dyskutują o „funkcjach" w codziennych rozmowach. Zazwyczaj jednak te domniemane funkcje rzadko są testowane empirycznie. Niektórzy przekonują na przykład, że opieka społeczna, której zamiarem jest pomoc ubogim, w rzeczywistości szkodzi im na różne sposoby. Twierdzi się czasem, że opieka społeczna tworzy dewiacyjne subkultury używające przemocy i skonfliktowane z głównym nurtem w społeczeństwie. Z tego punktu widzenia rzeczywistą konsekwencją programów opieki społecznej jest wzrastający wskaźnik przestępczości.

Lance Hannon i James Defronzo (1998) postanowili sprawdzić to ostatnie twierdzenie. Pracując nad danymi zebranymi w 406 zurbanizowanych hrabstwach w USA, zbadali związek między poziomem wypłat z opieki społecznej a wskaźnikiem przestępczości. W przeciwieństwie do przekonań niektórych, ich dane wskazują, że wyższe wypłaty z opieki społecznej wiązały się z niższymi wskaźnikami przestępczości. Innymi słowy, programy opieki społecznej pełniły raczej funkcję ograniczania niż nasilania bezprawia.

Paradygmaty feministyczne

Kiedy Ralph Linton, formułując wnioski swego klasycznego antropologicznego dzieła, *The Study of Man* (1937), mówił o „zasobach wiedzy, które obiecują dać człowiekowi (*man*) życie lepsze od tego, jakie kiedykolwiek znał", nikt nie narzekał, że pominął kobiety. Linton korzystał z językowych konwencji swoich czasów, włączał kobiety milcząco we wszystkich swych odniesieniach do mężczyzn. Czy rzeczywiście?

Gdy feministki po raz pierwszy zaczęły się przeciwstawiać używaniu męskich zaimków i rzeczowników wszędzie tam, gdzie płeć była nieokreślona, ich zaniepokojenie uważano często za małostkowe, nawet głupie. Na ogół sądzono, że w sprawie tej chodzi o zranione kobiece uczucia i posiniaczone kobiece ego. Ale bądźmy szczerzy: kiedy przeczytałeś słowa Lintona, co sobie wyobraziłeś? Amorficzną, bezpłciową istotę ludzką, hermafrodytę – jednocześnie mężczyznę i kobietę, czy też osobę płci męskiej?

Podobnie jak poprzednio, również badacze patrzący na świat od strony paradygmatu feministycznego zwrócili uwagę na te aspekty życia społecznego, których nie ujawniają inne paradygmaty. Częściowo teorie i badania feministyczne skupiały się na różnicach między płciami oraz na tym, w jaki sposób wiążą się one z resztą społecznej organizacji. Zwróciły uwagę na ucisk kobiet w wielu społeczeństwach, co z kolei rzuciło światło na samo zjawisko ucisku.

Paradygmat feministyczny zmierzył się też z powszechnie używanymi pojęciami dotyczącymi konsensusu w społeczeństwie. Większość opisów dominujących przekonań, wartości i norm społecznych jest stworzona przez ludzi reprezentujących tylko część społeczeństwa. Na przykład w USA takie analizy są zazwyczaj pisane przez białych mężczyzn należących do klasy średniej – nic dziwnego, że ich autorzy opisali przekonania, wartości i normy, które sami podzielali. I podczas gdy George Herbert Mead mówił o „uogólnionym innym", którego każdy z nas sobie uświadamia i którego „postawę może przyjąć", to feministyczne paradygmaty w ogóle kwestionują istnienie takiego uogólnionego innego.

Mead używał przykładu nauki gry w baseball, aby zilustrować, jak dowiadujemy się o uogólnionym innym, tymczasem badania Janet Lever sugerują, że zrozumienie doświadczeń chłopców może niewiele powiedzieć nam o dziewczętach.

„Gry i zabawy dziewcząt są całkiem różne. Najczęściej są spontaniczne, wykorzystują wyobraźnię, są wolne od struktur czy reguł. Zajęcia, w których wyznaczona jest kolejność, jak na przykład skakanka, mogą się odbywać bez ustalonych wyraźnych celów. Dziewczęta mają znacznie mniej doświadczeń w międzyosobowej rywalizacji. Styl ich współzawodnictwa jest raczej pośredni, nie odbywa się twarzą w twarz, jest bardziej indywidualny niż oparty na przynależności do zespołu. Role przywódcze są odgrywane rzadko albo nie ma ich wcale".

(Lever 1986, s. 86)

Nasilająca się wśród badaczy społecznych świadomość generalnych różnic intelektualnych między kobietami a mężczyznami doprowadziła do tego, że psycholog Mary Field Belenky i jej współpracownicy zaczęli mówić o kobiecych sposobach poznawania (*Women's Ways of Knowing*) (1986). Wywiady pogłębione przeprowadzone z 45 kobietami doprowadziły do wyróżnienia przez badaczy pięciu perspektyw poznawczych, które powinny podważyć pogląd, że zdobywanie wiedzy jest banalne i proste.

Cisza: Niektóre kobiety, szczególnie w młodym wieku, czuły się izolowane od świata wiedzy, ich życie w przeważającym stopniu determinowane było przez autorytety zewnętrzne.

Otrzymana wiedza: W ramach tej perspektywy kobiety uważały się za zdolne do przyjmowania i zachowywania wiedzy, pochodzącej od zewnętrznych autorytetów.

Wiedza subiektywna: Ta perspektywa otwiera możliwość osobistej, subiektywnej wiedzy, włączającej intuicję.

Wiedza proceduralna: Niektóre kobiety czują, że doskonale opanowały sposoby zdobywania wiedzy za pomocą obiektywnych procedur.

Wiedza konstruowana: Autorzy opisują tę perspektywę jako „sytuację, w której kobiety postrzegają całą wiedzę jako kontekstualną, doświadczają siebie jako twórców wiedzy oraz cenią zarówno subiektywne, jak i obiektywne strategie poznawcze" (Belenky i in., 1986, s. 15).

„Wiedza konstruowana" jest szczególnie interesująca w kontekście paradygmatów. W ramach pozytywistycznego paradygmatu Comte'a nie znalazłoby się miejsce ani dla „subiektywnej wiedzy", ani dla pomysłu, że prawda może różnić się w zależności od kontekstu. Paradygmat etnometodologiczny z kolei pomieściłby w sobie obie te idee.

Racjonalna obiektywność raz jeszcze

Zaczęliśmy tę dyskusję o paradygmatach od stwierdzenia Comte'a, że społeczeństwo można badać racjonalnie i obiektywnie. Od jego czasów rozwój nauki i techniki, w połączeniu z relatywnym ograniczeniem przesądów, umieszczał racjonalność coraz bliżej centrum społecznego życia. Choć jednak racjonalność jest dla większości z nas tak fundamentalna, niektórzy uczeni współcześni stawiają ją pod znakiem zapytania.

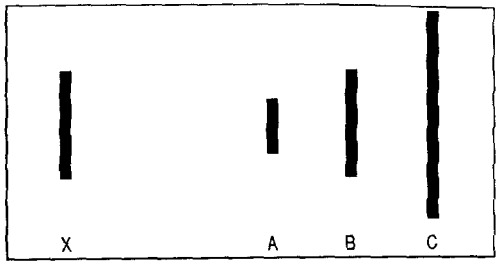
Na przykład pozytywistyczni badacze popełniali czasem błąd, przyjmując, że rzeczywistość społeczna może być wyjaśniona w racjonalnych kategoriach, ponieważ ludzie zawsze działają racjonalnie. Jestem pewien, że twoje własne doświadczenie służy ci wystarczającymi dowodami na rzecz tezy przeciwnej. Tymczasem wiele współczesnych modeli ekonomicznych zasadniczo zakłada, że ludzie będą dokonywać racjonalnych wyborów w sektorze ekonomicznym: wybiorą najlepiej płatną pracę, zapłacą najniższą cenę itd. Założenie to lekceważy siłę tradycji, lojalności, wizerunku i innych czynników, które rywalizują z rozumem i kalkulacją w określaniu ludzkiego zachowania.

Bardziej wyrafinowany pozytywista twierdziłby, że możemy racjonalnie zrozumieć i przewidzieć nawet nieracjonalne zachowanie. Przykładem tego jest słynny „eksperyment Ascha" (Asch 1958). W eksperymencie tym grupie badanych pokazany zostaje na ekranie zestaw linii, następnie prosi się ich, aby znaleźli dwie linie równej długości.

Wyobraź sobie siebie jako badanego w takim eksperymencie. Siedzisz w pierwszym rzędzie w sali, w grupie sześciu osób badanych. Zestaw linii wyświetlony zostaje na ścianie przed tobą (patrz rys. 2.1). Eksperymentator prosi każdego z was, po kolei, o rozpoznanie tej linii z prawej strony (A, B czy C), która odpowiada długości linii *X*. Właściwa odpowiedź

Rysunek 2.1

Eksperyment Ascha



(B) jest dla ciebie całkiem oczywista. Jednak, ku swemu zdziwieniu, dowiadujesz się, że wszyscy pozostali badani zgodnie wybierają inną odpowiedź!

Eksperymentator oświadcza, że wszyscy oprócz jednej osoby odpowiedzieli poprawnie. Ponieważ jesteś jedyną osobą, która wybrała B, *równa się to stwierdzeniu, że się pomyliłeś.* Następnie zaprezentowany jest nowy zestaw linii i doświadczasz tego samego. Odpowiedź, która wydaje się z oczywistych względów właściwa, uznana jest przez wszystkich pozostałych za błędną.

Jak się okazuje, rzecz jasna, jesteś jedynym prawdziwym badanym w tym eksperymencie – wszyscy inni współpracują z eksperymentatorem. Celem badania jest sprawdzenie, czy nacisk opinii publicznej spowoduje, że zgodzisz się na niewłaściwą odpowiedź. W swych pierwszych eksperymentach – do każdego z nich zaangażowani byli młodzi ludzie – Asch odkrył, że trochę więcej niż jedna trzecia badanych właśnie tak postąpiła.

Wybieranie w prostym eksperymencie odpowiedzi, która z oczywistych względów jest błędna, to przykład nieracjonalnego zachowania. Ale, jak miał to później pokazać Asch, eksperymentatorzy mogą zbadać okoliczności, które sprawiły, że mniej lub więcej badanych zgodziło się na niewłaściwą odpowiedź. Na przykład w kolejnych badaniach Asch różnicował rozmiar grupy i liczbę „odmieńców", którzy wybierali „złą" (tzn. dobrą) odpowiedź. A więc można badać nieracjonalne zachowanie w sposób racjonalny i naukowy.

Ujmując to bardziej radykalnie, możemy wątpić, czy życie społeczne w ogóle kieruje się racjonalnymi zasadami. W naukach przyrodniczych nowe idee, takie jak teoria chaosu, logika rozmyta[6] i układy złożone sugerują, że należy chyba przemyśleć w sposób zasadniczy problem uporządkowania zdarzeń świata fizycznego. Z pewnością ład świata społecznego może okazać się nie większy niż przyrodniczego.

Współczesne wyzwania dla pozytywizmu wykraczają poza pytanie, czy ludzie działają racjonalnie. Krytyka pozytywizmu podaje w wątpliwość m.in. myśl, że naukowcy mogą być tak obiektywni, jak to zakładał ideał pozytywistyczny. Większość naukowców zgodziłaby się, że osobiste uczucia mogą wpływać i że w istocie wpływają na wybór problemów badawczych, na wybór obserwacji i na wnioski, do jakich dochodzi się na podstawie owych obserwacji.

Istnieje jeszcze bardziej radykalna krytyka ideału obiektywności. Jak przelotnie o tym wspominaliśmy, omawiając feminizm i etnometodologię, niektórzy współcześni badacze sugerują, że subiektywność w pewnych sytuacjach może nawet być bardziej wskazana. Powróćmy na chwilę do dialektyki subiektywności i obiektywności.

Zacznijmy od tego, że wszystkie nasze doświadczenia są nieuchronnie subiektywne. Nie ma na to sposobu. Możemy patrzeć wyłącznie naszymi własnymi oczami i to, co w nich specyficzne, będzie nadawało kształt temu, co widzimy. Możemy słyszeć dźwięki tylko w taki sposób, w jaki nasze uszy przekazują, a mózg interpretuje fale dźwiękowe. Ty i ja do pewnego stopnia słyszymy i widzimy różne rzeczywistości. I doświadczamy fizycznych „rzeczywistości" całkiem różnych od tych doświadczanych przez, powiedzmy, nietoperze. W tym, co dla nas jest absolutną ciemnością, nietoperz, emitując dźwięk niesłyszalny dla człowieka, „widzi" takie rzeczy jak latające owady. Odbicie tego dźwięku tworzy „obraz dźwiękowy" wystarczająco dla nietoperza precyzyjny, by naprowadził go na poruszającego się owada i umożliwił schwytanie go zębami. Podobnie naukowcy na planecie

[6] Logika rozmyta (*fuzzy logic*) – logika zbiorów, do których przynależność jest określona nieprecyzyjnie, za pomocą funkcji prawdopodobieństwa; jedna z logik wielowartościowych (przyp. tłum.).

Xandu mogą rozwinąć teorie świata fizycznego oparte na aparacie zmysłowym, jakiego my, ludzie, nie możemy sobie nawet wyobrazić. Może widzą promienie Rentgena albo słyszą kolory.

Pomimo tej nieuchronnej subiektywności naszego doświadczenia my, ludzie, zdajemy się energicznie dążyć do porozumienia w kwestii tego, co rzeczywiście jest rzeczywiste, co obiektywnie jako takie istnieje. Obiektywność jest próbą pojęciowego wykroczenia poza nasze indywidualne poglądy. Ostatecznie jest to kwestia komunikacji, ponieważ próbujemy znaleźć wspólny grunt dla naszych subiektywnych doświadczeń. Gdy tylko poszukiwania się powiodą, mówimy, że mamy do czynienia z obiektywną rzeczywistością. To właśnie jest rzeczywistość umowy omówiona w rozdziale 1.

Nasza subiektywność jest indywidualna, natomiast nasze szukanie obiektywności jest społeczne. Potwierdza się to w każdym aspekcie życia, nie tylko w nauce. Choć wolimy inne rodzaje jedzenia, musimy w pewnym stopniu uzgodnić, co do jedzenia się nadaje, a co nie. W przeciwnym razie nie byłoby restauracji ani sklepów spożywczych. Ten sam argument można podać w odniesieniu do każdej innej formy konsumpcji. Bez rzeczywistości umowy nie byłoby kin, telewizji ani sportów.

Również badacze społeczni odkryli zalety pojęcia społecznie uzgodnionej, obiektywnej rzeczywistości. Ludzie, chcąc doświadczać życia w sposób uporządkowany, uznają za korzystne osiąganie tego celu w działaniu zbiorowym. Jakie są przyczyny uprzedzeń i lekarstwa na nie? Pracując razem, badacze społeczni znaleźli kilka odpowiedzi, które wytrzymały intersubiektywne sprawdzanie. Niezależnie od twojego subiektywnego doświadczenia możesz na przykład samodzielnie odkryć, że gdy wykształcenie wzrasta, uprzedzenia na ogół przyjmują tendencję spadkową. Mówimy, że jest to prawda obiektywna, ponieważ każdy z nas może ją odkryć niezależnie.

A jednak od XVII aż do połowy XX wieku w nauce dominowała wiara w rzeczywistość obiektywną niezależną od jednostkowego postrzegania. W przeważającym stopniu uważana była nie za użyteczny paradygmat, lecz za Prawdę. Termin *pozytywizm*, ogólnie rzecz biorąc, wyrażał wiarę w logicznie uporządkowaną, obiektywną rzeczywistość, którą dzięki nauce możemy poznawać coraz lepiej. Ten pogląd jest dziś podważany m.in. przez postmodernistów.

Niektórzy twierdzą, że ideał obiektywności zasłania równie wiele, co odsłania. Jak widzieliśmy wcześniej, w minionych latach większość tego, co w naukach społecznych na Zachodzie uznawane było za obiektywne, w istocie było zgodą panującą głównie między mieszkającymi w Europie i należącymi do klasy średniej białymi mężczyznami. Równie rzeczywiste doświadczenia, wspólne kobietom, mniejszościom etnicznym, kulturom innym niż zachodnia albo ludziom ubogim niekoniecznie znajdowały w tej rzeczywistości swoje odzwierciedlenie.

Pierwsi antropologowie są więc dziś krytykowani za częste „wyjaśnianie" wierzeń i praktyk niepiśmiennych plemion całego świata w kategoriach nowoczesnego Zachodu, co czasami oznaczało przedstawianie badanych jako przesądnych dzikich. Przekazywane ustnie przekonania dotyczące odległej przeszłości często nazywamy „mitem stworzenia", podczas gdy o naszych własnych przekonaniach mówimy: „historia". Obecnie upowszechnia się wymóg znajdowania lokalnej logiki, za pomocą której różne ludy nadają sens swemu życiu, i rozumienia jej w jej własnych kategoriach.

Ostatecznie nigdy nie będziemy w stanie w pełni odróżnić obiektywnej rzeczywistości od naszego subiektywnego doświadczenia. Nie możemy wiedzieć, czy nasze pojęcia odpowiadają obiektywnej rzeczywistości, czy są po prostu użyteczne, pozwalając nam przewidywać i kontrolować otoczenie. Nasza potrzeba poznania, co jest rzeczywiście rzeczywiste, przybiera jednak tak desperacką postać, że zarówno pozytywiści, jak i postmoderniści skłaniają się niekiedy ku przekonaniu, że ich własne poglądy są rzeczywiste i prawdziwe. Tkwi w tym podwójna ironia. Przede wszystkim pozytywistyczne przekonanie, że nauka precyzyjnie odzwierciedla obiektywny świat, musi w ostateczności opierać się na wierze; nie może go dowieść „obiektywna" nauka, ponieważ właśnie o niej się tu mówi. A postmoderniści, którzy twierdzą,

że nic nie jest obiektywne, a wszystko w ostateczności jest subiektywne, muszą przynajmniej czuć, że tak jest naprawdę.

Na szczęście, jako badacze społeczni, nie jesteśmy zmuszeni całkowicie opowiedzieć się za jednym z tych ujęć. Zamiast tego możemy traktować je jako dwie różne strzały w naszym kołczanie. Każde z nich rekompensuje słabości drugiego, proponując komplementarne perspektywy, które mogą wprowadzić badania na interesujące tory.

Podsumowując, na badania życia społecznego może wpływać bardzo wiele teoretycznych paradygmatów. W obrębie każdego z tych fundamentalnych układów odniesienia skonstruować można użyteczne teorie. Przejdziemy teraz do kilku zagadnień związanych z konstruowaniem teorii, którymi interesują się i posługują wszyscy badacze społeczni, począwszy od pozytywistów, a skończywszy na postmodernistach.

Elementy teorii społecznej

Jak widzieliśmy, paradygmaty to ogólne ramy czy punkty widzenia, dosłownie: „miejsca, z których się patrzy". Dostarczają sposobów patrzenia na życie i opierają się na zestawach założeń dotyczących natury rzeczywistości.

Natomiast teorie są systemami powiązanych wzajemnie twierdzeń mających wyjaśnić niektóre aspekty życia społecznego. Zatem teorie doprecyzowują i dookreślają paradygmaty. Paradygmat oferuje sposób patrzenia, celem teorii zaś jest wyjaśnienie tego, co widzimy.

Przyjrzyjmy się teraz trochę uważniej kilku elementom teorii. Jak wspomniałem w rozdziale 1, nauka opiera się na obserwacji. W badaniach społecznych *obserwacja* zazwyczaj odnosi się do patrzenia, słyszenia i – rzadziej – dotykania. Jest z tym powiązane pojęcie *faktu*. Choć dla filozofów „fakt" jest pojęciem równie złożonym co „rzeczywistość", badacze społeczni na ogół wykorzystują je, odnosząc do jakiegoś zjawiska, które zaobserwowano. Jest na przykład faktem, że Bill Clinton pokonał Roberta Dole'a w wyborach prezydenckich w 1996 roku.

Naukowcy dążą do tego, by uporządkować wiele faktów według „reguł" nazywanych *prawami*. Abraham Kaplan (1964, s. 91) zdefiniował *prawa* jako uniwersalne uogólnienia dotyczące klas faktów. Prawo ciążenia jest klasycznym przykładem: ciała przyciągają się wzajemnie proporcjonalnie do swych mas i odwrotnie proporcjonalnie do odległości, która je dzieli.

Prawa muszą jednak mieć naprawdę uniwersalny charakter, a nie być tylko przypadkowymi prawidłowościami odkrytymi dla specyficznego zbioru faktów. Jest faktem, jak wskazuje Kaplan (1964, s. 92), że w każdych wyborach prezydenckich w USA od 1920 do 1960 roku wygrywał główny kandydat mający najdłuższe nazwisko. Nie jest to jednak prawem, jak pokazały kolejne trzy elekcje. Wcześniejsza prawidłowość była zbiegiem okoliczności.

Prawa, nazywane czasem zasadami, to ważne stwierdzenia dotyczące tego, co istnieje. Mówimy o nich, że są „odkrywane", przyznając oczywiście, że nasze paradygmaty wpływają na wybór tego, czego chcemy szukać i na to, co widzimy. Prawa same z siebie nie wyjaśniają niczego. Są jedynie streszczeniem tego, jak się sprawy mają. Wyjaśnianie, jak wkrótce zobaczymy, jest funkcją teorii.

Nie ma praw w naukach społecznych, które miałby pretensje do uniwersalnej pewności nauk przyrodniczych. Badacze społeczni dyskutują między sobą, czy kiedykolwiek takie prawa mogą zostać odkryte. Być może życie społeczne ze swej natury nie podlega niezmiennym prawom. Nie znaczy to jednak, że jest ono tak chaotyczne, że opiera się przewidywaniu i wyjaśnianiu. Jak widzieliśmy w rozdziale 1, zachowania społeczne wykazują prawidłowości, a te prawidłowości bardzo często są doskonale zrozumiałe, choć bywa, że, aby odkryć ich logikę, musimy zajrzeć pod powierzchnię.

Jak właśnie wskazałem, praw nie należy mieszać z teoriami. Prawo jest zaobserwowaną prawidłowością, *teoria* natomiast jest systematycznym wyjaśnieniem obserwacji, które dotyczą jakiegoś konkretnego aspektu życia. Na przykład ktoś może zaproponować teorię przestępczości młodocianych, uprzedzeń czy rewolucji politycznej.

Teorie wyjaśniają obserwacje za pomocą pojęć. Jonathan Turner (*Theory Building...* 1989, s. 5) nazywa pojęcia „podstawowymi klockami teorii" (*basic building blocks of theory*). *Pojęcia* są abstrakcyjnymi elementami reprezentującymi klasy zjawisk z danej dziedziny badań. Na przykład pojęciami odpowiadającymi teorii przestępczości młodocianych będą na początek „młodociani" i „przestępczość". „Grupa rówieśnicza" – ludzie, z którymi się spotykasz i identyfikujesz – to kolejne adekwatne pojęcie. „Klasa społeczna" i „przynależność etniczna" są bez wątpienia pojęciami odpowiednimi dla teorii przestępczości wśród młodzieży. „Wyniki w szkole" mogą być również odpowiednie.

Zmienna jest szczególnym rodzajem pojęcia. Jak widzieliśmy w rozdziale 1, każda zmienna zawiera zestaw wartości, a zatem przestępczość, w najprostszym wydaniu, jest stworzona z cech: „popełnił przestępstwo" i „nie popełnił przestępstwa". Teoria przestępczości miałaby na celu wyjaśnienie, dlaczego część młodzieży popełnia przestępstwa, a część nie.

Aksjomaty albo *założenia* to podstawowe twierdzenia uznane za prawdziwe, na których opiera się teoria. W teorii przestępczości młodocianych zaczęlibyśmy od aksjomatów typu: „każdy pragnie żyć w dostatku" i „szansa bogatych na dostatnie życie zgodne z prawem jest większa niż szansa biednych". Stąd przeszlibyśmy do *tez* – szczegółowych wniosków dotyczących relacji między pojęciami, które wyciąga się z podstawowych założeń. Wychodząc ze wstępnych założeń dotyczących przestępczości młodocianych, moglibyśmy na przykład postawić sensowną tezę, że biedna młodzież częściej niż bogata łamie prawo, żeby umożliwić sobie życie w dostatku.

Teza ta zgadza się akurat z klasyczną próbą wyjaśnienia patologii społecznych dokonaną przez Roberta Mertona. Merton (1957, s. 139–157) mówił o celach i środkach uzgodnionych przez społeczeństwo. W jego modelu nie są dewiantami ci, którzy akceptują społeczną umowę dotyczącą pożądanych celów (takich jak nowy samochód) i środków przypisanych do ich osiągnięcia (poprzez kupno). Jeden typ dewianta – Merton nazywa go „innowatorem" – zgadza się co do pożądanych celów, ale nie ma dostępu do ustalonych środków ich osiągania. Innowatorzy szukają innej metody, na przykład przestępstwa, by osiągnąć pożądany cel.

Z tez możemy z kolei wyprowadzić *hipotezy*. **Hipoteza** jest szczegółowym, sprawdzalnym oczekiwaniem wobec rzeczywistości, które wynika z ogólniejszej tezy. Zatem badacz mógłby sformułować hipotezę: „Wśród biednej młodzieży wskaźnik przestępczości będzie wyższy niż wśród bogatej". Badanie jest tak zaplanowane, by sprawdziło tę hipotezę. Innymi słowy, badanie potwierdzi teorię (lub jej nie potwierdzi) tylko w sposób pośredni – sprawdzając konkretne hipotezy wyprowadzone z teorii i tez.

Przyjrzyjmy się uważniej, jak współzależą od siebie teoria i badania.

Powrót do dwóch systemów logicznych

W rozdziale 1 wprowadziłem pojęcia rozumowania dedukcyjnego i indukcyjnego, obiecując, że wrócimy do nich później. Oto i owo później.

Tradycyjny model nauki

Większość z nas ma trochę wyidealizowany obraz „metody naukowej", który, szczególnie dzięki naukom przyrodniczym, przekazywano nam w toku kształcenia, począwszy od szkoły podstawowej. Choć ten tradycyjny model nauki jest jednostronny, przydatne jest zrozumienie jego logiki.

W tradycyjnym modelu nauki istnieją trzy główne elementy: teoria, operacjonalizacja i obserwacja. Jesteśmy już teraz odpowiednio dobrze zaznajomieni z pojęciem teorii. Według tradycyjnego modelu nauki naukowcy zaczynają od czegoś, z czego wyprowadzają sprawdzalną hipotezę. A więc na przykład jako badacze społeczni moglibyśmy mieć teorię dotyczącą przyczyn przestępczości młodocianych. Przyjmijmy, że doszliśmy do hipotezy, zgodnie z którą przestępczość jest odwrotnie związana z klasą społeczną. Czyli im wyższa klasa społeczna, tym przestępczość mniejsza.

Aby przetestować jakąkolwiek hipotezę, musimy szczegółowo określić znaczenia wszystkich zmiennych związanych z nią w kolejnych obserwacjach. W omawianym przypadku te zmienne to *klasa społeczna* i *przestępczość*. By dać tym terminom szczegółowe znaczenie, moglibyśmy zdefiniować przestępczość jako „bycie aresztowanym za przestępstwo", „bycie skazanym za przestępstwo" albo w jakiś inny możliwy do przyjęcia sposób, podczas gdy klasę społeczną, dla celów tych konkretnych badań, można by określić w kategoriach dochodu gospodarstwa domowego.

Gdy już zdefiniowaliśmy zmienne, musimy ustalić, jak będziemy je mierzyć. (Przypomnijmy sobie stwierdzenie z rozdziału 1, że nauka według klasycznego ideału zależy od mierzalnych obserwacji). **Operacjonalizacja** dosłownie oznacza określenie dokładnych działań (operacji) związanych z pomiarem zmiennej. Istnieje wiele sposobów testowania naszej hipotezy, z których każdy dopuszcza różne sposoby pomiaru zmiennych.

Przyjmijmy dla uproszczenia, że zamierzamy przeprowadzić sondaż wśród uczniów szkół średnich. Moglibyśmy zoperacjonalizować przestępczość w formie pytania: „Czy kiedykolwiek coś ukradłeś?". Ci, którzy odpowiedzą „tak", zostaną w naszym badaniu zaklasyfikowani jako przestępcy; ci, którzy odpowiedzą „nie", będą zaklasyfikowani jako nie-przestępcy. Podobnie moglibyśmy zoperacjonalizować klasę społeczną pytając respondentów: „Jaki był dochód twojej rodziny w zeszłym roku?" i podając im zestaw kategorii dochodu rodziny: mniej niż 10 000 USD, 10 000–24 999 USD, 25 000–49 999 USD, oraz 50 000 USD i więcej.

W tym miejscu ktoś mógłby sprzeciwić się, mówiąc, że „przestępczość" może oznaczać coś innego lub coś więcej niż jedną czy drugą kradzież, albo że klasa społeczna niekoniecznie jest dokładnie tym samym, co dochód rodziny. Niektórzy rodzice mogą uważać, że przekłuwanie różnych części ciała jest oznaką zachowań przestępczych, nawet jeżeli ich dzieci nie kradną, a dla innych „klasa społeczna" może mieścić w sobie element prestiżu czy pozycji społecznej, tak samo ważny, jak ilość posiadanych przez rodzinę pieniędzy. Jednak dla badacza testującego hipotezę znaczenie zmiennych jest dokładnie takie i tylko takie, jak określa to **definicja operacyjna**.

Pod tym względem badacze bardzo przypominają Humpty Dumpty'ego z *Alicji w krainie czarów* Lewisa Carrolla. „Gdy *ja* używam jakiegoś słowa", mówi Humpty Dumpty do Alicji, „oznacza ono dokładnie to, co mu każę oznaczać... ni mniej, ni więcej".

„Pozostaje pytanie", odpowiada Alicja, „czy *potrafisz* nadawać słowom tak wiele rozmaitych znaczeń?". Na co Humpty Dumpty odpowiada, „Pozostaje pytanie, kto ma być panem... to wszystko"[7].

Naukowcy muszą być „panami" swych definicji operacyjnych przez wzgląd na precyzję ich obserwacji, pomiarów i porozumiewania się. W przeciwnym razie nigdy nie wiedzielibyśmy, czy jakieś badania zaprzeczyły naszym tylko dlatego, że wykorzystano w nich inny zestaw procedur do zmierzenia jednej ze zmiennych, co zmieniło znaczenie testowanej hipotezy. Oczywiście znaczy to również, że aby ocenić wnioski z badań dotyczących przestępczości młodocianych i klas społecznych, czy jakichkolwiek innych zmiennych, musimy dowiedzieć się, jak te zmienne zoperacjonalizowano.

Sposób, w jaki zoperacjonalizowaliśmy zmienne w naszych wymyślonych badaniach, mógłby jednak prowadzić do kolejnych problemów. Prawdopodobnie niektórzy respondenci będą kłamać w odpowiedzi na pytanie, czy coś ukradli, w których to przypadkach nie zaklasyfikujemy ich jako przestępców. Niektórzy nie będą znali dochodów swych rodzin i udzielą błędnych odpowiedzi; inni mogą się zawstydzić i skłamać. Szczegółowo omówimy takie problemy w części 2.

Zoperacjonalizowana przez nas hipoteza głosi teraz, że przestępczość najczęściej występować będzie wśród respondentów, którzy wybrali najniższą kategorię dochodu rodziny (poniżej 10 000 USD); niższy procent przestępców znajdziemy w kategorii 10 000–24 999 USD;

[7] Lewis Carroll, *O tym, co Alicja odkryła po drugiej stronie lustra*, tłum. M. Słomczyński, Wydawnictwo Zielona Sowa, Kraków 1998, s. 68 (przyp. tłum.).

jeszcze mniej przestępców będzie w kategorii 25 000–49 999 USD; a najniższy procent przestępców znajdzie się w kategorii 50 000 USD i więcej. Jesteśmy teraz gotowi na ostatni krok w tradycyjnym modelu nauki – obserwację. Gdy już osiągnęliśmy teoretyczną klarowność i doszliśmy do szczegółowych oczekiwań, gdy stworzyliśmy strategię obserwowania, pozostaje nam tylko popatrzeć, jak jest naprawdę.

Przypuśćmy, że nasz sondaż dał następujące wyniki:

	Procent przestępców
Poniżej 10 000 USD	20
10 000–24 999 USD	15
25 000–49 999 USD	10
50 000 USD i więcej	5

Obserwacje, które dostarczyłyby takich danych, potwierdziłyby naszą hipotezę. Ale przypuśćmy, że wyniki byłyby następujące:

	Procent przestępców
Poniżej 10 000 USD	15
10 000–24 999 USD	15
25 000–49 999 USD	15
50 000 USD i więcej	15

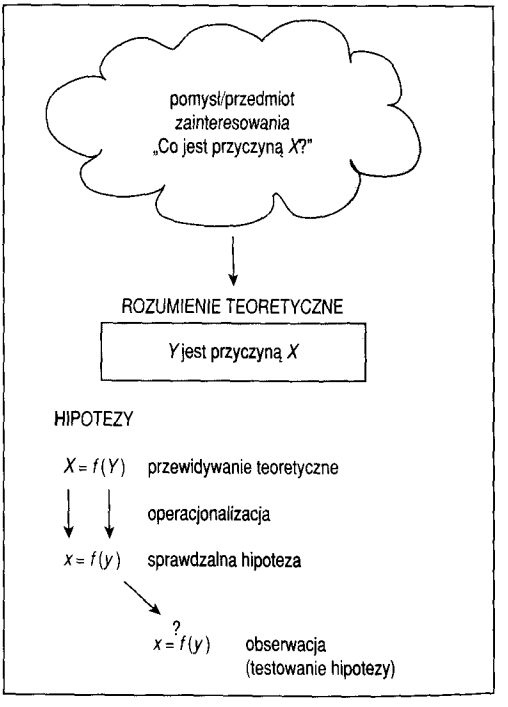
Takie wyniki obaliłyby naszą hipotezę dotyczącą dochodu rodziny i przestępczości. Obalalność – możliwość, że obserwacje mogą nie potwierdzić oczekiwań – jest zasadniczą własnością każdej hipotezy. Innymi słowy, jeżeli nie ma szans, że nasza hipoteza zostanie obalona, to znaczy, że nie mówi ona niczego sensownego.

Na przykład hipoteza, że „młodociani przestępcy" popełniają więcej przestępstw niż „nie-przestępcy", nie może ostatecznie zostać obalona, ponieważ zachowanie przestępcze zawiera się w pojęciu przestępczości. Nawet jeżeli zorientujemy się, że niektórzy młodzi ludzie popełniają przestępstwa, lecz nie zostają złapani i napiętnowani jako przestępcy, nie może to zagrozić naszej hipotezie, ponieważ nasze obserwacje i tak doprowadziłyby nas do wniosku, że są oni posłuszni prawu i nie są przestępcami.

Rysunek 2.2 pokazuje schemat tradycyjnego modelu badań naukowych. Widzimy na nim, że badacz wychodzi od zainteresowania jakimś zjawiskiem (takim jak przestępczość młodocianych). Następnie dąży do teoretycznego rozumienia, w tym przypadku tego, że pojedyncze pojęcie (takie jak klasa społeczna) może wyjaśniać inne. Efektem tych teoretycznych rozważań jest stwierdzenie, że jeśli teoria jest poprawna, powinny zostać zaobserwowane oczekiwane przez nas zjawiska. Zapis $X=f(Y)$ jest standardowym sposobem powiedzenia, że X (na przykład przestępczość) jest funkcją (zależy od) Y (na przykład klasy społecznej). Na tym jednak poziomie X i Y mają dość ogólne znaczenie, co może prowadzić do całkiem różnych obserwacji i pomiarów. Operacjonalizacja określa procedury, których użyje się przy pomiarze zmiennych. Na przykład na rysunku 2.2 mała litera x jest ściśle mierzalnym wskaźnikiem dużego X. Efektem tego procesu operacjonalizacji jest sformułowanie sprawdzalnej hipotezy: na przykład przyznanie, że się coś

Rysunek 2.2

Tradycyjne wyobrażenie nauki



ukradło, jest funkcją dochodu rodziny. Obserwacje mające przetestować, czy stwierdzenie to poprawnie opisuje rzeczywistość, są częścią tego, co zazwyczaj nazywa się *testowaniem hipotezy* (patrz ramka „Wskazówki, jak stawiać hipotezy", gdzie znajdziesz więcej uwag na temat procesu formułowania hipotez).

Rozumowanie dedukcyjne i indukcyjne – zilustrowane na przykładzie

Jak się być może zorientowałeś, opisany właśnie tradycyjny model nauki jest dobrym przykładem rozumowania dedukcyjnego: z ogólnego wyjaśnienia teoretycznego badacz wyprowadza (dedukuje) spodziewane zjawiska, a na koniec sprawdzalną hipotezę. Na obrazku wszystko jest wzorowo uporządkowane, lecz w rzeczywistości nauka posługuje się również rozumowaniem indukcyjnym. Rozważmy przykład prawdziwych badań, który posłuży nam do porównania dedukcyjnych i indukcyjnych powiązań między teorią a badaniami. Przed laty Charles Glock, Benjamin Ringer i ja (1967) zamierzaliśmy odkryć, co jest przyczyną zróżnicowań w poziomie zaangażowania religijnego członków amerykańskiego Kościoła Episkopalnego. Wiele teoretycznych czy quasi-teoretycznych perspektyw sugerowało możliwe odpowiedzi. Skupię się tutaj tylko na jednej, którą nazwaliśmy „hipotezą pocieszenia".

Jedną ze wskazówek był dla nas chrześcijański nakaz troszczenia się o „chromych, ułomnych i niewidomych" oraz tych, którzy są „znużeni i pogrążeni w smutku"[8]. Jednocześnie, jak na ironię, zwróciliśmy uwagę na stwierdzenie Marksa, że religia jest „opium dla mas". Uwzględniwszy to wszystko, oczekiwalibyśmy w naszej hipotezie, że „parafianie, których sytuacja życiowa w największym stopniu pozbawiała ich satysfakcji i spełnienia w społeczności świeckiej, będą się zwracać do kościoła po pocieszenie i nagrody zastępcze" (Glock i in. 1967, s. 107–108).

Gdy sformułowaliśmy tę ogólną hipotezę, zaczęliśmy ją testować. Czy osoby niespełnione w sferze świeckiej były istotnie bardziej religijne niż ci, którym sfera ta przynosiła więcej satysfakcji? By odpowiedzieć na to pytanie, potrzebowaliśmy wyróżnić osoby pozbawione satysfakcji. W celu przetestowania „hipotezy pocieszenia" skonstruowaliśmy kwestionariusz zawierający pytania, które wydawały się wskaźnikami tego, czy dla danego parafianina społeczność świecka była źródłem deprywacji, czy zaspokojenia.

Zaczęliśmy rozumowanie od tego, że w naszym zdominowanym przez płeć męską społeczeństwie mężczyźni w większym stopniu cieszą się prestiżem niż kobiety. Konkluzja ta, choć mało oryginalna, dała podstawę do przetestowania „hipotezy pocieszenia". Jeśli nasza hipoteza była poprawna, kobiety powinny okazać się bardziej religijne niż mężczyźni. Gdy dane sondażowe zostały zebrane i przeanalizowane, nasze oczekiwania dotyczące płci i religijności uzyskały wyraźne potwierdzenie. Według trzech oddzielnych miar zaangażowania religijnego – rytualnego (na przykład uczęszczanie do kościoła), organizacyjnego (na przykład przynależność do kościelnych organizacji) oraz intelektualnego (na przykład czytanie kościelnych wydawnictw) – kobiety były bardziej religijne niż mężczyźni. W naszym pomiarze ogólnym wynik kobiet był o 50% wyższy od wyniku mężczyzn.

W innym teście naszej „hipotezy pocieszenia" rozumowaliśmy, że w społeczeństwie ceniącym młodość ludzie starzy będą w większym stopniu pozbawieni satysfakcji w sferze świeckiej niż ludzie młodzi. Ponownie dane potwierdziły nasze oczekiwania. Najstarsi parafianie byli bardziej religijni niż ci w średnim wieku, a ci z kolei byli bardziej religijni od młodych.

Klasa społeczna – mierzona wykształceniem i dochodem – pozwoliła na następny test „hipotezy pocieszenia". I znów test się powiódł. Osoby o niskim statusie społecznym były bardziej zaangażowane w sprawy kościoła niż osoby o wysokim statusie.

Nasza hipoteza została nawet potwierdzona w teście, który stał w sprzeczności ze zdroworozsądkowymi oczekiwaniami każdego z nas. Pomimo kościelnych plakatów poka-

[8] Mt (15,30), Łk (14,13) (przyp. tłum.).

Wskazówki, jak stawiać hipotezy

Riley E. Dunlap

Wydział Socjologii, Washington State University

Hipoteza jest podstawowym twierdzeniem, które testuje się w badaniu. Zazwyczaj hipoteza stwierdza związek między dwiema zmiennymi. (Choć możliwe jest korzystanie z większej liczby zmiennych, na razie powinieneś trzymać się dwóch). Ponieważ hipoteza formułuje prognozy dotyczące związku między dwiema zmiennymi, musi być sprawdzalna, aby można było, gdy prześledzi się wyniki uzyskane w badaniu, określić, czy przewidywanie jest słuszne, czy nie. Aby można było dokładnie przetestować hipotezę, musi być ona sformułowana w sposób jednoznaczny. Poniżej znajdują się sugestie, jak tworzyć sprawdzalne hipotezy.

Przyjmijmy, że interesuje cię próba przewidywania takiego zjawiska jak „postawy wobec wyzwolenia kobiet" oraz, że możesz zmierzyć te postawy na następującym kontinuum: „przeciwny wyzwoleniu kobiet", „neutralny", „popierający wyzwolenie kobiet". Przyjmijmy też, że z braku teorii zmienne, które mogłyby być związane z postawami wobec wyzwolenia kobiet, znajdujesz polegając na „wyczuciu".

W pewnym sensie, możesz myśleć o konstruowaniu hipotezy jak o wypełnianiu pustych miejsc: „_____ wiąże się z postawami wobec wyzwolenia kobiet". Twoje zadanie polega na wymyśleniu zmiennej, która będzie się w sposób wiarygodny wiązała z takimi postawami, a potem na sformułowaniu hipotezy, która stwierdzi związek między tymi dwiema zmiennymi (zmienną, która wpasuje się w „puste miejsce" i „postawami wobec wyzwolenia kobiet"). Musisz zrobić to w sposób precyzyjny, abyś mógł jasno określić, gdy przeanalizujesz wyniki (w tym przypadku najprawdopodobniej wyniki sondażu), czy hipoteza została potwierdzona, czy nie.

Sprawą kluczową jest uważny dobór słów w hipotezie, dzięki czemu formułowane przewidywania są jasne zarówno dla ciebie, jak i dla innych. Jeśli wykorzystujesz wiek, zauważ, że stwierdzenie „wiek wiąże się z postawami wobec wyzwolenia kobiet" nie określa precyzyjnie, jak twoim zdaniem obie te zmienne są związane (właściwie jedynym sposobem na obalenie tej hipotezy byłoby nieznalezienie żadnego statystycznie istotnego związku między wiekiem a postawami wobec wyzwolenia kobiet). W tym przypadku konieczne jest podjęcie kilku kroków. Masz dwie opcje.

1. „Wiek wiąże się z postawami wobec wyzwolenia kobiet, przy czym wśród dorosłych osoby młodsze częściej niż starsze popierają wyzwolenie kobiet". (Albo można stwierdzić rzecz przeciwną, jeśli ktoś sądzi, że to u starszych poparcie jest większe).

2. „Wiek jest ujemnie związany z poparciem dla wyzwolenia kobiet". Zauważ, że wymieniam tu konkretnie „poparcie" dla wyzwolenia kobiet (PWK) i że przewiduję odwrotny związek – tzn.: im bardziej zaawansowany wiek, tym przewidywane PWK będzie mniejsze.

Trzeba zaznaczyć, że w tej hipotezie obie zmienne (wiek, zmienna niezależna albo prawdopodobna „przyczyna", oraz PWK, zmienna zależna albo prawdopodobny „skutek") uszeregowane są od wartości niskich do wysokich. Ta cecha obu zmiennych pozwala używać określeń „ujemny" (lub „dodatni") przy opisie omawianego związku.

Zauważ, co stanie się, jeśli hipoteza dotyczyć będzie związku między płcią a PWK. Ponieważ płeć (jak dowiemy się z rozdziału 5) jest zmienną nominalną, nie jest uszeregowana od wartości niskich do wysokich – ludzie są albo mężczyznami, albo kobietami (są to dwie wartości zmiennej *płeć*). W konsekwencji trzeba zachować ostrożność, aby stawiana hipoteza była sformułowana jednoznacznie:

1. „Płeć jest dodatnio (lub ujemnie) związana z PWK" nie jest właściwą hipotezą, ponieważ nie określa charakteru oczekiwanego związku między płcią a PWK, a więc czy uważasz, że poparcie dla wyzwolenia kobiet będzie większe wśród mężczyzn, czy raczej wśród kobiet.
2. Kusi nas, by powiedzieć coś w rodzaju: „Bycie kobietą jest dodatnio związane z PWK", ale naprawdę nie wchodzi to w grę, ponieważ bycie kobietą jest tylko jedną z wartości, a nie pełną zmienną (tą zmienną jest płeć).
3. Zalecałbym takie sformułowanie: „Płeć jest związana z PWK, przy czym poparcie wśród kobiet jest większe niż wśród mężczyzn". Można też powiedzieć: „przy czym poparcie u męż-

czyzn jest słabsze niż u kobiet", z czego wynikają identyczne przewidywania. (Oczywiście, gdyby ktoś chciał, można przewidywać rzecz odwrotną, że poparcie wśród mężczyzn jest większe niż wśród kobiet).

4. Tak samo uprawnione byłoby „U kobiet widać większą skłonność do popierania wyzwolenia kobiet niż u mężczyzn". (Zauważmy, jak potrzebne jest drugie „u", bez niego można by rozumieć, że według twojej hipotezy kobiety chętniej popierają wyzwolenie kobiet niż mężczyzn – myśl to nie całkiem ta sama).

W powyższych przykładach stawiane były hipotezy dotyczące związku między „cechą" („wiek" lub „płeć") a „orientacją" (postawami wobec wyzwolenia kobiet). Ponieważ porządek przyczynowy jest całkiem jasny (jest oczywiste, że wiek i płeć pojawiają się przed postawami, i że trudniej ulegają zmianie), moglibyśmy formułować hipotezy, jak ja to uczyniłem, i każdy doszedłby do wniosku, że są to hipotezy dotyczące przyczyn.

Można wreszcie natknąć się, głównie w statystyce, na odwołania do **hipotezy zerowej**. Hipoteza taka przewiduje, że nie istnieje żaden związek (technicznie ujmując, że nie istnieje żaden statystycznie istotny związek) między dwiema zmiennymi, i zawsze zakłada się to milcząco przy testowaniu hipotez. Zasadniczo, jeśli stworzyłeś hipotezę mówiącą o dodatnim (lub ujemnym) związku, masz nadzieję, że wyniki pozwolą ci odrzucić hipotezę zerową i potwierdzić hipotetyczny związek.

zujących pobożne młode rodziny i opatrzonych hasłem: „Rodzina, która modli się razem, pozostaje razem" – „hipoteza pocieszenia" sugerowała, że parafianie, którzy pobrali się i mieli dzieci – co było bezspornie amerykańskim ideałem tamtych czasów – osiągali świecką satysfakcję w tej dziedzinie. W konsekwencji powinni być oni mniej religijni niż ci, którym brakowało jednego lub obu elementów życia rodzinnego. Postawiliśmy więc hipotezę, że parafianie bezdzietni stanu wolnego będą najbardziej religijni; ci, którzy mają dzieci albo współmałżonka, będą trochę mniej religijni; a małżeństwa mające dzieci – reprezentujące ideał przedstawiany na każdym z tych plakatów – powinny być najmniej religijne ze wszystkich. Takie właśnie wyniki uzyskaliśmy.

Wreszcie, z „hipotezy pocieszenia" wynika też, że różne rodzaje deprywacji w sferze świeckiej powinny się kumulować. Osoby mające wszystkie cechy kojarzone z deprywacją powinny być najbardziej religijne; osoby bez żadnej z tych cech – najmniej. Gdy połączyliśmy cztery pojedyncze miary deprywacji w miarę łączną, teoretyczne przewidywania zostały całkowicie potwierdzone. Porównując obie skrajności, odkryliśmy, że niezamężne, bezdzietne, stare i pochodzące z niższej klasy parafianki osiągały trzy razy wyższe wyniki według miary zaangażowania kościelnego niż młodzi, żonaci i pochodzący z klasy wyższej ojcowie. Tak oto potwierdzona została „hipoteza pocieszenia".

Podoba mi się przykład tych badań, ponieważ tak przejrzyście ilustruje logikę modelu dedukcyjnego. Wyjście od ogólnych, teoretycznych wyjaśnień dotyczących wpływu społecznej deprywacji na zaangażowanie w sprawy kościoła umożliwiło wyprowadzenie konkretnych hipotez łączących konkretne mierzalne zmienne, takie jak wiek i uczęszczanie do kościoła. Następnie mogliśmy przeanalizować rzeczywiste dane empiryczne, by sprawdzić, czy te dedukcyjne oczekiwania znajdą potwierdzenie w rzeczywistości.

Opisując w przykładzie powyżej, jak zrealizowano to badanie, niestety oszukiwałem. Prawdę mówiąc, choć od początku interesowało nas odkrycie, co spowodowało zróżnicowanie zaangażowania wśród parafian Kościoła Episkopalnego, w rzeczywistości nie zaczęliśmy od „hipotezy pocieszenia" ani od żadnej innej. (Mówiąc całkiem szczerze, to Glock i Ringer zainicjowali te badania, a ja dołączyłem do nich po wielu latach od momentu zebrania danych). Kwestionariusz zbudowano pod kątem gromadzenia danych, które mogłyby rzucić trochę światła na problem zróżnicowanego zaangażowania parafian, lecz do jego skonstruowania nie prowadziła żadna konkretna te-

oria dedukcyjna. Gdy dane zostały zebrane, zadanie wyjaśnienia różnic w religijności rozpoczęto od analizy zmiennych, które miały znaczący wpływ na życie ludzi, wliczając w to płeć, wiek, klasę społeczną i status rodziny. Odkryto, że każda z tych czterech zmiennych silnie wiąże się z zaangażowaniem w sprawy kościoła, tak jak to przedstawiono powyżej. Ich działanie istotnie się kumulowało, jak to opisałem. Zamiast dobrych wieści oznaczało to dylemat.

Glock wspomina, jak dyskutował o tych wynikach z kolegami, jedząc lunch w klubie wydziałowym Uniwersytetu Columbia. Gdy pokazał tabele przedstawiające wpływ każdej pojedynczej zmiennej oraz ich silnego łącznego efektu, jeden z kolegów spytał: „Co to wszystko znaczy, Charlie?". Glock był zakłopotany. Dlaczego wszystkie te zmienne *były* tak bardzo związane z zaangażowaniem w sprawy kościoła?

Pytanie to uruchomiło proces rozumowania: co tych kilka zmiennych miało ze sobą wspólnego oprócz wpływu na religijność. W końcu dostrzegliśmy, że każda z nich odzwierciedlała także *zróżnicowany status w społeczności świeckiej*. Glock wpadł na myśl, że mogła się z tym wiązać sprawa „pocieszenia". A zatem proces indukcyjny przeszedł od konkretnych obserwacji do ogólnego wyjaśnienia teoretycznego.

Kontrast ujęty graficznie

Jak ilustruje to opisany wyżej przypadek, teorią i badaniami można z korzyścią zajmować się zarówno w drodze dedukcji, jak i indukcji. Rysunek 2.3 pokazuje porównanie graficzne obu tych podejść zastosowanych do badań nad zwyczajami uczenia się i wynikami z egzaminów. W obu przypadkach interesuje nas związek między liczbą godzin spędzonych na uczeniu się do egzaminu a stopniem uzyskanym z tego egzaminu. Korzystając z metody dedukcyjnej, zaczęlibyśmy od logicznej analizy tego problemu. Dobrze zdany egzamin odzwierciedla zdolność studenta do przypominania sobie informacji i posługiwania się nimi. Obie te zdolności powinny się wzmacniać dzięki kontaktowi z informacjami przed egzaminem. W ten sposób dochodzimy do hipotezy sugerującej dodatnią zależność między liczbą godzin spędzonych na uczeniu się a stopniem zdobytym na egzaminie. Mówimy, że zależność jest dodatnia, ponieważ oczekujemy, że stopnie będą się poprawiać wraz ze wzrostem liczby godzin uczenia się. Gdyby większa liczba godzin powodowała niższe stopnie, nazwalibyśmy to zależnością ujemną albo odwrotną. Hipoteza ta przedstawiona jest za pomocą linii w części 1(a) rysunku 2.3.

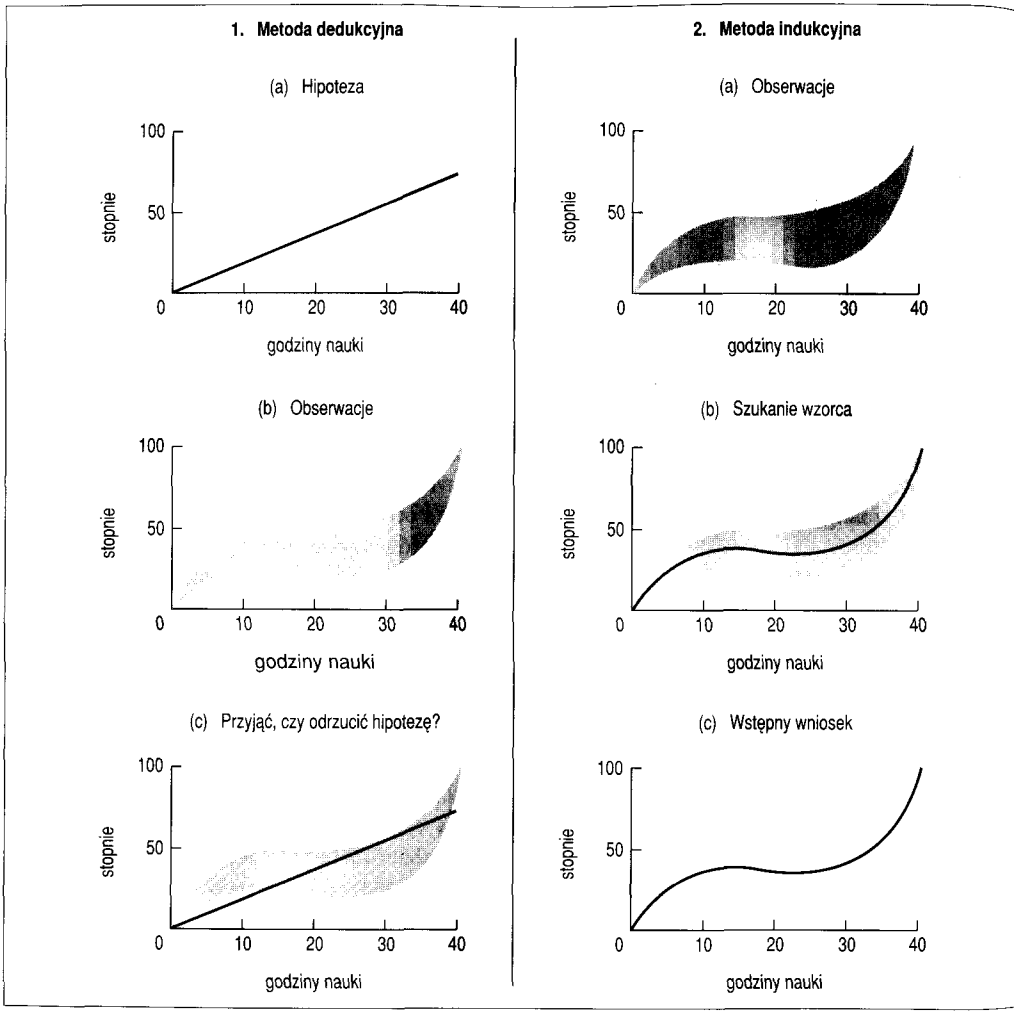
Naszym następnym krokiem będzie dokonanie odpowiednich obserwacji umożliwiających przetestowanie naszej hipotezy. Zacienione pole w części 1(b) na rysunku oznaczać może setki obserwacji przeprowadzonych na różnych studentach, dotyczących liczby godzin nauki i otrzymanych stopni. Wreszcie, w części 1(c) porównujemy hipotezę i obserwacje. Ponieważ obserwacje dokonane w realnym świecie rzadko, o ile w ogóle kiedykolwiek, w sposób doskonały zgadzają się z przewidywaniami, musimy zdecydować, czy zgodność jest wystarczająco duża, aby uznać hipotezę za potwierdzoną. Ujmując to inaczej, czy możemy wnioskować, że hipoteza ta opisuje istniejącą ogólną prawidłowość, dopuszczającą pewne odchylenia w rzeczywistym świecie? Czasami, by odpowiedzieć na to pytanie, konieczne jest zastosowanie metod analizy statystycznej, które omówione będą w części 4.

Przypuśćmy teraz, że odnosząc się do tego samego pytania stosujemy metodę indukcyjną. W tym przypadku zaczęlibyśmy od zbioru obserwacji, tak jak w części 2(a) na rysunku 2.3. Ciekawi związku między godzinami spędzonymi na uczeniu się a zdobytymi stopniami, moglibyśmy po prostu zadbać o zgromadzenie odpowiednich danych. Następnie szukalibyśmy wzorca, który najlepiej oddawałby czy podsumowywał nasze obserwacje. W części 2(b) na rysunku wzorzec ten pokazany jest jako krzywa biegnąca przez środek naszych obserwacji.

W tym przypadku odkryty wśród punktów wzorzec sugeruje, że gdy nauka trwa od 1 do 15 godzin, każda kolejna godzina przynosi na ogół wyższy stopień na egzaminie. Jednak jeśli

Rysunek 2.3

Metody dedukcyjne i indukcyjne



nauka trwa od 15 do 25 godzin, jej przedłużanie, jak się wydaje, odrobinę obniża stopień. Natomiast uczenie się dłużej niż przez 25 godzin powoduje powrót do pierwotnego wzorca: więcej godzin daje lepsze stopnie. Korzystając z metody indukcyjnej, kończymy wstępnym wnioskiem dotyczącym wzorca relacji między dwiema zmiennymi. Wniosek jest wstępny, ponieważ poczynionych obserwacji nie można uznać za test dla tego wzorca – stworzyliśmy go na ich podstawie.

Jak pisałem w rozdziale 1, w rzeczywistych badaniach teoria i badania są wzajemnie powiązane w niekończących się przejściach od indukcji do dedukcji i odwrotnie. Dobrym tego przykładem jest klasyczna praca Émile'a Durkheima poświęcona samobójstwom ([1897] 1951). Gdy Durkheim ślęczał nad kolejnymi tabelami urzędowych statystyk wskaźnika samobójstw w różnych regionach, uderzył go fakt, że kraje protestanckie nieodmiennie mają wyższy wskaźnik samobójstw niż kraje katolickie. Dlaczego tak miałoby być? Jego wstępne obserwacje sprawiły, że stworzył metodą indukcji teorię religii, integracji społecznej, anomii i samobójstwa. Z kolei jego wyjaśnienia te-

oretyczne metodą dedukcji poprowadziły do dalszych hipotez i obserwacji.

Podsumowując, naukowa norma logicznego rozumowania zaopatruje nas w pomost, o dwustronnym ruchu, między teorią a badaniami. W praktyce badania naukowe zazwyczaj wymagają występowania na przemian dedukcji i indukcji. Obie metody wymagają współgrania logiki i obserwacji. I obie prowadzą do konstruowania teorii społecznych.

Choć zarówno indukcja, jak i dedukcja są uprawnione w naukach społecznych, poszczególne osoby mogą się swobodniej czuć w ramach jednego z tych podejść. Rozważmy taką wymianę zdań z opowieści Sir Arthura Conan Doyle'a *Skandal w Czechach*, gdzie Sherlock Holmes odpowiada dociekliwemu Dr. Watsonowi:

> „«Jakże ci się zdaje, *co by to miało znaczyć?*»
> «Nie mam jeszcze żadnych danych, a błędem jest wielkim tworzyć teorię bez realnych podstaw. Mimo woli zaczyna człowiek podporządkowywać fakty owemu z góry wypowiedzianemu twierdzeniu, podczas gdy przeciwnie – twierdzenie powinno z faktów wypływać»"[9].

Niektórzy badacze społeczni w mniejszym lub większym stopniu zgodziliby się z tym indukcyjnym stanowiskiem, natomiast inni przyjęliby bardziej dedukcyjny punkt widzenia. Jednak większość uznałaby oba podejścia za prawomocne.

Skoro rozumiemy już te dedukcyjne i indukcyjne związki między teorią a badaniami, sięgnijmy teraz nieco głębiej i zapytajmy, jak za pomocą tych dwóch podejść konstruowane są teorie.

Budowa teorii dedukcyjnej

Aby zobaczyć, co wiąże się z konstruowaniem teorii dedukcyjnej i testowaniem hipotez, wyobraźmy sobie, że zamierzasz zbudować taką teorię. Jak byś się do tego zabrał?

[9] Arthur Conan Doyle, *Skandal w Czechach*, tłum. St. Zawirska-Michalska, Nakładem Wacława Pawłowskiego, Warszawa 1922, s. 8 (przyp. tłum.).

Zaczynamy

Pierwszym krokiem w konstruowaniu teorii dedukcyjnej jest wybranie tematu, który cię interesuje. Temat ten może być bardzo szeroki: „Jaka jest struktura społeczeństwa?" albo węższy, jak: „Dlaczego ludzie popierają lub odrzucają ideę prawa kobiet do aborcji?". Bez względu na to, jaki to temat, musi on dotyczyć rzeczy, których wyjaśnienie i zrozumienie cię interesuje.

Gdy wybrałeś temat, następnym krokiem jest zestawienie tego, co już o nim wiadomo. Częściowo oznacza to spisanie swych własnych obserwacji i pomysłów. Oznacza to również sprawdzenie, jak inni uczeni wypowiedzieli się w danej sprawie. Możesz rozmawiać z ludźmi, możesz też zechcieć przeczytać literaturę naukową poświęconą temu zagadnieniu. Załącznik A dostarcza wskazówek, jak korzystać z biblioteki – prawdopodobnie spędzisz w niej dużo czasu.

Twoje wstępne badania ukażą zapewne trwałe prawidłowości, które odkryli już wcześniej uczeni. Na przykład zmienne dotyczące religii i polityki będą wyróżniały się jako ważne determinanty postaw wobec aborcji. Tego typu wyniki przydadzą ci się bardzo, gdy będziesz tworzył własną teorię.

W procesie tym należy doceniać wartość introspekcji. Zawsze, kiedy jesteśmy w stanie przyjrzeć się procesom, które zachodzą w nas samych – włączając w to reakcje, lęki i uprzedzenia – możemy skutecznie wniknąć w naturę ludzkiego działania jako takiego. Nie chcę przez to powiedzieć, że każdy myśli tak jak ty czy ja, ale że introspekcja bywa użytecznym źródłem informacji, które mogą stać się natchnieniem dla naszych badań.

Tworzenie własnej teorii

Teraz, gdy już przejrzałeś wcześniejsze prace dotyczące danego tematu, jesteś gotów, by zacząć konstruować swoją teorię. Choć budowanie teorii nie przypomina maszerowania równym krokiem, to proces ten, ogólnie rzecz biorąc, wymaga podjęcia pewnych kroków przedstawionych poniżej.

1. Określ temat.
2. Określ zakres zjawisk, do których odnosi się twoja teoria. Czy będzie ona mieć zastosowanie do całości życia społecznego człowieka, czy będzie mieć zastosowanie tylko do obywateli USA, tylko do ludzi młodych, czy do czegoś innego?
3. Rozpoznaj i określ twoje główne pojęcia i zmienne.
4. Poszukaj tego, co wiadomo (tezy) o związkach między tymi zmiennymi.
5. Rozumując logicznie, przejdź od tych tez do konkretnego tematu, który badasz.

Omówiliśmy już punkty od 1 do 3, skupmy się więc teraz na 4 i 5. Skoro określiłeś już odpowiednie pojęcia i znalazłeś to, czego się dotąd o nich dowiedziano, możesz zacząć tworzyć strukturę tez, które wyjaśnią badane zagadnienie.

Przyjrzyjmy się teraz, jak te klocki dopasowane są do siebie w budowanej teorii dedukcyjnej i w badaniach empirycznych.

Przykład teorii dedukcyjnej: sprawiedliwość dystrybutywna

Pojęcie sprawiedliwości dystrybutywnej jest zagadnieniem znajdującym się w centrum zainteresowań naukowców – czy ludzie uważają, że życie traktuje ich sprawiedliwie, że dostają to, co „się im należy". Guillermina Jasso opisuje teorię sprawiedliwości dystrybutywnej w sposób bardziej formalny. Czytamy:

> „Teoria ta dostarcza matematycznego opisu procesu, za pomocą którego jednostki, gdy zastanawiają się nad posiadanymi, cenionymi przez siebie dobrami (takimi jak uroda, inteligencja czy bogactwo), porównują się z innymi ludźmi, odczuwając natychmiast, jak fundamentalne znaczenie ma ocena sprawiedliwości (S). W ocenie tej uchwycone jest ich poczucie, że zostali sprawiedliwie bądź niesprawiedliwie potraktowani przy dystrybucji dóbr naturalnych i społecznych".
>
> (1988, s. 11)

Zauważmy, że Jasso przypisała symboliczną reprezentację swej kluczowej zmiennej: S będzie oznaczać sprawiedliwość dystrybutywną. Czyni to, aby utrzymać swój zamiar sformułowania teorii za pomocą wzorów matematycznych. Choć często teorie są wyrażane matematycznie, nie będziemy zgłębiać tu zbytnio tej praktyki.

Jasso wskazuje, że w jej teorii istnieją trzy rodzaje założeń. „Pierwsze artykułuje zasadniczy aksjomat stanowiący faktyczny punkt wyjścia dla teorii". Rozwija to następująco: „Teoria wychodzi od przyjętego aksjomatu porównywania, który formalizuje od dawna utrzymywany pogląd, że rozległa klasa zjawisk, włącznie ze szczęściem, poczuciem własnej wartości i poczuciem sprawiedliwości dystrybutywnej, może być rozumiana jako produkt procesu porównywania" (1988, s. 11).

A zatem twoje poczucie tego, czy otrzymałeś „należną ci" część dobrych rzeczy w życiu, wypływa z porównywania się z innymi. Fakt, że może ci się to wydawać oczywiste, nie stanowi wady aksjomatu. Pamiętaj, że aksjomaty to oczywistości, które są początkiem teorii.

Jasso kontynuuje tworzenie podstaw swej teorii. Po pierwsze, wskazuje, że nasze poczucie sprawiedliwości dystrybutywnej jest funkcją „faktycznego posiadania dóbr (F)" i „porównawczego posiadania (P)" jakiegoś dobra. Rozważmy na przykład pieniądze. Moje poczucie sprawiedliwości pod tym względem jest funkcją tego, ile ich faktycznie mam w porównaniu z tym, ile mają inni. Określając oba składniki porównania, Jasso może użyć ich jako zmiennych w swojej teorii.

Następnie Jasso przedstawia „zasadę pomiaru", która jeszcze dokładniej określa, jak obie te zmienne, F i P, zostaną skonceptualizowane. Krok ten jest potrzebny, ponieważ tylko część dóbr, które mają być badane, to dobra konkretne i mierzone powszechnie (jak pieniądze). Pozostałe są mniej namacalne (jak szacunek). Pierwszy rodzaj, jak mówi Jasso, będzie mierzony konwencjonalnie, natomiast drugi „za pomocą względnego uszeregowania przez jednostkę (...) w ramach specjalnie wybranej grupy porównawczej". Teoria dostarczy wzoru do tworzenia tych pomiarów (1988, s. 13).

Kontynuując w podobny sposób, Jasso wprowadza dodatkowe elementy, wplatając je we wzory matematyczne, których będzie używać przy przewidywaniu funkcjonowania sprawiedliwości dystrybutywnej w różnych układach społecznych. Oto tylko wybrana część wniosków, do których zaprowadziła ją jej teoria (1988, s. 14–15).

- Ludzie będą woleli ukraść coś członkowi swojej grupy niż osobie z zewnątrz, przy założeniu, że wszystkie pozostałe warunki są identyczne.
- Preferencja ta wyraźniej zaznacza się w grupach biednych niż bogatych.
- W przypadku kradzieży informatorzy pojawiają się tylko, gdy dotyka ona osób z innej grupy oraz są oni członkami grupy złodzieja.
- Osoby, które spóźniają się na letni obóz albo na rozpoczęcie pierwszego roku na studiach, z większym prawdopodobieństwem staną się przyjaciółmi osób lubiących gry, w których wygrana zależy od szczęścia, niż tych, którzy wolą gry wymagające umiejętności.
- Wraz ze wzrostem zamożności jakiegoś społeczeństwa staje się ono bardziej podatne na deficyt w wydatkach.
- Społeczeństwa, które z zadowoleniem przyjmują wzrost liczby ludności, to społeczeństwa, w których do grupy cenionych dóbr należy przynajmniej jedno dobro mające charakter ilościowy, na przykład dobrobyt.

Z teorii Jasso wynika wiele innych tez, ale ta próbka powinna pozwolić zorientować się, dokąd może nas zaprowadzić teoretyzowanie dedukcyjne. Aby mieć pojęcie o tym, jaką drogę przeszło rozumowanie Jasso, by dojść do takich tez, przyjrzyjmy się krótko logice dwóch tez dotyczących kradzieży w obrębie własnej grupy i poza nią.

- Ludzie będą woleli ukraść coś członkowi swojej grupy niż osobie z zewnątrz, przy założeniu, że wszystkie pozostałe warunki są identyczne.

Wychodząc z założenia, że złodzieje chcą maksymalizować swą względną zamożność, zapytajmy, czy cel ten lepiej osiągnięty zostanie dzięki okradaniu tych, z którymi się porównujemy, czy tych z zewnątrz. W obu przypadkach kradzież zwiększy ilość faktycznie posiadanych przez nas dóbr, ale co z naszym posiadaniem porównawczym?

Po chwilowym namyśle okaże się, że kradzież w ramach grupy, do której członków się porównujemy, zmniejszy ich stan posiadania, zwiększając dodatkowo naszą względną zamożność. Aby to uprościć, wyobraź sobie, że w grupie, z którą się porównujesz, są tylko dwie osoby: ty i ja. Przypuśćmy, że obaj mamy po 100 USD. Jeśli ukradniesz 50 USD komuś z zewnątrz, zwiększysz swą względną zamożność o 50% w stosunku do mnie: 150 USD *versus* 100 USD. Ale jeśli ukradniesz mi 50 USD, zwiększysz swą względną zamożność o 200%: 150 USD w porównaniu z moimi 50 USD. Łatwiej osiągniesz swój cel, kradnąc w obrębie grupy, do której członków się porównujesz.

- W przypadku kradzieży informatorzy pojawiają się tylko, gdy dotyka ona osób z innej grupy oraz są oni członkami grupy złodzieja.

Czy rozumiesz, dlaczego dla informatorów sens ma (1) ujawnianie się tylko w przypadku kradzieży osobom z innej grupy oraz (2) przynależność do grupy złodzieja? Teza ta ponownie zależy od podstawowego założenia, że każdy chce względnie polepszyć swą pozycję. Przypuśćmy, że ty i ja jesteśmy w tej samej grupie, lecz teraz należą do niej jeszcze inni ludzie. Jeśli ukradniesz coś komuś innemu z naszej grupy, moja względna pozycja w grupie nie zmienia się. Choć twoja zamożność wzrosła, średnia zamożność w grupie pozostała taka sama (ponieważ zamożność kogoś innego zmniejszyła się o tę samą sumę). A więc moja względna pozycja pozostaje taka sama. Nic nie zyskuję, donosząc na ciebie.

Jeśli jednak ukradniesz coś komuś spoza grupy, z *której członkami się porównujemy*, dochód z twej niegodziwości zwiększy całkowitą zamożność w naszej grupie. Teraz moja własna zamożność zestawiona z tą całkowitą miarą zmniejsza się. Ponieważ moja względna zamożność ucierpiała, jestem bardziej skłonny

donieść na ciebie, by położyć kres twym kradzieżom. Informatorzy pojawiają się zatem tylko w przypadku kradzieży dotyczącej osób z innych grup.

Ten ostatni wniosek zaczyna też wyjaśniać, dlaczego informatorzy pochodzą z własnej grupy złodzieja. Widzieliśmy właśnie, w jaki sposób twoja kradzież obniżyła moją pozycję względem grupy. A co z członkami tej drugiej grupy (z wyjątkiem osoby, którą okradłeś)? Każdy z nich w rzeczywistości osiąga zysk z tej kradzieży, ponieważ obniżyłeś całkowitą miarę, z którą się porównują. Nie mają zatem żadnego powodu, aby na ciebie donosić. Teoria sprawiedliwości dystrybutywnej przewiduje więc, że informatorzy pojawią się w grupie, z której członkami porównuje się złodziej.

Ten krótki rzut oka na wywody Jasso powinien dać ci pewne pojęcie o przedsięwzięciu, jakim jest teoria dedukcyjna. Oczywiście żadne z przedstawionych przewidywań nie jest przez tę teorię zagwarantowane. Rolą badań jest przetestowanie ich i określenie, czy to, co brzmi sensownie (logika), rzeczywiście pojawia się w praktyce (obserwacja).

Budowa teorii indukcyjnej

Jak widzieliśmy, dość często badacze rozpoczynają konstruowanie teorii metodą indukcji, najpierw obserwując pewne aspekty życia społecznego. Potem próbują odkryć prawidłowości, które mogą wskazywać na pewne względnie uniwersalne zasady. Barney Glaser i Anselm Strauss (1967) ukuli w odniesieniu do tej metody termin *teoria ugruntowana*.

Badania terenowe – bezpośrednia obserwacja zachodzących zdarzeń – są często stosowane w celu rozwinięcia teorii poprzez obserwację. W swej długiej i bogatej tradycji antropologia z dużym pożytkiem korzystała z tej metody.

Wśród współczesnych badaczy społecznych żaden nie osiągnął takiej biegłości w poszukiwaniu prawidłowości ludzkiego zachowania za pomocą obserwacji, jak Ervin Goffman:

> „Gra taka jak szachy stwarza dla pojętnych graczy nadający się do zamieszkania wszechświat, płaszczyznę istnienia, obsadę postaci z nieskończoną ilością różnych sytuacji i zachowań, poprzez które można realizować swoją naturę i powołanie. A jednak prawie wszystko to daje się zredukować do zestawu kilku wzajemnie zależnych reguł i praktyk. Jeśli pełnia znaczeń codziennego działania w podobny sposób uzależniona jest od zamkniętego, skończonego zestawu reguł, wówczas wyjaśnienie ich dostarczyłoby nam skutecznych środków analizy życia społecznego".
>
> (1974, s. 5)

Podejmując rozmaite wysiłki badawcze, Goffman odkrył reguły tak odmiennych zachowań, jak życie w zakładzie psychiatrycznym (1961) i radzenie sobie osób zeszpeconych ze swoją „zniszczoną tożsamością" (*spoiled identity*) (1963). W każdym przypadku Goffman dogłębnie obserwował dane zjawisko i rozsupływał reguły rządzące zachowaniem. Badania Goffmana dostarczają wspaniałego przykładu jakościowych badań terenowych jako źródła teorii ugruntowanej.

Nasze wcześniejsze omówienie „hipotezy pocieszenia" i zaangażowania w sprawy kościoła pokazuje, że jakościowe badania terenowe nie są jedyną metodą obserwacji właściwą dla rozwinięcia teorii indukcyjnej. Oto inny szczegółowy przykład mający dokładniej przedstawić budowę teorii indukcyjnej z wykorzystaniem metod jakościowych.

Przykład teorii indukcyjnej: dlaczego ludzie palą marihuanę?

W latach 60. i 70. używanie marihuany w campusach amerykańskich uniwersytetów było przedmiotem poważnej dyskusji w prasie. Niektórzy ludzie martwili się popularnością marihuany; inni przyjmowali to z zadowoleniem. Nas interesuje tu, dlaczego jedni studenci palili marihuanę, a inni nie. Sondaż wśród studentów Uniwersytetu Hawajskiego (Takeuchi 1974) dostarczył danych pozwalających odpowiedzieć na to pytanie.

W czasie trwania tych badań proponowano mnóstwo wyjaśnień odnoszących się do używania narkotyków. Ludzie, którzy byli jego przeciwnikami, często na przykład sugerowali, że palacze marihuany to uniwersyteccy wykolejeńcy, którzy próbowali uniknąć rygorów życia akademickiego. Z kolei zwolennicy marihuany często mówili o poszukiwaniu nowych wartości: palacze marihuany, twierdzono, to ludzie, którzy na wylot przejrzeli hipokryzję wartości klasy średniej.

Jednak David Takeuchi (1974), analizując dane zebrane wśród studentów Uniwersytetu Hawajskiego, nie potwierdził żadnego z proponowanych wyjaśnień. Osoby, które oświadczyły, że palą marihuanę, miały zasadniczo takie same wyniki na studiach, jak ci, którzy nie palili, i obie grupy były tak samo zaangażowane w tradycyjne szkolne zajęcia.

Obie grupy wydawały się równie dobrze zaadaptowane do życia w campusie.

Istniały jednak między nimi inne różnice:

1. Kobiety były mniej skłonne do palenia marihuany niż mężczyźni.
2. Studenci z Azji (duży procent wśród wszystkich studentów) byli mniej skłonni do palenia marihuany niż pozostali studenci.
3. Studenci mieszkający w domu mniej byli skłonni do palenia marihuany niż studenci mieszkający w akademikach.

Tak jak w przypadku religijności, trzy zmienne niezależnie wpływały na prawdopodobieństwo palenia przez studenta marihuany. Marihuanę paliło około 10% Azjatek mieszkających w domu, w zestawieniu z 80% mężczyzn nie pochodzących z Azji, którzy mieszkali w akademikach. I tak jak w badaniach dotyczących religijności, badacze odkryli silną prawidłowość w używaniu narkotyków, jeszcze zanim znaleźli dla niej wyjaśnienie.

W tym przypadku wyjaśnianie przybrało szczególny obrót. Zamiast wyjaśnić, dlaczego niektórzy studenci palą marihuanę, badacze wyjaśnili, dlaczego niektórzy jej nie palą. Przyjmując, że wszyscy studenci mieli jakieś pobudki, by spróbować narkotyków, badacze zasugerowali, że studenci różnią się poziomem „społecznych ograniczeń", zabezpieczających ich przed uleganiem tym pobudkom.

Społeczeństwo amerykańskie jako całość, gdy chodzi o zachowania odbiegające od normy, jest bardziej permisywnie nastawione wobec mężczyzn niż wobec kobiet. Rozważmy przykład grupy pijanych i hałasujących mężczyzn. Mamy skłonność do lekceważenia takich zachowań, odwołując się do „kumplostwa" i „dobrej zabawy", podczas gdy grupa kobiet zachowujących się podobnie byłaby zapewne postrzegana z wielką dezaprobatą. Mamy powiedzenie „kawalerska fantazja"[10], ale żadnego porównywalnego w odniesieniu do kobiet. Badacze wnioskowali zatem, że kobiety, paląc marihuanę, miałyby więcej do stracenia niż mężczyźni. Innymi słowy, bycie kobietą stanowiło ograniczenie w paleniu marihuany.

Studenci mieszkający w domu, w przeciwieństwie do studentów żyjących na własną rękę, mieli oczywiste ograniczenia w paleniu marihuany. Poza różnicami w okazjach, studenci mieszkający w domu uznani zostali za bardziej zależnych od rodziców, a zatem łatwiej podlegających dodatkowym karom za złamanie prawa.

Wreszcie azjatycka subkultura na Hawajach tradycyjnie bardziej niż inne kultury nagradzała posłuszeństwo wobec prawa, a więc studenci pochodzenia azjatyckiego mieliby więcej do stracenia, gdyby przyłapano ich na łamaniu tegoż prawa przez palenie marihuany.

Podsumowując, teoria „społecznych ograniczeń" oferowała wyjaśnienie zaobserwowanych różnic w prawdopodobieństwie palenia marihuany. Im więcej student miał ograniczeń, tym mniejsze było prawdopodobieństwo, że będzie palić marihuanę. Warto powtórzyć, że badacze w ogóle nie myśleli o tego rodzaju teorii, gdy zaczynały się ich badania. Była ona wynikiem analizy danych.

Związki między teorią a badaniami

Przez cały ten rozdział przyglądaliśmy się różnym aspektom powiązań między teorią a badaniami w naukach społecznych. W modelu

[10] W oryg. „boys will be boys" (przyp. tłum.).

dedukcyjnym badania są wykorzystywane do testowania teorii. W modelu indukcyjnym teorie rozwijane są na podstawie analizy danych z badań. W tej ostatniej części uważniej przyjrzymy się, w jaki sposób powiązane są teoria i badania w prawdziwych analizach prowadzonych w ramach nauk społecznych.

Przedyskutowaliśmy dwa wyidealizowane, logiczne modele powiązań między teorią a badaniami, lecz w praktyce nauk społecznych stworzono wiele wariacji na te dwa tematy. Czasami teoretyczne zagadnienia są wprowadzane wyłącznie jako tło dla analiz empirycznych. Kiedy indziej cytuje się wybrane dane empiryczne, aby podeprzeć argumenty teoretyczne. W żadnym z tych przypadków nie ma między teorią a badaniami prawdziwej interakcji, która miałaby na celu stworzenie nowych wyjaśnień. Niektórzy badacze nie korzystają wcale z teorii, wyraźnie zaplanowawszy na przykład etnograficzny opis konkretnej społecznej sytuacji, powiedzmy, antropologiczną relację dotyczącą jedzenia i ubioru w konkretnym społeczeństwie.

Gdy jednak czytamy raporty z badań społecznych, bardzo często dowiadujemy się, że autorzy są świadomi implikacji, jakie ich badania będą miały dla teorii społecznej i *vice versa*. Oto kilka przykładów, które zilustrują tę myśl.

Kiedy W. Lawrence Neuman (1998) przystąpił do badania problemu monopoli („problem trustów") w historii USA, dostrzegł jego związki z teoriami wyjaśniającymi, jak ruchy społeczne przekształcają społeczeństwo („transformacja państwa"). Przekonał się jednak, że istniejące teorie nie pasują do zadania, które miał przed sobą.

> „Teoria transformacji państwa łączy ruchy społeczne z procesami kształtowania polityki państwa, skupiając się na roli kulturowego znaczenia zorganizowanych zmagań politycznych. Pomimo podobieństwa pojęć i problemów idee konstrukcjonistyczne znalezione w literaturze dotyczącej problemów społecznych, ruchów społecznych i symboli politycznych nie zostały włączone do tej teorii. W pracy posługuję się wspomnianymi trzema rodzajami literatury, by ulepszyć teorię transformacji państwa".
>
> (1998, s. 315)

Zmodyfikowawszy w ten sposób teorię transformacji państwa, Neuman miał teoretyczne narzędzie, które mogło wprowadzić jego badania i analizę w zagadnienie taktycznych działań politycznych związanych z monopolami, rozpoczętych w latach 80. XIX wieku i trwających do I wojny światowej. Teoria posłużyła zatem jako wsparcie dla badań i jednocześnie została przez nie zmodyfikowana.

W trochę podobnym badaniu Alemseghed Kebede oraz J. David Knottnerus (1998) chcieli prześledzić narodziny rastafarianizmu na Karaibach. Uważali jednak, że ostatnio powstałe teorie ruchów społecznych stały się zbyt pozytywistyczne w swym skupieniu na mobilizacji zasobów. Uznali, że teoria mobilizacji zasobów nie docenia

> „motywacji, spostrzeżeń i zachowań osób uczestniczących w ruchu... i koncentruje się zamiast tego na «dlaczego» i «jak» samej mobilizacji. Na ogół teoretyczne i badawcze problemy obejmują kwestie: w jaki sposób wyłaniająca się struktura organizacyjna ruchu dąży do mobilizacji i rutynizacji przepływu zasobów, oraz jak istniejący aparat polityczny wpływa na organizację zasobów".
>
> (1998, s. 500)

Badacze uznali, że aby studiować rastafarianizm w sposób właściwszy, potrzebne jest włączenie kilku pojęć ze współczesnej psychologii społecznej. W szczególności poszukiwali modeli, które można by wykorzystać przy rozwiązywaniu kwestii znaczenia i przekonań zbiorowych.

Frederika E. Schmitt i Patricia Yancey Martin (1999) szczególnie zainteresowały się tym, co wpływa na skuteczność funkcjonowania Rape Crisis Centres (centra niosące pomoc ofiarom gwałtu) i jak radzą sobie one z otoczeniem organizacyjnym i politycznym, w którym działają. Autorki badania znalazły konstrukty teoretyczne odpowiednie dla ich badania.

> „Niniejsze studium przypadku dyskretnej mobilizacji prowadzonej przez Południowokalifornijskie Centrum Kryzysowe dla Ofiar Gwałtu wykorzystuje dane pochodzące z archiwów, obserwacji i wywiadów w celu zbadania, jak

jedna z organizacji feministycznych pracowała nad zmianami w policji, szkołach, prokuraturze i w niektórych organizacjach stanowych i państwowych w latach 1974–1994. Pojęcie teorii ulicznej (*street theory*) Mansbridge oraz pojęcia dyskretnej mobilizacji (*unobtrusive mobilization*) i polityki dyskursywnej (*discursive politics*) Katzenstein wskazują drogę analizom".

(1999, s. 364)

Podsumowując, nie istnieje prosta recepta na prowadzenie badań w ramach nauk społecznych. Jest to zajęcie, które może się rozwijać na wiele różnych sposobów, o wiele więcej, niż to sugeruje tradycyjne spojrzenie na naukę. Ostatecznie, nauka zależy od dwóch typów działalności: logicznego rozumowania i obserwacji. Jak zobaczymy w tej książce, mogą być one wzajemnie dopasowane według różnych wzorców.

Główne zagadnienia

- Badacze społeczni, aby usystematyzować swój sposób rozumienia i badania życia społecznego, korzystają z rozmaitych paradygmatów.
- Rozróżnienie typów teorii, odnoszące się do różnych paradygmatów, przeciwstawia makroteorie (teorie dotyczące złożonych cech społeczeństwa) mikroteoriom (teoriom dotyczącym mniejszych jednostek czy cech społeczeństwa).
- Paradygmat pozytywistyczny zakłada, że możemy w sposób naukowy odkryć reguły rządzące życiem społecznym.
- Paradygmat darwinizmu społecznego widział w życiu społecznym postępującą ewolucję.
- Paradygmat konfliktu koncentruje się na dążeniu jednostek lub grup do zdominowania innych jednostek lub grup i do uniknięcia ich dominacji.
- Paradygmat symbolicznego interakcjonizmu bada, jak uznawane znaczenia i wzorce społeczne są rozwijane w toku społecznej interakcji.
- Etnometodologia koncentruje się na sposobach, jakie stosują ludzie, by zrozumieć życie społeczne w procesie jego przeżywania, tak jak gdyby każdy był badaczem prowadzącym badania.
- Paradygmat funkcjonalizmu strukturalnego (albo systemów społecznych) usiłuje odkryć funkcje, jakie liczne elementy społeczeństwa pełnią w całym systemie.
- Paradygmaty feministyczne, poza tym, że kierują uwagę na ucisk kobiet w większości społeczeństw, uwidaczniają, że dotychczasowe wyobrażenia rzeczywistości społecznej często wynikały z doświadczeń mężczyzn i dodatkowo je wzmacniały.
- Kilku współczesnych teoretyków i badaczy podważyło długo utrzymującą się wiarę w obiektywną rzeczywistość, która podporządkowana jest racjonalnym regułom. Wskazują oni, że możliwa jest zgoda na rzeczywistość „intersubiektywną".
- Do składników teorii społecznej należą obserwacje, fakty i prawa (które odnoszą się do obserwowanej rzeczywistości) oraz pojęcia, zmienne, aksjomaty czy założenia, tezy i hipotezy (które są logicznymi klockami samej teorii).
- W tradycyjnym wyobrażeniu nauki badacze przechodzą od teorii przez operacjonalizację do obserwacji. Wyobrażenie to nie jest jednak poprawnym obrazem tego, jak w rzeczywistości prowadzi się badania naukowe.
- Teoria i badania w naukach społecznych są połączone za pomocą dwóch metod logicznych: dedukcji (wyprowadzanie z teorii przewidywań i hipotez) oraz indukcji (formułowanie uogólnień na podstawie konkretnych obserwacji).
- W praktyce nauka jest procesem obejmującym następowanie po sobie dedukcji i indukcji.
- Teoria sprawiedliwości dystrybutywnej Guillerminy Jasso obrazuje, jak formalne rozumowanie może prowadzić do rozmaitych teoretycznych przewidywań, które można przetestować za pomocą obserwacji.
- Badania Davida Takeuchiego dotyczące czynników wpływających na palenie marihuany przez studentów Uniwersytetu Hawajskiego pokazują, jak zbieranie obserwacji może prowadzić do uogólnień i teorii wyjaśniającej.

- W praktyce istnieje wiele możliwych powiązań teorii z badaniami oraz wiele sposobów zajmowania się badaniami społecznymi.

Kluczowe pojęcia

Oto terminy, które w tym rozdziale zostały zdefiniowane w odpowiednich kontekstach, i które można też znaleźć w słowniczku na końcu książki:

paradygmaty
makroteoria
mikroteoria
hipoteza
operacjonalizacja
definicja operacyjna
hipoteza zerowa

Pytania i ćwiczenia

1. Rozważ możliwy związek między wykształceniem a uprzedzeniami, o którym wspomniano w rozdziale 1. Opisz, jak mógłbyś zbadać ten związek za pomocą metody (1) dedukcyjnej i (2) indukcyjnej.
2. Przypomnij sobie omówione w tym rozdziale powiązania między teorią a badaniami. Wybierz z jakiegoś czasopisma naukowego artykuł zawierający omówienie badań i określ związek między teorią a badaniami, który tam znajdziesz.
3. Korzystając z jednej z wielu wyszukiwarek (takich jak Excite, HotBot, Infoseek, Lycos, Netscape, WebCrawler lub Yahoo), znajdź w sieci informacje dotyczące przynajmniej trzech z poniżej wymienionych paradygmatów. Podaj adresy internetowe i napisz o teoretykach, których dotyczą znalezione przez ciebie omówienia.

Teoria konfliktu	Funkcjonalizm
Teoria wymiany	Interakcjonizm
Etnometodologia	Pozytywizm
Feminizm	Postmodernizm

Dodatkowe lektury

Chafetz Janet, *A Primer on the Construction and Testing of Theories in Sociology*, Peacock, Itasca, Ill. 1978. Jedna z nielicznych książek poświęconych budowie teorii, napisana wyraźnie dla studentów. Chafetz dzięki prostemu językowi i przykładom z życia codziennego umożliwia zrozumienie filozofii nauki w stopniu podstawowym. Omawia istotę wyjaśniania, rolę założeń i pojęć, opisuje budowanie i testowanie teorii.

Denzin Norman K., Lincoln Yvonna S., *Handbook of Qualitative Research*, Sage, Newbury Park, Cal. 1994. Różni autorzy omawiają przebieg badań jakościowych z perspektywy różnych paradygmatów, pokazując, jak wpływają one na charakter badań. Redaktorzy zbioru krytykują również pozytywizm z perspektywy postmodernizmu.

DeVault Marjorie L., *Liberating Method: Feminism and Social Research*, Temple University Press, Philadelphia 1999. Książka ta dokładnie analizuje niektóre metody związane z paradygmatem feministycznym i opowiada się zarówno za rygorystycznymi badaniami, jak i za wykorzystaniem badań społecznych do walki z uciskiem.

Kuhn Thomas, *Struktura rewolucji naukowych*, tłum. H. Ostromęcka, J. Nowotniak, Warszawa 2001. Ekscytujące i innowacyjne przeformułowanie istoty rozwoju naukowego. Kuhn przeciwstawia się pojęciu stopniowej zmiany i modyfikacji w nauce, argumentując w zamian, że ustanowione paradygmaty mają skłonność opierać się aż do chwili, gdy siła zaprzeczających im dowodów sprawi, że zostaną odrzucone i zastąpione przez nowe paradygmaty. Ta krótka książka jednocześnie stymuluje i dostarcza informacji.

Lofland John, Lofland Lyn H., *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*, Wadsworth, Belmont, Cal. 1995. Wspaniały tekst dotyczący tego, jak prowadzić jakościowe badania społeczne, mając na oku odkrycie zasad życia społecznego. Zawiera krytykę postmodernizmu.

McGrane Bernard, *The Un-TV and 10 mph Car: Experiments in Personal Freedom and Everyday Life*, The Small Press, Fort Bragg, Cal. 1994.

Kilka wspaniałych i pomysłowych przykładów etnometodologicznego podejścia do społeczeństwa i do rzemiosła socjologa. Książka ta jest pożyteczna zarówno dla studentów, jak i dla profesjonalistów.

Reinharz Shulamit, *Feminist Methods in Social Research*, Oxford University Press, New York 1992. Książka ta z perspektywy feministycznej ujmuje kilka technik badań społecznych (takich jak wywiady, eksperymenty i analiza treści).

Ritzer George, *Sociological Theory*, Knopf, New York 1988. Jest to wspaniały przegląd głównych tradycji teoretycznych w socjologii.

Theory Building in Sociology: Assessing Theoretical Cumulation, Jonathan H. Turner (red.), Sage, Newbury Park, Cal. 1989. Zbiór esejów poświęconych budowie teorii socjologicznej, który szczególnie skupia się na pytaniu postawionym przez Turnera w rozdziale wstępnym: „Czy socjologia może być nauką kumulatywną?".

Turner Stephen Park, Turner Jonathan H., *The impossible Science: An Institutional Analysis of American Sociology*, Sage, Newbury Park, Cal. 1990. Dwaj autorzy proponują dwa bardzo odmienne spojrzenia na historię wysiłków podejmowanych przez socjologów amerykańskich w celu stworzenia nauki o społeczeństwie.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział
3

Pojęcie przyczynowości w badaniach społecznych

Przegląd holograficzny

Wyjaśnianie w naukach społecznych opiera się na idei przyczynowości. Zarówno gdy prowadzimy własne badania, jak i gdy oceniamy wnioski z badań przeprowadzonych przez innych, potrzebujemy dokładnego zrozumienia, czym jest związek przyczynowy i jak ustala się go w badaniach społecznych.

Wprowadzenie

Determinizm a nauki społeczne
Przyczynowość w naukach przyrodniczych
Znajdowanie przyczyn w naukach społecznych
Nie ma powodu bez powodu
Determinizm we właściwych proporcjach

Przyczynowość w idiograficznych i nomotetycznych modelach wyjaśniania

Kryteria przyczynowości

Przyczyny konieczne i wystarczające

Błędy w rozumowaniu na temat związków przyczynowych
Prowincjonalizm
Pochopne wnioski
Wątpliwa przyczyna
Przemilczane dowody
Fałszywy dylemat

Połączenie pomiaru z ustalaniem związków
Praktyczne problemy przy pomiarze zmiennych i ustalaniu związków między nimi
Przykład pomiaru i związku między zmiennymi

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Jednym z głównych celów, jakie stawiają sobie badacze – nie tylko przedstawiciele nauk społecznych – jest wyjaśnienie, dlaczego świat wygląda tak, jak wygląda. Na ogół polega to na określaniu przyczyn obserwowanych zjawisk: jakie zjawiska są przyczyną innych zjawisk.

Pojęcie przyczynowości zawierało się *implicite* w większości naszych dotychczasowych rozważań. Teraz musimy dogłębniej je zbadać, szczególnie od strony zastosowań w badaniach społecznych. Zaczniemy od spojrzenia na rolę determinizmu w naukach społecznych przeciwstawionego pojęciu wolnej woli. Potem przyjrzymy się przyczynowości w kontekście idiograficznych i nomotetycznych sposobów wyjaśniania przedstawionych w rozdziale 1. Po rozpatrzeniu kryteriów rozstrzygania, czy między dwoma zjawiskami zachodzi zależność przyczynowa, prześledzimy kilka błędów, które są powszechnie popełniane przez ludzi przy wnioskowaniu o związkach przyczynowych. Zakończymy nasze omówienie, rozważając krótko związek między pomiarem zmiennych a określaniem, jak są one mierzone.

Determinizm a nauki społeczne

Wyjaśnianie w nauce opiera się na myśli, że stany i zdarzenia mają swe przyczyny. Gdy już te przyczyny zrozumiemy, pojmiemy też, dlaczego zjawiska zachodzą tak, a nie inaczej: zważywszy na występowanie ich przyczyn, nie mogą zachodzić w żaden inny sposób. W kategoriach filozoficznych perspektywę tę nazywa się *determinizmem*: występowanie zdarzeń jest determinowane, czy inaczej: powodowane; zdarzeniom „nie wolno" przydarzać się w żaden inny sposób.

W naukach społecznych przeciwstawia się tę deterministyczną perspektywę pojęciu wolnej woli, którą w życiu codziennym traktujemy jako oczywistość. Studenci zajmujący się badaniami społecznymi mogą więc uznać podstawowe założenie nauki za kłopotliwe. Zasadniczy problem jest następujący: czy nasze zachowanie jest wytworem naszych swobodnie podjętych decyzji i zachcianek, czy też jest wytworem sił i czynników działających w świecie, których nie możemy kontrolować i z których być może nawet nie zdajemy sobie sprawy?

Nasze wyjaśnienia w tej kwestii rozpoczniemy od przyjrzenia się przyczynowości w naukach przyrodniczych. Następnie rozważymy, jak to podstawowe pojęcie jest wykorzystywane w naukach społecznych.

Przyczynowość w naukach przyrodniczych

Przykłady deterministycznego modelu wyjaśniania znajdujemy w naukach przyrodniczych na każdym kroku. Przedmiot upuszczony z wieży spada na ziemię z prędkością, którą możemy dokładnie określić, jeśli znamy wszystkie wpływające na nią czynniki. W nauce o istotach żywych wzrost również jest spowodowany przez kilka czynników. Możemy wpływać na wzrost roślin, różnicując otrzymywaną przez nie ilość światła, wody i substancji odżywczych; czynniki te, obok innych, takich jak dziedzictwo genetyczne, określają, jak rośliny będą rosły. Wiemy również, że na wskaźnik wzrostu ludzi wpływają otrzymywane przez nich substancje odżywcze. Pragnienie, by rosnąć albo by nie rosnąć, jest w tym samym stopniu pozbawione znaczenia u ludzi, co u roślin. Przyznajemy, że pod tym względem genetyka i odżywianie w ogromnym stopniu przysłaniają naszą wolną wolę.

Przyczynowo-skutkowy, deterministyczny model nauk przyrodniczych jest więc często stosowany w odniesieniu do istot ludzkich, podobnie jak do roślin i przedmiotów nieożywionych. W tych przypadkach w przeważającej mierze akceptujemy ten deterministyczny model jako właściwy. Uznajemy, że nasza wolna wola jest w pewien sposób skrępowana i ograniczona.

Znajdowanie przyczyn w naukach społecznych

Zasadniczo ten sam model jest stosowany w naukach społecznych. Zazwyczaj jest tak ukryty, że łatwo możemy zapomnieć, jaki jest

jego charakter. Rozważmy na przykładzie, w jaki sposób nauki społeczne opierają się na modelu deterministycznym.

Wyobraź sobie, że w Narodowej Fundacji Nauki udało ci się uzyskać grant w wysokości miliona dolarów na zbadanie przyczyn uprzedzeń. Masz nadzieję, że twoje odkrycia przyczynią się do powstania idei pozwalających ograniczyć uprzedzenia. Jest to z pewnością cel godny pochwały i często wiele rządów oraz prywatnych fundacji wykazuje wolę wspierania takich badań. Przypuśćmy, że otrzymałeś pieniądze, ukończyłeś badania i wysyłasz swój raport do Fundacji. Oto jak on brzmi:

> „Po gruntownych badaniach nad tym zagadnieniem odkryliśmy, że niektórzy ludzie są uprzedzeni, a powodem tego jest fakt, że wolą być uprzedzeni. Inni znów ludzie nie są uprzedzeni, a powodem tego jest fakt, że wolą być nieuprzedzeni".

Dlaczego fundacja odrzuciłaby tę konkluzję jako niezadowalającą? Celem projektu było znalezienie przyczyn uprzedzeń, tymczasem twoje wnioski nie tłumaczą, *dlaczego* pojawiają się uprzedzenia. Jedynie wpychają z powrotem wyjaśnienie do tajemniczej, czarnej skrzynki z napisem „wybór". Nie wiemy więcej niż przed przeprowadzeniem badań.

Gdy szukamy przyczyn uprzedzeń, szukamy powodów czy przyczyn, które sprawiają, że jedni ludzie są uprzedzeni, a inni nie są. Jeśli „wyjaśnieniem" ma być wybór własnej postawy, wówczas chcemy wiedzieć, dlaczego ludzie dokonują różnych wyborów. Do zadowalających powodów mogłyby się zaliczać: konkurencja ekonomiczna, ideologia religijna, poglądy polityczne, doświadczenia z dzieciństwa oraz długość trwania edukacji i jej rodzaj. Badania wykazują na przykład, że wykształcenie sprzyja ograniczeniu uprzedzeń. To jest właśnie rodzaj użytecznego wyjaśnienia przyczynowego, który akceptujemy jako końcowy produkt badań społecznych.

Przyjrzyjmy się nieco bliżej logice tego rodzaju wyjaśniania. Co o ludziach, którzy byli przedmiotem badań, mówi nam wyjaśnianie uprzedzeń w kategoriach wykształcenia i innych czynników? Przede wszystkim mówi, że stali się uprzedzeni lub nieuprzedzeni w efekcie działania czynników, których nie kontrolowali i nie wybierali. Tak jakby doszli do rozstaju dróg – jedna ścieżka prowadzi do uprzedzeń, a druga do ich braku – i zostali popchnięci na jedną z nich przez takie siły, jak doświadczenia z dzieciństwa, poziom wykształcenia, konkurencja ekonomiczna i tym podobne czynniki, nad którymi nie panowali, i których nawet nie uznawali za przyczyny swych postaw. Zatem w stopniu, w jakim czynniki te odpowiadają za uprzedzenia, ludzie ci stali się uprzedzeni lub nieuprzedzeni z powodów niezależnych od swej woli.

Gdy naukowcy badają przestępczość młodocianych, podstawowy model jest taki sam: przestępczość jest powodowana przez czynniki inne niż wolny wybór przestępcy. Następnie przyjmuje się, że czynniki te mogą zostać odkryte i być może tak przekształcone, aby w przyszłości dawały w efekcie mniej przestępców.

Ten sam model badacze stosują zajmując się zachowaniami uznawanymi przez nich lub przez innych za społecznie pożądane. Jakie czynniki sprawiają, że ktoś jest altruistą skupionym na innych, osobą odpowiedzialną? Gdybyśmy znali odpowiedź – tak w domyśle przebiega to rozumowanie – moglibyśmy sprawić, że pojawi się więcej takich osób.

Nie ma powodu bez powodu

Niekiedy ludzie sprzeciwiają się tej linii rozumowania, twierdząc, że jednostki jednak same wybierają to, co determinuje, w jakim stopniu są uprzedzone, w jakim stopniu są altruistami lub przestępcami. Powiedzmy na przykład, że raczej nie masz uprzedzeń. Badacze zaś wywnioskowali, że ich brak jest w twoim przypadku funkcją całego odebranego przez ciebie wykształcenia. Czyż nie zdecydowałeś się na pójście do szkoły? Czy to zatem nie twoje własne wybory są zasadniczą przyczyną tego, że nie jesteś uprzedzony?

Nie ma powodu bez powodu – oto problem, który się z tym przekonaniem wiąże. Dlaczego poszedłeś do szkoły? Początkowo, prawdopodobnie dlatego, że ktoś cię do tego

zmusił. Lecz w pewnym momencie sam podjąłeś decyzję, by kontynuować swą edukację przynajmniej w szkole wyższej. Powiedzmy, że chciałeś poznać otaczający cię świat i pomyślałeś, że szkoła wyższa jest dobrym na to sposobem. Brzmi sensownie.

Brzmi to na tyle sensownie, że właściwie moglibyśmy nawet powiedzieć, iż to twoje pragnienie poznawania świata było przyczyną kontynuowania przez ciebie nauki. Zważywszy, że miałeś takie pragnienie, jak miałbyś nie pójść na studia?

Mógłbyś rzec: „Byłbym nie poszedł, gdybym nie miał wystarczającej ilości pieniędzy". To prawda – ten czynnik mógł cię zmusić do trzymania się z dala od szkoły. Lecz wobec tego przypuśćmy, że twoje pragnienie poznawania świata było tak silne, że poradziłeś sobie z brakiem pieniędzy. Może dostałeś stypendium albo na pewien czas poszedłeś do pracy. Wówczas powracamy do „twego silnego pragnienia" jako przyczyny pójścia na studia.

Ach, ale dlaczego tak bardzo pragnąłeś kontynuować naukę i w ten sposób poznawać otaczający cię świat? Przecież zamiast tego mogłeś uciułać trochę i zaoszczędzić na wycieczkę dookoła świata. Może wyrastałeś w rodzinie, w której każdy, pokolenie za pokoleniem, szedł na studia i gdybyś nie zrobił tego samego, poczułbyś, że zawiodłeś swych bliskich. Albo też pochodzisz z rodziny, w której nikt nigdy nie studiował i wszyscy byli dumni z faktu, że ty możesz być tym pierwszym. W obu przypadkach istniały czynniki, które wywołały twoje mocne pragnienie pójścia na studia, a pragnienie to spowodowało, że na nie poszedłeś. Jeśli przyjmiemy, że pójście na studia sprawiło, iż jesteś wolny od uprzedzeń, to ten pomyślny skutek jest ostatecznie wytworem czynników, których nie wybierałeś.

Oczywiście przykład ten przesadnie upraszcza złożone powody, dla których ludzie decydują się na coś takiego, jak pójście na studia. Możesz jednak zastosować ten tok rozumowania w odniesieniu do samego siebie i zobaczysz, że twoje zachowanie miało przyczyny wyjaśniające, dlaczego robiłeś to, co robiłeś. Ponadto niezależnie od tego, jaki powód miałeś na każdym poszczególnym etapie tego procesu, powód ten również był czymś powodowany. Ostateczny wniosek z tych rozważań jest taki, że nasze postawy i zachowania można wyprowadzić z długiego i skomplikowanego łańcucha powodów tłumaczących, dlaczego staliśmy się tacy, jacy jesteśmy. Gdyby było inaczej, nie moglibyśmy formułować tego typu wyjaśnień, jakie formułujemy w naukach społecznych, ani zobaczyć, jak potwierdzają się one w badaniach.

Zawsze gdy rozpoczynamy badania wyjaśniające w ramach nauk społecznych, *implicite* przyjmujemy model ludzkiego zachowania, który zakłada, że ludzie mają mało indywidualnej wolności wyboru. Czy tym samym opowiadamy się za poglądem, że nie istnieje nic takiego jak wolny wybór? Zamiast odpowiadać na to pytanie prostym „tak" lub „nie", przedstawmy założenia determinizmu, zachowując odpowiednie proporcje.

Determinizm we właściwych proporcjach

Zagadnienie determinizmu i wolności jest dość złożone – należy do tych, nad którymi filozofowie debatują od tysięcy lat i prawdopodobnie będą debatować przez kolejne tysiące. Jest to chyba jedno z tych „otwartych pytań", które mają większą wartość, gdy się je stawia, niż gdy się na nie odpowiada. I oczywiście nie zamierzamy tutaj rozstrzygać tej kwestii.

Niemniej jednak jako badacze społeczni musimy uznać rolę determinizmu w badaniach społecznych i zmierzyć się z jego konsekwencjami. Młodzi badacze często żywią obawę, czy nie uczą się po to, by wykazać, że oni sami nie mają wolnej woli czy osobistej wolności decydowania o przebiegu własnego życia. W stopniu, w jakim obawa ta narasta i umacnia się, przeszkadza ona w opanowywaniu umiejętności i technik analitycznych. Jest zatem ważne, byśmy stanęli oko w oko z tym problemem.

Choć wyjaśnienia w naukach społecznych opierają się na modelu deterministycznym, musimy dobrze zrozumieć, co *nie* jest częścią tego modelu. Po pierwsze, badacze społeczni nie muszą wierzyć, że wszystkie ludzkie działania, myśli i uczucia są determinowane, ani też nie

żyją tak, jak gdyby w to wierzyli. Jednak gdy usiłują wyjaśnić zjawiska nadające się na przedmiot naukowych badań społecznych, koniecznie muszą używać pojęć przyczyny i skutku. Po drugie, jak już sugerowałem, model deterministyczny nie zakłada, że wzory związków przyczynowych są proste. Nie zakłada również, że wszyscy ulegamy wpływowi tych samych sił i czynników: powody, dla których poszedłeś na studia (i powody tych powodów), z pewnością różnią się w pewnym stopniu od moich. Po trzecie, deterministyczny model leżący u podstaw wyjaśniania w naukach społecznych nie sugeruje, że znamy wszystkie odpowiedzi dotyczące przyczyn wszystkiego, ani że kiedykolwiek je poznamy. Właściwie wiele pożytecznych badań zmierza do odkrycia związków czy relacji między zmiennymi, nie usiłując wykazać związku przyczynowego.

I wreszcie, jak zauważyliśmy wcześniej, nauki społeczne z reguły opierają się na modelu przyczynowym, który ma charakter probabilistyczny. Zamiast więc przewidywać, na przykład, że konkretny absolwent szkoły wyższej nie będzie miał uprzedzeń, przewiduje się w badaniach, że uprzedzenia z większym lub mniejszym prawdopodobieństwem zależą od poziomu wykształcenia ludzi. Nie oznacza to, że wszyscy mający doktorat są wolni od uprzedzeń. Nie oznacza to też, że wszyscy, którzy wcześnie porzucili szkołę, są uprzedzeni. Oznacza to natomiast, że w naszym społeczeństwie uprzedzenia z mniejszym prawdopodobieństwem zaobserwujemy wśród osób dobrze wykształconych, a z większym wśród tych, którzy uczyli się krócej.

Podsumowując, sposób rozumienia zjawisk, do którego dążą badacze społeczni, konstruując wyjaśnienia przyczynowe, nieuchronnie jest związany z deterministycznym modelem ludzkiego zachowania. Model ten milcząco zakłada, że badane cechy i działania są determinowane przez siły oraz przez czynniki, które można rozpoznać w toku badania. Badacze społeczni nie muszą wierzyć, że każdy aspekt ludzkiego życia jest w całości determinowany, ale na pewno, szukając wyjaśnień dla interesujących ich zjawisk, muszą chcieć stosować deterministyczną logikę.

Przyczynowość w idiograficznych i nomotetycznych modelach wyjaśniania

Dotychczasowe rozważania, w których nacisk był położony na wielość przyczyn mogących prowadzić do danego zachowania, stanowią przykład idiograficznego modelu wyjaśniania. Celem tego modelu jest wyjaśnianie za pomocą zestawienia rozlicznych powodów, które stoją za konkretnym działaniem lub wydarzeniem. Choć w praktyce wykaz ten nigdy nie jest naprawdę wyczerpujący, model idiograficzny stosowany jest często w wielu różnych kontekstach.

Na przykład historycy tradycjonaliści chętnie korzystają z idiograficznego modelu, wyliczając każdą z poszczególnych przyczyn rewolucji francuskiej lub trwających, najwyraźniej bez końca, konfliktów na Bałkanach. Psychologowie kliniczni mogą stosować ten model, szukając wyjaśnień aberracyjnych zachowań pacjenta. Sąd karny, odpowiadając obronie powołującej się na okoliczności łagodzące, może chcieć zbadać każdy z różnych czynników, które odegrały rolę w działaniach oskarżonego.

Jak wskazałem w rozdziale 1, modelu tego używają czasem również badacze społeczni. Zastanówmy się nad wydarzeniami z 13 maja 1985 roku w Filadelfii. Miejska policja próbowała tam wręczyć nakazy aresztowania aktywistom radykalnej grupy znanej jako MOVE i walczącej o prawa obywatelskie. Po ewakuacji mieszkańców z sąsiedztwa i otoczeniu głównej siedziby organizacji przez 500 uzbrojonych oficerów przypuszczono atak. Napaść policji, która zaczęła się od użycia broni maszynowej, armatek wodnych i gazu łzawiącego, osiągnęła kulminację, gdy helikopter zrzucił materiały wybuchowe na dach, zabijając 11 osób i burząc dwa bloki mieszkalne.

Wstrząśniętym głosem, którego echa słychać było jeszcze w latach 90., w Waco w stanie Teksas i w Ruby Ridge w stanie Idaho podczas podobnych oblężeń wymuszających podporządkowanie się prawu, ludzie zadawali sobie pytanie, jak taka tragedia mogła się wydarzyć. Robin Wagner-Pacifici (1995), aby odkryć przyczyny, podjął się badań. Zastosował technikę znaną jako *analiza dyskursu*, która wymaga pod-

dania drobiazgowej analizie podstawowych znaczeń znajdowanych w różnych formach komunikacji. W tym przypadku oświadczenia policji można najłatwiej zrozumieć jako „dyskurs wojenny". Walka, która z niego wynikła, z policyjnego punktu widzenia nabiera sensu.

O ile idiograficzny model wyjaśniania jest często stosowany w życiu codziennym i w badaniach społecznych, o tyle inne sytuacje oraz inne cele domagają się modelu nomotetycznego. Nie pociąga on za sobą wyliczania wszystkich okoliczności, których rezultatem było konkretne działanie czy wydarzenie. Raczej ma on na celu odkrycie tych okoliczności, które są najważniejsze przy wyjaśnianiu ogólnych typów działań czy zdarzeń.

Przypuśćmy, że chcieliśmy dowiedzieć się, dlaczego ludzie głosowali tak, jak głosowali w wyborach prezydenckich w 1996 roku. Każda osoba, z którą rozmawialiśmy, mogła podać wiele powodów, dla których głosowała albo na demokratę Billa Clintona, albo na republikanina Boba Dole'a. Wyobraźmy sobie, że ktoś podał nam 99 różnych powodów, dla których głosował na Clintona. Myślelibyśmy prawdopodobnie, że mamy dość pełne wyjaśnienie głosu oddanego przez tę osobę. Właściwie, znajdując kogoś innego, kto zgodziłby się z tymi 99 powodami, czulibyśmy się całkiem pewni przewidując, że również on zagłosował na Clintona. To podejście odpowiada idiograficznemu modelowi wyjaśniania.

Model nomotetyczny z kolei wymaga wyodrębnienia relatywnie niewielu czynników, które dostarczą cząstkowego wyjaśnienia zachowań wyborczych wielu lub wszystkich ludzi. Na przykład orientacja polityczna – liberalna lub konserwatywna – byłaby prawdopodobnie czynnikiem mającym główne znaczenie przy determinowaniu wyborczych zachowań elektoratu jako całości. Większość z tych, którzy należą do grupy określanej cechą „liberał", prawdopodobnie zagłosowała w 1996 roku na Clintona, a większość z tych, którzy należą do grupy określanej cechą „konserwatysta", prawdopodobnie zagłosowała na Dole'a. W takim zakresie orientacja polityczna jest jedną z przyczyn zachowań wyborczych. Jednak to jedno stwierdzenie nie dałoby pełnego wyjaśnienia zachowań wyborczych wszystkich jednostek. Niektórzy liberałowie zagłosowali na Dole'a, niektórzy konserwatyści na Clintona. Model nomotetyczny ma na celu dostarczenie jak najdokładniejszego wyjaśnienia przy jak najmniejszej liczbie zmiennych przyczynowych, tak by odkryć ogólne wzorce przyczynowo-skutkowe.

Model nomotetyczny w swym podejściu do związków przyczynowych jest nieuchronnie probabilistyczny. Określenie kilku zmiennych przyczynowych nieczęsto, jeśli w ogóle, może dostarczyć pełnego wyjaśnienia złożonych zachowań. W najlepszym z możliwych światów model nomotetyczny wskazuje na bardzo wysokie (lub bardzo niskie) prawdopodobieństwo, że dane działanie wystąpi, gdy tylko pojawi się ograniczona liczba określonych czynników. Dodawanie do równania dużej ilości poszczególnych czynników z reguły zwiększa stopień wyjaśnienia, ale jednocześnie czyni je bardziej złożonym – może nawet tak bardzo, że przestaje być ono użyteczne.

Kiedy zatem Eric Plutzer i John Zipp (1996) zamierzali wyjaśnić liczbę głosów otrzymanych w wyborach 1992 roku przez feministki, poszukiwali zmiennych, które mogły odgrywać rolę nie w konkretnych przypadkach, lecz na poziomie ogólnym. Odkryli na przykład, jak można się było spodziewać, że na feministki chętniej głosowały kobiety niż mężczyźni. Interesowało ich jednak również to, w jakim stopniu głosowanie wynikające z płci kolidowało z lojalnością wobec partii. Chodzi tu o sytuacje, w których niektórzy mężczyźni demokraci opuszczali feministyczne kandydatki swej partii, a niektóre kobiety republikanki wspierały feministki z partii demokratycznej. Cel nomotetyczny polegał więc na odkryciu zmiennych, które miały największy wpływ na decyzje wyborcze.

Inny przykład wyjaśniania nomotetycznego dał Jeremy Hein (1993), przeglądając zestaw badań, których celem było odróżnienie doświadczeń uchodźców od doświadczeń innych imigrantów. Zwróćmy uwagę na język związku przyczynowego.

> „W przypadku uchodźców i imigrantów te same zmienne demograficzne pozwalają przewidywać rodzaj zatrudnienia i dochody. Obie grupy przystosowują się w ramach gospodarstw domo-

wych i uzyskują dochód z wielu źródeł. Kobiety pełnią zasadniczą funkcję, budując pomosty między sieciami społecznymi oraz sferą ekonomiczną. Jednak interwencja państwa raz jeszcze stwarza pewne ważne różnice, szczególnie w dostępie do systemu pomocy społecznej. (Uchodźcy uzyskują uprawnienia do tej pomocy natychmiast, podczas gdy inni imigranci muszą czekać pięć lat)".

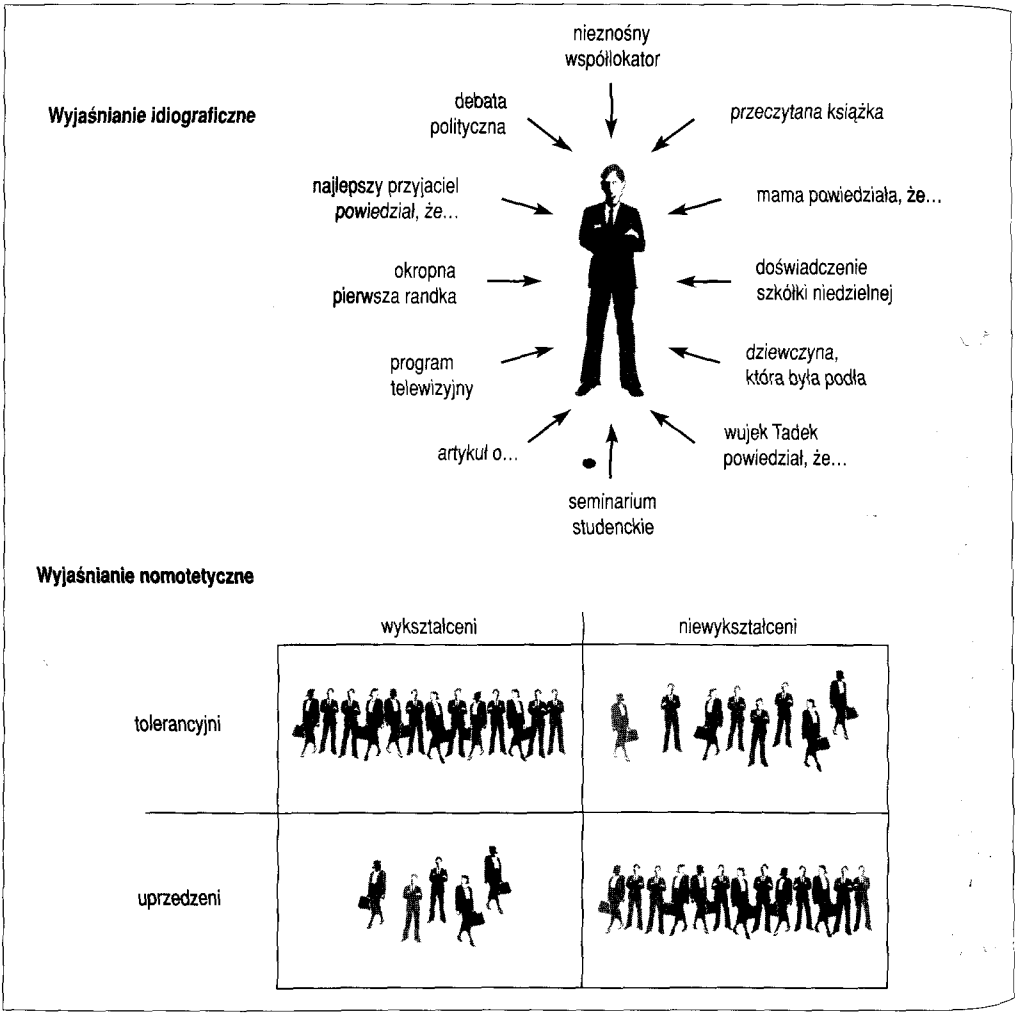
(1993, s. 55).

Słowa takie jak *przewidywać* i *uprzedzenia* sygnalizują wyjaśnienia przyczynowe. Zauważmy jednak, że uogólnienia w tych twierdzeniach pozostawiają mnóstwo miejsca na indywidualne interpretacje. Badacze opisują użyteczne, ogólne wzorce związków przyczynowo-skutkowych, które dotyczą szerokiego zakresu przypadków.

Rysunek 3.1 oferuje graficzne przedstawienie różnych podejść reprezentowanych przez wyjaśnianie nomotetyczne i idiograficzne. Diagram ten obrazuje dwa tryby wyjaśniania. W wyjaśnianiu idiograficznym próbujemy zrozumieć, dlaczego ta właśnie osoba jest uprzedzona, uwzględniamy przy tym fakt, że wiele specyficznych doświadczeń i okoliczności miało udział

Rysunek 3.1

Wyjaśnianie idiograficzne *versus* nomotetyczne: dwa przykłady uprzedzeń



w kształtowaniu jej poglądów. W wyjaśnianiu nomotetycznym chcemy na poziomie ogólnym odkryć czynniki, które wpływają na intensywność uprzedzeń. W tym przypadku widzimy, że lepiej wykształceni ludzie są na ogół mniej uprzedzeni niż osoby gorzej wykształcone.

Krytykuje się czasem badaczy społecznych za to, że dehumanizują ludzi, których badają. Oskarżenie to wnoszone jest szczególnie przeciwko modelowi nomotetycznemu; ciężar oskarżenia wzrasta, gdy badacze społeczni analizują kwestie mające dla ludzi duże znaczenie. Osoby religijne na przykład często czują się ograbione z ludzkiej indywidualności, gdy w badaniach społecznych stwierdza się, że ich religijność jest w znacznym stopniu funkcją ich płci, wieku, stanu cywilnego i klasy społecznej. Każda religijna osoba od razu zakomunikuje nam, że siła jej przekonań bierze się z wielu innych rzeczy. I tak właśnie jest, jak wykazałby model idiograficzny użyty w przypadku każdego pojedynczego człowieka. Czy jednak model idiograficzny jest choć trochę mniej dehumanizujący od modelu nomotetycznego?

Jeśli cecha taka jak religijność (lub bycie wolnym od uprzedzeń) może być wyjaśniana w kategoriach podstawowych, warunkujących ją czynników, to czy w jakimkolwiek stopniu bardziej dehumanizujące jest poszukiwanie cząstkowych, lecz ogólnych wyjaśnień przy wykorzystaniu tylko kilku z tych czynników, od poszukiwania pełnego wyjaśnienia z użyciem ich wszystkich? Logicznie rzecz biorąc, pełne ujęcie czyjejś religijności wciąż zakłada, że jest ona funkcją innych przyczyn – tyle że może ich być bardzo wiele. Prawdopodobnie prawdziwą przyczyną przykrości wywoływanej przez wyjaśnienia nomotetyczne jest to, że deterministyczny charakter tego modelu jest bardziej oczywisty. Jednak staranne zestawienie wszystkich tych powodów, dla których dana osoba jest religijna, albo dla których zagłosowała na kandydata X, albo dla których jest wolna od uprzedzeń, wymaga perspektywy deterministycznej. Pod tym względem model idiograficzny jest nie mniej deterministyczny niż model, który pozwala nam na wyszczególnienie czterech zmiennych będących ogólnie najważniejszymi przyczynami religijności wśród ludzi.

Różnica między wyjaśnianiem idiograficznym a nomotetycznym wiąże się z innym rozróżnieniem przedstawionym po raz pierwszy w rozdziale 1: różnicą między danymi jakościowymi a ilościowymi. Dane jakościowe, obejmujące informacje szczegółowe i w większym stopniu pogłębione, dają się łatwo zastosować w wyjaśnieniach idiograficznych. Z kolei dane ilościowe są właściwsze w wyjaśnieniach nomotetycznych. A więc na przykład wywiad pogłębiony z jednym bezdomnym może przynieść pełne (idiograficzne) zrozumienie przyczyn jego obecnego stanu, podczas gdy analiza ilościowa może powiedzieć nam, czy to wykształcenie, czy płeć pozwala lepiej (nomotetycznie) przewidywać bezdomność.

Podsumowując, zarówno idiograficzny, jak i nomotetyczny tryb wyjaśniania wiąże się z pojęciem przyczynowości. Oba są uprawnione i pożyteczne. To, czego dowiemy się o przypadkach indywidualnych, może ponadto zasugerować nam ogólne, przyczynowe zależności między zmiennymi, a znów te ogólne związki mogą nam pomóc ukierunkować analizę poszczególnych przypadków. Najkorzystniejsze jest łączenie tych dwóch podejść – jeżeli nie w jednym badaniu, to w obrębie całej społeczności badaczy.

Kryteria przyczynowości

Jak już wskazałem, większość – choć w żadnym razie nie całość – badań społecznych jest nastawiona ostatecznie na odkrywanie przyczyn zjawisk społecznych. Ale sam fakt, że zauważamy pewien związek między zmiennymi, nie dowodzi jeszcze, iż jest to związek przyczynowo-skutkowy. Gdybyśmy dowiedzieli się, że większość zarejestrowanych demokratów woli lody truskawkowe od waniliowych, mielibyśmy prawo podejrzewać, że prawdopodobnie to nie ich preferencje polityczne powodują, że wolą truskawki od wanilii, a tym bardziej, że to nie ich upodobanie do truskawek czyni ich demokratami. Najprawdopodobniej każdy taki zaobserwowany związek między tymi zmiennymi byłby czysto przypadkowy.

Podobnie, moglibyśmy dowiedzieć się, że ludzie z wyższym wykształceniem częściej mają luksusowe samochody niż ludzie gorzej wykształceni. Możemy być całkiem pewni, że posiadanie luksusowego samochodu nie sprawi, że dana osoba otrzyma lepsze wykształcenie, a nie istnieje też żaden oczywisty powód, dla którego lepsze wykształcenie, samo w sobie, miałoby być przyczyną posiadania przez ludzi luksusowych samochodów.

W tym miejscu można by wtrącić, że ludzie lepiej wykształceni są zazwyczaj w stanie zarabiać więcej pieniędzy niż ich gorzej wykształceni rówieśnicy i że to tłumaczy wyższy odsetek wykształconych ludzi mających luksusowe samochody. Jest to w istocie wiarygodne wyjaśnienie przyczynowe. Powrócimy wkrótce do tego rozróżnienia zaobserwowanych związków i wyjaśnień przyczynowych, które za nimi stoją. Teraz zaś chodzi o to, że aby określić związek przyczynowy, potrzebne jest coś więcej niż tylko stwierdzenie zależności między zmiennymi.

Jak widzieliśmy, nauka jest działalnością wymagającą zarówno obserwacji, jak i logiki. W powyższych przykładach logika sugerowała, że zaobserwowane relacje czy związki nie mają przyczynowego charakteru. A więc przy użyciu jakich kryteriów badacze społeczni określają, że jedna rzecz jest przyczyną drugiej?

Weźmy najpierw wyjaśnienia idiograficzne. Joseph Maxwell (1996, s. 87–88) twierdzi, że główne kryteria oceniania trafności wyjaśnienia są następujące: (1) jego wiarygodność oraz (2) to, czy wyjaśnienia alternatywne („hipotezy konkurencyjne") zostały poważnie rozważone i nie zdały egzaminu. Pierwsze kryterium odnosi się do logiki jako jednej z podstaw nauki: wymagamy, by nasze wyjaśnienia miały sens, nawet jeśli czasem ich logika jest złożona. Drugie kryterium przypomina nam o słynnym powiedzeniu Sherlocka Holmesa: gdy wszystkie możliwości zostały wyeliminowane, ta, która pozostaje, musi być prawdziwa.

Zatem jeśli zaproponuję wyjaśnienie czyjejś religijności, które wydaje się pełne, to moje ujęcie, jeśli jest sensowne, zdaje pierwszy test. Jeżeli jednak nie rozważyłem innych wyjaśnień, ktoś może zaproponować inne, równie wiarygodne. Zanim jednego z nich nie uzna się za wadliwe, nikt nie może twierdzić, że odkrył „prawdę". Być może wyjaśnienie najpełniejsze i najbardziej poprawne łączyć będzie w sobie cechy obu.

Jeśli chodzi o wyjaśnianie nomotetyczne, Paul Lazarsfeld (1959) zaproponował trzy szczegółowe kryteria przyczynowego związku między zmiennymi. (W kolejnych częściach tej książki zobaczymy, jak kryteria te wykorzystuje się w praktyce badawczej). Pierwszy wymóg brzmi, że przyczyna poprzedza skutek. Nie ma żadnego sensu wyobrażanie sobie, że coś zostało spowodowane przez coś innego, co nastąpiło później. Jest jasne, że kula wylatując z lufy pistoletu nie spowodowała, że proch strzelniczy wybuchł; działa to odwrotnie. Posiadanie luksusowego samochodu nie jest przyczyną tego, że zarabia się wystarczająco dużo, by sobie na niego pozwolić.

Choć to kryterium wydać się może proste i oczywiste, to zorientujesz się, że przy analizie danych w naukach społecznych stwarza ono niekończące się problemy. Często kolejność dwóch zmiennych jest po prostu niejasna. Co pojawia się wcześniej: uprzedzenia czy cecha znana jako autorytaryzm (nadmierna uległość wobec władzy, której towarzyszy sztywność myślenia)? Nawet jeżeli kolejność w czasie wydaje się zasadniczo jasna, często można znaleźć wyjątki. Na przykład z reguły twierdzilibyśmy, że poziom wykształcenia rodziców będzie stanowił przyczynę poziomu wykształcenia ich dzieci. A przecież niektórzy rodzice mogą wrócić do szkoły, ponieważ ich dzieci zdobyły lepsze wykształcenie.

Drugim wymogiem stawianym związkom przyczynowym jest empiryczna korelacja między zmiennymi (*empiryczna* oznacza „oparta na doświadczeniu"). **Korelacja** zachodzi między zmiennymi wówczas, gdy zaobserwujemy, że są one powiązane. To znaczy, że gdy jedna z nich pojawia się lub zmienia, to samo dzieje się z drugą. Nie miałoby sensu twierdzenie, że wybuchający proch strzelniczy spowodował wylatywanie kul z lufy pistoletu, jeżeli w obserwowanej rzeczywistości kule nie wylatywałyby po tym, jak proch wybuchł (albo jeżeli wylatywałyby one nawet wtedy, gdy proch nie wybuchł).

Badania społeczne popadają w trudności wynikające także z tego na pozór oczywistego wymogu. Istnieje niewiele doskonałych korelacji, przynajmniej w probabilistycznym świecie wyjaśnień nomotetycznych. W 1996 roku większość konserwatystów zagłosowała na Boba Dole'a, lecz nie wszyscy. Na kandydatów feministycznych w 1992 roku zagłosowało więcej kobiet niż mężczyzn, choć niektóre kobiety tego nie zrobiły – w przeciwieństwie do wielu mężczyzn. Prowadzi to nas nieuchronnie do pytania, jak silny musi być empiryczny związek, aby można go było uznać za przyczynowy. Powrócimy do tego problemu w kolejnych częściach książki.

Trzeci wymóg zależności przyczynowej mówi, że skutek nie może być wyjaśniany w kategoriach jakiejś trzeciej zmiennej. Widzieliśmy już przykład zjawiska „trzeciej zmiennej" w przypadku wykształcenia i posiadania luksusowego samochodu. Weźmy inny przykład. Istnieje korelacja między sprzedażą lodów a zgonami z powodu utonięć: im więcej sprzedaje się lodów, tym więcej jest utonięć, i odwrotnie. Nie istnieje jednak bezpośredni związek między lodami a utonięciami. Trzecią zmienną, która tu oddziałuje, jest pora roku lub temperatura. Większość zgonów przez utonięcie zdarza się latem – w szczytowym okresie sprzedaży lodów.

Oto kilka innych przykładów **zależności pozornych** czy takich, które nie są prawdziwe. Istnieje negatywny związek między liczbą mułów w różnych miejscowościach a liczbą osób mających doktorat: im więcej mułów, tym mniej doktorów, i odwrotnie. Pewnie przychodzi ci na myśl zmienna, która wyjaśniłaby ten rzekomy związek. Odpowiedź brzmi: wiejskie albo miejskie otoczenie. Jest więcej mułów (i mniej osób z doktoratem) w okolicach wiejskich, podczas gdy w miastach – wprost przeciwnie.

Albo rozważmy dodatnią korelację między rozmiarem obuwia a zdolnościami matematycznymi u uczniów. Tu trzecią zmienną, która wyjaśnia ten zagadkowy związek, jest wiek. Średnio rzecz biorąc, starsze dzieci mają większe stopy i bardziej zaawansowane umiejętności matematyczne niż młodsze dzieci.

Rysunek 3.2 podsumowuje różnice między korelacją, związkami przyczynowymi i pozornymi zależnościami przyczynowymi. Obserwowane związki pokazane są za pomocą strzałek czarnych, a związki przyczynowe – przerywanych. Zauważ, że obserwowane zależności zachodzą w obu kierunkach. Zatem gdy jedna zmienna pojawia się lub zmienia, to samo robi druga. Lecz sam fakt istnienia związku nie mówi nam, która zmienna jest przyczyną której, a właściwie, czy w ogóle istnieje zależność przyczynowa między nimi. Jak się okazuje, w zestawieniu A naprawdę istnieje między zmiennymi związek przyczynowy: pierwsza zmienna jest przyczyną drugiej, co tłumaczy występowanie obserwowanego związku. W zestawieniu B obserwowany związek między rozmiarem buta a zdolnościami matematycznymi jest całkiem prawdziwy, choć przyczynowe łączenie tych zmiennych byłoby pozorne. W rzeczywistości to trzecia zmienna wyjaśnia obserwowany związek. Ramka zatytułowana „Korelacja i związek przyczynowy" dodatkowo ilustruje twierdzenie, że sama korelacja nie dowodzi określonego związku przyczynowego.

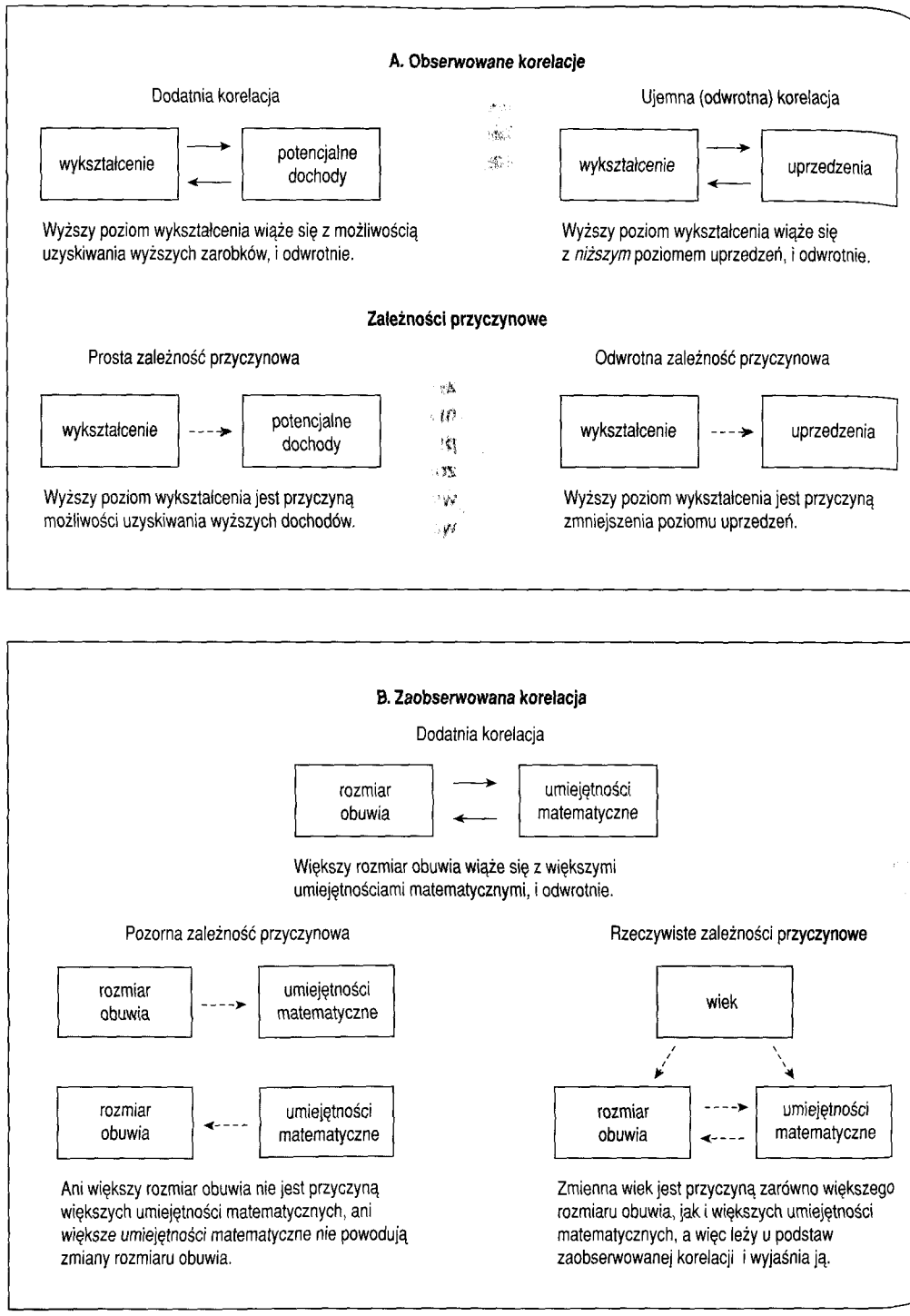
Jak przestrzegają John i Lyn Loflandowie (1995, s. 138–139), ważne jest, by odróżnić testowanie zależności przyczynowych od domysłów na ich temat. Choć odnotowywanie naszych przeczuć czy niesprawdzonych hipotez dotyczących procesów przyczynowych zachodzących w obserwowanych przez nas zjawiskach jest całkiem uprawnione, musimy odróżnić takie przypuszczenia od dowiedzionych wniosków. Nieco dalej wejdziemy w szczegóły tego, jak badacze udowadniają, że przypuszczalne związki przyczynowe są rzeczywiste.

Podsumowując, większość badaczy społecznych uznaje, że między dwiema zmiennymi występuje zależność przyczynowa, jeśli (1) przyczyna poprzedza skutek, (2) występuje między nimi korelacja empiryczna oraz (3) okazuje się, że zależność ta nie jest rezultatem działania jakiejś trzeciej zmiennej.

Jak to pokrótce omówimy, ludzie czasami stosują nieodpowiednie kryteria, wnioskując o zależnościach przyczynowych. Dlatego ważne jest, by pamiętać, że każdy związek spełniający opisane właśnie trzy kryteria ma charakter przyczynowy oraz że są to jedyne kryteria.

Rysunek 3.2

Korelacja, związki przyczynowe i pozorne zależności przyczynowe



Korelacja i związek przyczynowy

Charles Bonney

Wydział Socjologii, Eastern Michigan University

Wielu ludzi (włączając w to czasem także badaczy, którzy przecież powinni być nieco mądrzejsi), dowiódłszy związku statystycznego między hipotetyczną „przyczyną" a jej domniemanym „skutkiem", zbyt ochoczo ogłasza „dowód" istnienia zależności przyczynowej. Weźmy jakiś przykład, aby zobaczyć, dlaczego „niekoniecznie tak musi być".

Wyobraź sobie, że przeprowadziłeś badania na studentach i odkryłeś odwrotną korelację między paleniem marihuany (zmienna M) a średnią ocen (zmienna O) – a więc ci, którzy palą, częściej mają niższą średnią niż ci, którzy nie palą, a im więcej ktoś pali, tym niższa jest jego średnia. Mógłbyś zatem twierdzić, że palenie marihuany obniża oceny danej osoby (w formie symbolicznej: M→O), podając na przykład wyjaśnienie, że marihuana źle wpływa na pamięć, co oczywiście szkodzi ocenom.

Jeżeli jednak oprócz odwrotnej korelacji nie masz żadnych dowodów, istnieje inna możliwość. Otrzymywanie słabych stopni jest frustrujące; frustracja często prowadzi do zachowań ucieczkowych; „upalenie się" jest rozpowszechnionym sposobem ucieczki; *ergo*, złe stopnie są przyczyną palenia marihuany (O→M)! Zanim nie ustalisz, co było pierwsze, palenie czy niskie stopnie, wyjaśnienie to jest tak samo zadowalające jak poprzednie.

Wprowadźmy inną zmienną do tego obrazka: obecność i/lub skalę problemów emocjonalnych (zmienna E). Z pewnością można przekonująco dowodzić, że problemy emocjonalne mogą prowadzić do zachowań ucieczkowych, łącznie z paleniem marihuany. Również sensowna wydaje się sugestia, że problemy emocjonalne mogą źle wpływać na oceny. Korelacja między paleniem marihuany a złymi stopniami może istnieć z tego samego powodu, z jakiego zakatarzone nosy i chore gardła zazwyczaj pojawiają się razem – *żadne* nie jest przyczyną tego drugiego, raczej oba są konsekwencją działania jakiejś trzeciej zmiennej ($E \nearrow M, E \searrow O$). Jeśli nie możesz wykluczyć takiej trzeciej zmiennej, dane potwierdzają to wyjaśnienie w tym samym stopniu, co wyjaśnienia poprzednie.

Ponadto, może studenci palą marihuanę dlatego, że ich przyjaciele palą, a otrzymują złe stopnie, ponieważ po prostu nie są tak bystrzy lub tak dobrze przygotowani, lub tak pilni, jak koledzy, natomiast fakt, że w obu przypadkach chodzi o tych samych studentów z twojej próby, jest czystym zbiegiem okoliczności. Jeżeli uzyskana przez ciebie korelacja nie jest wystarczająco silna i wystarczająco spójna, aby zbieg okoliczności był wysoce nieprawdopodobny, to również tej ostatniej możliwości, mimo iż nie ma oparcia w zebranych przez ciebie danych, nie można z góry wykluczyć. Nawiasem mówiąc, przykład ten wybrałem z dwóch powodów. Po pierwsze, *każde* z powyższych wyjaśnień takiej ujemnej korelacji pojawiło się kiedyś w krajowej prasie. A po drugie, *każde* z nich jest prawdopodobnie skazane na porażkę, ponieważ, jak się okazuje, większość badań wykazała dodatnią korelację: wśród studentów osoby z wyższą średnią częściej są palaczami marihuany! A zatem z przymrużeniem oka możemy powtórnie przeanalizować nasze wyniki:

1. Marihuana działa na człowieka odprężająco, usuwa pozostałe stresy, pozwalając tym samym na bardziej wydajną naukę, zatem M→O
   lub
2. Marihuana jest stosowana jako nagroda za porządne kucie czy za sukcesy („O rany, człowieku! Dostałeś 6! Chodź się upalimy"); zatem O→M.
   lub
3. Większa ciekawość świata (C) jest rzeczą zdecydowanie cenną przy zdobywaniu wiedzy i dobrych stopni, ale może jednocześnie skłaniać człowieka do badania tematów tabu; zatem $C \nearrow M, C \searrow O$
   lub
4. Znów zbieg okoliczności, lecz tym razem zdarzyło się, że do próby weszło po prostu bardzo dużo bystrych, pilniejszych studentów, których przyjaciele palą marihuanę.

Oczywisty wniosek jest taki: skoro *wszystko* to są możliwe wyjaśnienia związku między dwiema zmiennymi, zatem nie należy zbyt pospiesznie wyróżniać *jednego* z nich. Ustalenie, że dwie zmienne zazwyczaj pojawiają się razem, jest warunkiem *koniecznym* dla wykazania związku przyczynowego, lecz samo w sobie nie jest warunkiem *wystarczającym*. Jest na przykład faktem, że przyrost naturalny jest wyższy w tych częściach Europy, w których występuje dużo bocianów, jednak co do znaczenia tego związku...!

Przyczyny konieczne i wystarczające

Jednym z błędów popełnianych przez ludzi w odniesieniu do związków przyczynowych jest zakładanie, że wymagają one istnienia doskonałej korelacji między zmiennymi. W rzeczywistości, jeśli o to chodzi, doskonałe powiązanie nie jest kryterium przyczynowości w naukach społecznych ani w nauce jako takiej. Ujmując to inaczej, wyjątki, choć nie potwierdzają reguły, niekoniecznie jej zaprzeczają. W modelach probabilistycznych prawie zawsze istnieją wyjątki od zakładanej zależności. Jeśli niektórzy liberałowie zagłosowali na Boba Dole'a, a niektórzy konserwatyści na Billa Clintona, to przypadki te nie zaprzeczają ogólnemu związkowi przyczynowemu między orientacją polityczną a głosowaniem w wyborach.

W ramach tego probabilistycznego modelu przydatne jest rozróżnienie przyczyn koniecznych i wystarczających. *Przyczyna konieczna* oznacza warunek, który *musi* być spełniony, aby nastąpił określony skutek. Na przykład aby otrzymać stopień, konieczne jest chodzenie na zajęcia. Gdyby zlikwidowano zajęcia, nigdy więcej nie byłoby stopni. Jednak zwykła obecność na zajęciach nie wystarcza, by otrzymać stopień. Należy wybrać odpowiednie zajęcia i zaliczyć je. Podobnie jak bycie kobietą jest koniecznym, choć niewystarczającym warunkiem zajścia w ciążę. W przeciwnym razie wszystkie kobiety zachodziłyby w ciążę.

Rysunek 3.3 przedstawia związek między zmiennymi „płeć" i „ciąża" w postaci macierzy, która pokazuje, jakie są możliwe wyniki ich kombinacji.

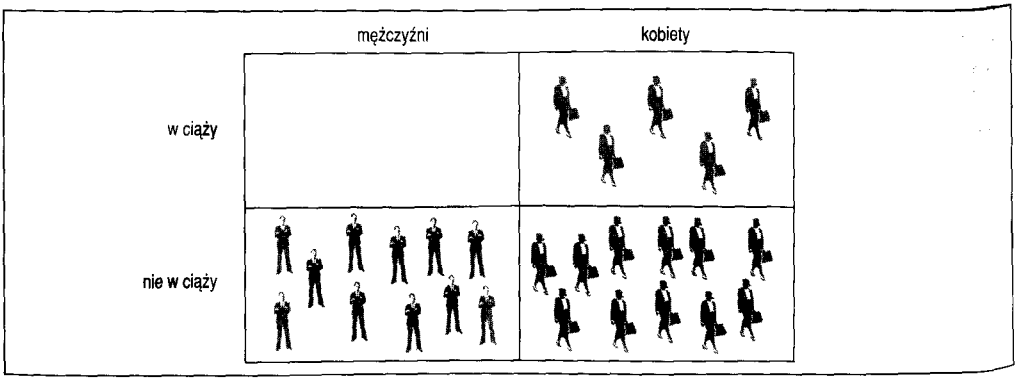
Z kolei *przyczyna wystarczająca* oznacza warunek, który, jeżeli zostanie spełniony, gwarantuje oczekiwany skutek. Nie należy przez to rozumieć, że wystarczająca przyczyna jest *jedyną* możliwą przyczyną danego skutku. Na przykład ucieczka z egzaminu kończącego zajęcia byłaby wystarczającą przyczyną oblania go, choć studenci mogą go oblać również na inne sposoby. Zatem jakaś przyczyna może być wystarczająca, choć nie konieczna. Rysunek 3.4 pokazuje związek między pójściem lub niepójściem na egzamin a zdaniem go lub oblaniem.

Odkrycie przyczyny *zarówno* koniecznej, *jak i* wystarczającej jest oczywiście wynikiem najbardziej satysfakcjonującym badacza. Jeżeli przestępczość młodocianych badalibyśmy jako skutek, miło byłoby odkryć pojedynczy warunek, który (1) musi być spełniony, aby przestępczość się rozwijała, i (2) którego konsekwencją zawsze są przestępstwa. W takim przypadku badacz z pewnością czułby, że dokładnie wie, co jest przyczyną przestępczości wśród młodzieży.

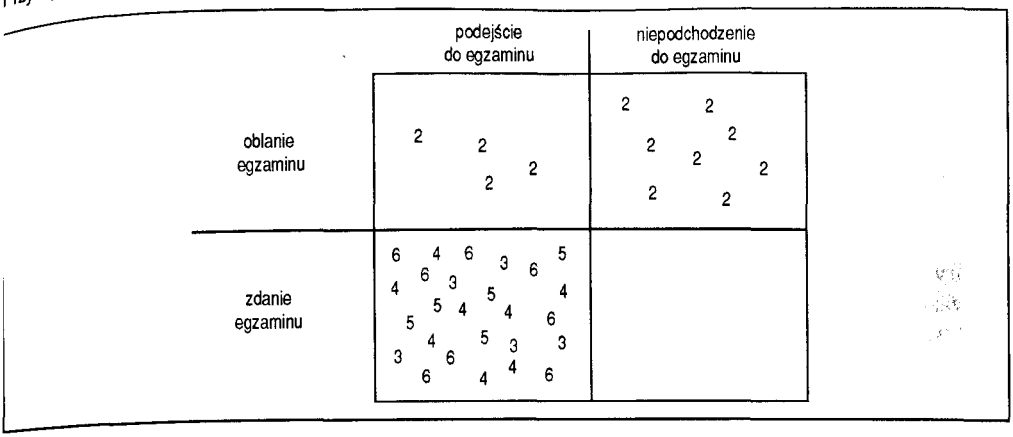
Niestety gdy analizujemy nomotetyczne związki między zmiennymi, nigdy nie odkrywamy pojedynczych przyczyn, które są absolutnie konieczne i absolutnie wystarczające. Nierzadko jednak wśród przyczyn znajdujemy czynniki, które są albo w 100% konieczne (trze-

Rysunek 3.3

Przyczyna konieczna



Rysunek 3.4
Przyczyna wystarczająca



ba być kobietą, aby zajść w ciążę), albo w 100% wystarczające (ucieczka z egzaminu nieodwołalnie spowoduje oblanie go).

W idiograficznej analizie pojedynczych przypadków można dojść do tak pogłębionego wyjaśnienia, że pozwoli ono w sposób uzasadniony twierdzić, iż sprawy nie mogły potoczyć się inaczej. Sugeruje to, iż określone zostały wystarczające przyczyny danego skutku. (Każdy, kto miałby genetyczne dziedzictwo we wszystkich szczegółach identyczne z twoim, identycznie zostałby wychowany i miał takie same późniejsze doświadczenia – skończyłby na studiach). Takie idiograficzne wyjaśnienia z definicji nie zakładają pojedynczych przyczyn i zawsze mogą istnieć inne przyczynowe ścieżki prowadzące do tego samego rezultatu. Zatem przyczyny są wystarczające, lecz nie są konieczne.

Ogólnie w naukach społecznych przedstawienie albo koniecznych, albo wystarczających przyczyn – nawet niedoskonałych – może być podstawą wniosku, że zmienne są powiązane przyczynowo.

Błędy w rozumowaniu na temat związków przyczynowych

Jak widzieliśmy, przyczyna i skutek mają zasadnicze znaczenie w wyjaśnianiu naukowym. Są również podstawą w naszym życiu codziennym. Jednak w naszych stwierdzeniach dotyczących przyczyn często popełniamy błędy. Przyjrzyjmy się kilku przykładom takiego wadliwego rozumowania. Przy każdym z nich, zanim przeczytasz moje wyjaśnienie, sprawdź, czy sam potrafisz wykryć błąd.

Na początku historii epidemii AIDS artykuł w „San Francisco Chronicle" donosił o badaniach, w których stwierdzano związek między AIDS a fluoryzacją wody pitnej. Częściowo podstawą tego twierdzenia była obserwacja, że „mimo że zasoby wodne dla połowy wszystkich miejscowości w kraju są fluoryzowane, a dla połowy nie, 90% przypadków AIDS pochodzi z terenów fluoryzowanych, a tylko 10% z terenów niefluoryzowanych". Czy dostrzegasz jakieś wady tego rozumowania? Zastanów się nad tym przez chwilę, zanim przejdziesz dalej.

Zacznijmy od tego, że zawsze powinniśmy uważać, gdy korzystając z danych dotyczących zbiorowości wyprowadza się wnioski dotyczące jednostek. W tym przykładzie „połowa wszystkich miejscowości w kraju" może nie oznaczać połowy populacji. Właściwie łatwo sobie wyobrazić, że duże miasta będą fluoryzowały wodę częściej niż małe, wiejskie miejscowości. Z punktu widzenia logiki byłoby możliwe, że korzystające z fluoryzacji społeczności stanowiłyby 90% populacji w kraju, w której to sytuacji powinny mieć 90% przypadków AIDS, nawet jeżeli nie istniałby żaden związek między AIDS a fluoryzacją.

Poza tym postępowe wartości społeczne, bardziej powszechne w miastach niż w małych miejscowościach, mogłyby wpływać zarówno na (1) style życia łączone z AIDS, jak i na (2) decyzję o fluoryzacji zasobów wodnych. Wówczas AIDS i fluoryzacja mogłyby być statystycznie skorelowane, nie będąc powiązane zależnością przyczynową.

Problemy wiążące się z błędnym rozumowaniem dotyczącym przyczyn nie ograniczają się do analiz statystycznych. Badania procesów historycznych są na nie tak samo podatne. Rozważmy taki przykład: na konferencji prasowej 5 sierpnia 1985 roku ówczesny prezydent Ronald Reagan odnotował 40 rocznicę zrzucenia bomby na Hiroszimę. Jak to ujął, ten straszliwy przykład zagłady nuklearnej posłużył za „środek odstraszający, który sprawił, że zachowaliśmy pokój przez najdłuższy czas w historii, 40 lat pokoju". Wielu ludzi opowiadających się za natychmiastowym zakończeniem wyścigu zbrojeń szybko zaprzeczyło istnieniu związku przyczynowego stwierdzonego przez prezydenta, mówiąc, że to nie zbombardowanie Hiroszimy utrzymało pokój. Czy potrafisz dostrzec kolejną skazę w twierdzeniu o związku przyczynowym?

Najbardziej wątpliwa nie jest tu przyczyna „40 lat pokoju", lecz samo istnienie tej błogiej ery. W okresie, o którym mowa, wydarzyły się takie wielkie konflikty, jak wojna w Korei lub wojna w Wietnamie – nie wspominając o mnóstwie mniejszych zbrojnych konfrontacji na całym świecie. Trudno jest utrzymywać, że zbombardowanie Hiroszimy było przyczyną pokojowego okresu, skoro nie było żadnego trwałego pokoju.

Czasem można wykryć błędy w rozumowaniu dotyczącym przyczyn, opierając się wyłącznie na logice, nawet jeżeli nie ma możliwości przeanalizowania danych empirycznych. Ponownie przykład podsuwa nam kwestia AIDS. Niektórzy ludzie przekonywali, że epidemia AIDS odzwierciedla niezadowolenie Boga z pewnych ludzkich zachowań, w szczególności homoseksualizmu i zażywania narkotyków. Ewangelista Don Boys wypowiedział się na temat seksu, redagując gościnnie artykuł redakcyjny w „USA Today" (z 7 października 1985 roku):

> „Epidemia AIDS wskazuje, że moralność zerwała swe cumy i dryfuje w kierunku bagiennych wyziewów, wytwarzając choroby, degenerację i śmierć... Według Bożego planu każdy człowiek ma mieć jedną kobietę – swoją żonę – na całe życie i ma być jej wierny".

Pomyłka nie polega tu na tym, że AIDS rozpowszechnia się coraz bardziej wśród heteroseksualnych par małżeńskich lub że młodzi ludzie cierpiący na hemofilię zarażają się AIDS przez transfuzję zainfekowanej krwi. Uprawnione jest stosowanie w tym kontekście probabilistycznego modelu przyczynowości, a homoseksualiści rzeczywiście częściej zarażają się AIDS niż heteroseksualiści. Jeżeli jednak AIDS jest oznaką „Bożego planu", to lesbijki muszą być grupą najbardziej faworyzowaną ze wszystkich, ponieważ wśród nich AIDS jest najrzadsze.

Żaden z tych przykładów wadliwego rozumowania przyczynowego nie ma być oskarżeniem o przestępstwo. Wszyscy popełniamy takie błędy; powiada się, że problem ze „zdrowym rozsądkiem" polega na tym, że nie jest on wcale tak powszechny, jak się sądzi. Natomiast cel tych przykładów był dwojaki. Po pierwsze, mogą one nas uwrażliwić na błędne rozumowanie przyczynowe, ujawniając niektóre sposoby jego przejawiania się w codziennym życiu. Rozbudzanie w sobie głębokiej świadomości takiego wadliwego rozumowania może pomóc nam rozpoznać podobne błędy w rozumowaniu innych i uniknąć popełniania ich przez nas samych. Po drugie, przykłady te służą jako tło dla zrozumienia siły starannej, naukowej argumentacji. Choć naukowcy nie są uodpornieni na błędy logiczne, poznanie naukowych procedur daje pewien stopień zabezpieczenia.

Wprawdzie nie ma żadnego prostego zestawu reguł logicznego myślenia, lecz znakomita książka Howarda Kahane'a *Logic and Contemporary Rhetoric* (1992) przedstawia wiele błędów powszechnie popełnianych przez ludzi. Oto niektóre z pułapek omówionych przez Kahane'a.

Prowincjonalizm

Każdy z nas patrzy na świat przez pryzmat swojej indywidualnej historii i obecnej sytuacji. Za-

wsze więc istnieje niebezpieczeństwo, że badacze będą interpretować ludzkie zachowanie tylko w taki sposób, jaki z ich punktu widzenia jest sensowny. Świat może wyglądać całkiem inaczej dla kobiet i dla mężczyzn, dla osób, które wyrosły w różnych klasach społecznych, dla ludzi religijnych i niereligijnych itd. Problem ten jest szczególnie wyraźny w badaniach międzykulturowych.

Jednocześnie Harry Wolcott (1995, s. 164–165), omawiając „sztukę badań terenowych", sugeruje, że nasze osobiste i kulturowe obciążenia mogą zdziałać wiele dobrego. Mogą skierować uwagę na jakiś problem badawczy i dostarczyć nowego spojrzenia, rozwijając punkt widzenia inny od tego, który mają obserwowane osoby. Kluczowa jest jednak świadomość swych osobistych poglądów i zachowanie gotowości do poszerzenia swych perspektyw przez uznanie faktu, że przekonania innych mogą być równoprawne.

Pochopne wnioski

Badacze, podobnie jak inni ludzie, są podatni na wyciąganie pochopnych wniosków. Gdy badacz proponuje jakąś interpretację danych, upewnij się, że oceniłeś „wagę" dowodów wskazujących na tę interpretację. Czy wniosek, zważywszy na stojące za nim dane, jest istotnie nieuchronny, czy też inne wnioski są tak samo uzasadnione? Należy również rozważyć wszystkie alternatywne wnioski, które mogłyby wynikać z twoich własnych danych.

Wątpliwa przyczyna

Kiedy tylko wydaje ci się, że *X* jest przyczyną *Y*, spytaj sam siebie, czy koniecznie musi tak być. Co jeszcze mogło spowodować *Y*? Kahane (1992, s. 63) podaje kilka przykładów ze świata ekonomii. Jeśli przedsiębiorstwo bankrutuje, ludzie często dochodzą do wniosku, że jego prezesowi zabrakło talentu do interesów – nawet gdy do bankructwa doszło w czasie poważnej recesji odznaczającej się upadkiem wielu przedsiębiorstw. Łatwo jest czasem podać wiarygodne wyjaśnienia przyczynowe, jednak ustalenie prawdziwej przyczyny wymaga wyeliminowania innych możliwych wyjaśnień.

Przemilczane dowody

W badaniach terenowych, w wyniku bezpośredniej obserwacji, wywiadów, pracy w bibliotece itd., gromadzi się mnóstwo informacji. Sformułowanie wniosków wymaga od nas odrzucenia części informacji oraz przeprowadzenia selekcji. Na ogół badacz usuwa informacje, które są „bez związku"; aczkolwiek jest to oczywiście samo w sobie kwestią oceny.

Czytając sprawozdania z badań, zwróć uwagę na te obserwacje, które, jak odnotowałeś, nie znalazły się później we wnioskach, a także na te obserwacje, o których nie wspomniano, choć, jak możesz mniemać, zostały poczynione. Przypuśćmy, że jakiś badacz wysunął wniosek, że członkowie grup neonazistowskich są wrogo nastawieni do Afroamerykanów ze strachu przed konkurencją ekonomiczną. Jeśli wniosek ten jest poprawny, oczekiwalibyśmy, że większość z tych członków będzie należała do klasy robotniczej lub niższej klasy średniej. Jeśli badacz nie wskazał na zawody członków tej grupy, możemy łatwo zakwestionować dowody, na których wspiera się jego wniosek.

Fałszywy dylemat

Wnioski badaczy, podobnie jak opinie nienaukowe, często oznaczają wybór jednego stanowiska z wielu. Często wybór jednego okazuje się jednoznaczny z odrzuceniem pozostałych, jednak nie musi tak być. *Fałszywy dylemat* to wybór, do którego jesteśmy zmuszeni tylko na pozór, a nie naprawdę. Kahane proponuje taki przykład: „ekonomia, a nie biologia, może wyjaśnić męską dominację". To śmiałe stwierdzenie odrzuca, jak się zdaje, wpływ polityki, wykształcenia, obyczajów, religii i mnóstwa innych możliwości. Jak ujmuje to Kahane (1992, s. 42):

> „Stwierdzenie to sugeruje, że istnieją tylko dwie możliwości: męską dominację (zauważmy, że sprawa jest przesądzona!) wyjaśnia albo biologia, albo sukces ekonomiczny. Sugeruje też, że ta druga możliwość, sukces ekonomiczny, «może» (chytre słowo) być prawdziwym wyjaśnieniem

męskiej dominacji. Jednak istnieje wiele innych możliwości, takich jak obyczaj społeczny, przekonania religijne oraz różne kombinacje czynników ekonomicznych i biologicznych. Kusząc nas, byśmy myśleli, że przyczyna męskiej dominacji jest albo ekonomiczna, albo biologiczna, cytat ten prowadzi do przeoczenia innych możliwości, a tym samym do popełnienia przez nas błędu polegającego na uznawaniu fałszywego dylematu".

David Silverman (1993, s. 205) wygłasza podobne ostrzeżenie, mówiąc, że podstawowa zasada przy analizie danych jakościowych brzmi: „nigdy nie odwołuj się do pojedynczego elementu jako wyjaśnienia". Nauka z tego taka: zawsze szukaj alternatywnych lub dodatkowych przyczyn. Powinniśmy wystrzegać się tej pułapki, czytając prace innych, ale również strzec się, byśmy nie wpadli w nią sami.

Połączenie pomiaru z ustalaniem związków

Jak już wiemy, jednym z koniecznych kryteriów stosowanych przy stwierdzaniu związku przyczynowego w nauce jest empiryczna korelacja między przyczyną a skutkiem. Ustalenie korelacji wymaga od nas pomiaru zmiennych i stwierdzenia, czy i jak są one powiązane. Nazbyt często jednak proces pomiaru zmiennych jest postrzegany w oddzieleniu od procesu ustalania związków między nimi. Przekonanie to, w najlepszym przypadku, jest zwodnicze. Korzystniej jest widzieć pomiar i zależności w ścisłym powiązaniu. Aby zobrazować ten pogląd, przyjrzyjmy się kilku przykładom problemów, które pojawiają się w trakcie definiowania, pomiaru i określania związków między zmiennymi w badaniach społecznych.

Praktyczne problemy przy pomiarze zmiennych i ustalaniu związków między nimi

Jak pamiętamy z rozdziału 2, tradycyjny model dedukcyjnego rozumowania przebiega od teorii do teoretycznych oczekiwań czy idei, które mają być sprawdzone w badaniu. Operacjonalizując zmienne, naukowiec dochodzi do sprawdzalnej hipotezy. Następnie zbiera dane i analizuje je, aby sprawdzić, czy hipoteza została potwierdzona.

Chociaż ten tradycyjny wizerunek badania naukowego jest pożytecznym, wyidealizowanym modelem, przyczynia się on do ukrywania pewnych praktycznych problemów, pojawiających się w pracy ze zmiennymi w trakcie rzeczywistych badań. Aby zobaczyć dlaczego, zastosujmy jakiś przykład.

Przypuśćmy, że badacz interesuje się zagadnieniem zachowań dewiacyjnych i tworzy teorię opartą na ogólnej socjologicznej teorii kontroli społecznej. Ta szeroko zakrojona teoria koncentruje się na przymusie i perswazji, jako na czynnikach ustanawiających i utrzymujących porządek społeczny. Jako taka, zawiera wiele różnych pojęć, które są odpowiednie dla teorii zachowań dewiacyjnych.

Przyjmijmy, że nasz badacz, na podstawie tej teorii, przewiduje, że przestępczość młodocianych jest funkcją nadzoru: wraz z nasileniem nadzoru maleje przestępczość. Następny krok to operacjonalizacja kluczowych zmiennych polegająca na szczegółowym określeniu, jak będą one mierzone. A więc, na przykład badacz może zoperacjonalizować „młodocianych" jako osoby poniżej 18 roku życia, „przestępczość" jako bycie zatrzymanym za popełnienie przestępstwa, a „nadzór" jako obecność w domu niepracującego dorosłego. Badacz ma teraz sprawdzalną hipotezę: wśród osób poniżej 18 roku życia ci, którzy mieszkają razem z niepracującym dorosłym, rzadziej będą aresztowani za przestępstwo niż ci, w których domu nie ma niepracującego dorosłego.

Badacz przechodzi teraz do zbierania danych empirycznych. Jedna z przykładowych technik polegałaby na przeprowadzaniu sondażu wśród osób poniżej 18 roku życia. W kwestionariuszu wywiadu pytano by ich między innymi o to, czy mieszkają w domu z niepracującym dorosłym i czy kiedykolwiek zostali zatrzymani za popełnienie przestępstwa.

Wreszcie badacz poddaje zebrane dane analizie statystycznej, aby sprawdzić postawioną

hipotezę. Testy statystyczne mają na celu wykazanie, czy w rzeczywistości – na podstawie zebranych danych – młodociani mieszkający z niepracującym dorosłym rzadziej niż inni są zatrzymywani za przestępstwa. Potwierdzenie lub obalenie tej hipotezy jest zatem wykorzystywane jako dowód służący zaakceptowaniu lub odrzuceniu teorii, z której hipoteza ta została wyprowadzona.

Tak jak w rozdziale 2, proces ten wydaje się zgrabny i logiczny. Ale istnieją dwa podstawowe problemy, które stoją na przeszkodzie łatwemu zastosowaniu tego modelu w praktyce.

Po pierwsze, pojęcia teoretyczne rzadko, o ile w ogóle, pozwalają na jednoznaczną operacjonalizację. Ponieważ pojęcia są abstrakcyjne i ogólne, każde uszczegółowienie wskaźników empirycznych musi być sformułowane w przybliżeniu. Stosując nasz przykład, mało prawdopodobne jest, że obecność niepracującego dorosłego w domu jest właściwym odpowiednikiem ogólnego pojęcia nadzoru. Sama bowiem obecność takiego dorosłego nie zapewnia, że młody człowiek jest rzeczywiście nadzorowany. Poza tym, w niektórych domach, gdzie brak jest takiego dorosłego, być może w inny sposób zorganizowano nadzór nad młodzieżą.

Zastanówmy się też nad definicją operacyjną „przestępcy". Bycie zatrzymanym za popełnienie przestępstwa z pewnością nie jest tym samym, co abstrakcyjne pojęcie przestępczości. Niektórzy młodociani mogą brać udział w przestępstwach nie będąc aresztowanymi; inni mogą być zatrzymani niesłusznie. Wreszcie, nawet określenie „młodocianych" jako osób poniżej 18 roku życia jest arbitralne. Można to było określić inaczej i prawdopodobnie żadna forma nie byłaby jednoznacznie poprawna. Problemem nie jest tu niekompetencja badacza. Każdy wskaźnik empiryczny ma jakieś wady; wszystko może zostać ulepszone, a poszukiwaniom lepszych wskaźników nie ma końca.

Po drugie, związki empiryczne między zmiennymi prawie nigdy nie są doskonałe. Na przykład, gdyby żaden młodociany mieszkający z niepracującym dorosłym nigdy nie był zatrzymany, a wszyscy bez takich dorosłych – byli, wówczas bezpiecznie wnioskowalibyśmy, że hipoteza została potwierdzona. Albo gdyby obie grupy miały identyczne wyniki, moglibyśmy wnioskować, że na podstawie danych hipoteza została odrzucona. Jednak w praktyce rezultat będzie mieszany. Prawie wszystkie zmienne są ze sobą w jakimś stopniu powiązane empirycznie. Jaki związek uznaje się za wystarczająco silny, aby pozwalał na trafne wnioski? W praktyce określanie szczegółowo poziomu zależności koniecznego, aby móc przyjąć lub obalić hipotezę, jest również arbitralnym aktem, w tym sensie, że również innych poziomów można by w sposób wiarygodny bronić. (Zagadnienie istotności statystycznej rozważymy szczegółowo w rozdziale 17).

Ostatecznie więc naukowcy stosują niedoskonałe wskaźniki teoretycznych pojęć, aby odkryć niedoskonałe zależności otwierające drogę dla niedoskonałych interpretacji. A te niedoskonałości spiskują ze sobą przeciwko nam. Przypuśćmy, że dokładnie określisz siłę związku, która decydować będzie o przyjęciu hipotezy, a empiryczna analiza nie spełni twoich oczekiwań. Całkiem zrozumiałe, że spytasz sam siebie, czy inne wskaźniki pojęć nie dałyby w efekcie związku odpowiednio silnego. Jeśli na przykład mierzyłeś przestępczość młodocianych jako bycie skazanym za przestępstwo, być może analiza przyniosłaby inne wyniki, gdyby jako wskaźnika użyć raczej zatrzymania niż skazania.

Główny morał, jaki z tego płynie, brzmi: pomiar i związek między zmiennymi są pojęciami wzajemnie od siebie uzależnionymi. Badacze muszą radzić sobie z oboma jednocześnie i czynić to w sposób logiczny. Zamiast przechodzić przez sztywno ustalone serie określonych kroków, w praktyce naukowcy nieustannie poruszają się wśród nich tam i z powrotem. Wszystkie hipotezy, definicje operacyjne i techniki statystyczne mogą być ulepszone, jeśli jeden „krok" potwierdza pozostałe. Teoretyczne konstrukcje są często budowane wokół wcześniej zaobserwowanych związków między empirycznymi wskaźnikami. Cząstkowe konstrukcje teoretyczne mogą podsunąć nowe dane empiryczne do analizy itd. Po każdej czynności badacze mają nadzieję nieco lepiej zrozumieć przed-

miot swych badań. Podobnie kolejne starania znajdują potwierdzenie w nowych badaniach, a nowe wyniki mogą zmienić interpretację poprzednich. „Krytyczny eksperyment", który za jednym zamachem przesądza o losie całej teorii, jest doprawdy rzadkością – w każdej nauce.

Przykład pomiaru i związku między zmiennymi

Jako ostatni przykład powiązań między pomiarem a związkiem między zmiennymi, rozważymy kontrowersję, która wynikła wiele lat temu, a dotyczyła związku między religijnością a uprzedzeniami. Charles Glock i Rodney Stark w książce zatytułowanej *Christian Beliefs and Anti-Semitism* (1967) relacjonowali badania empiryczne, wskazujące, że ortodoksyjni wyznawcy religii chrześcijańskiej (wierzący w Boga, Jezusa, Biblię itd.) częściej byli antysemitami niż chrześcijanie mniej ortodoksyjni. Jak można sobie wyobrazić, konkluzje tej książki wzbudziły poważną dyskusję wewnątrz kościołów. Wówczas inni badacze zaczęli się prześcigać w projektach na ten sam temat.

Jeden z tych projektów przyniósł wnioski całkowicie przeciwne w stosunku do wniosków Glocka i Starka. Badacze twierdzili, że wraz z nasilaniem się ortodoksji słabły uprzedzenia. Jednak po bardziej wnikliwej analizie okazało się, że ortodoksja była w tych drugich badaniach mierzona akceptacją przedstawionych w kwestionariuszu tez tradycyjnej chrześcijańskiej doktryny: „wszyscy ludzie są braćmi" i „kochaj swego bliźniego". Nie dziwi więc, że respondenci, którzy akceptowali te tezy, wydali się mniej uprzedzeni niż ci, którzy je odrzucili.

Wyniki tego badania powinny zostać (i zostały) podważone ze względu na „kontaminację": wskaźniki zastosowane przy operacjonalizacji ortodoksji religijnej i uprzedzeń w rzeczywistości mierzyły te same lub podobne cechy. Nazywanie jednego zestawu wskaźników „ortodoksją", a drugiego „uprzedzeniami" nie dowodzi, że w bardziej ogólnym sensie wraz z nasilającą się ortodoksją słabną uprzedzenia. (Oczywiście pomiar ortodoksji religijnej przy wykorzystaniu kategorii miłości braterskiej i równości może być niezmiernie przydatny w jakimś innym kontekście).

Rozważania tego rozdziału sugerują nieco inną reakcję na te dwa typy wyników badań. Pytając raczej o to, *jak* ortodoksja i uprzedzenia są ze sobą związane niż o to, *czy* są związane, moglibyśmy wysnuć następujący wniosek: (1) ortodoksja mierzona za pomocą wskaźników Glocka i Starka jest dodatnio związana z uprzedzeniami, a (2) ortodoksja mierzona w kategoriach przywiązania do norm miłości braterskiej i równości jest negatywnie związana z uprzedzeniami. Oba wnioski są empirycznie poprawne. Żaden nie odpowiada na bardziej ogólne pytanie, czy religia i uprzedzenia są ze sobą związane.

Ostatni krok, jaki nam pozostał, to oczywiście ocena przydatności tych wniosków. Odkrycie, że ortodoksja – zdefiniowana jako przywiązanie do norm miłości braterskiej – jest negatywnie związana z uprzedzeniami, zbliża się niebezpiecznie do prawdy logicznej czy tautologii. *Tautologia* jest to stwierdzenie, które jest prawdziwe z definicji: „ludzie o ortodoksyjnych poglądach (wierzą w miłość braterską) są nieuprzedzeni (wierzą w miłość braterską)".

Oczywiście podanie tego przykładu nie miało na celu rozstrzygnięcia kwestii religii i uprzedzeń: czytanie literatury im poświęconej może ci się spodobać i może nawet pewnego dnia sam będziesz prowadził w tej dziedzinie oryginalne badania. Natomiast przykład ten naprawdę ilustruje fakt, że procesy definiowania, pomiaru i ustalania związków między zmiennymi są wzajemnie powiązane. Innym wnioskiem jest to, że możesz zadać wiele wnikliwych pytań, zanim zaakceptujesz wyniki badań, o których czytasz w profesjonalnym piśmie, czy w powszechnie dostępnych mediach. Jak były mierzone zmienne? (Pamiętaj, że nie znaczą one nic więcej poza tym, w jaki sposób zostały zmierzone). Jak te pomiary, z punktu widzenia logiki, są ze sobą powiązane? Jak silny jest związek między zmiennymi? I czy wnioski, tak jak zostały sformułowane, dokładnie odzwierciedlają wyniki badań, czy też rozszerzają te wyniki w sposób nieuzasadniony zgromadzonymi danymi? Naturalnie wszyscy badacze powinni stawiać te same pytania także w odniesieniu do swej własnej pracy.

Główne zagadnienia

- Badania naukowe o charakterze wyjaśniającym zależą bezpośrednio od pojęcia przyczyny i skutku.
- Badania wyjaśniające w naukach społecznych zależą bezpośrednio od deterministycznego modelu ludzkiego zachowania.
- Zarówno idiograficzne, jak i nomotetyczne modele wyjaśniania oparte są na pojęciu związku przyczynowego. Model idiograficzny ma na celu pełne zrozumienie konkretnego zjawiska przy wykorzystaniu wszystkich adekwatnych czynników przyczynowych. Model nomotetyczny ma na celu wyjaśnienie ogólne – niekoniecznie pełne – klasy zjawisk przy wykorzystaniu niewielu odpowiednich czynników przyczynowych.
- Przy wyjaśnianiu idiograficznym zasadnicze kryteria ustalania związku przyczynowego to: (1) wiarygodność wyjaśnienia oraz (2) wykazanie, że inne możliwe wyjaśnienia należy odrzucić.
- Istnieją trzy główne kryteria przy ustalaniu związków przyczynowych w analizach nomotetycznych: (1) zmienne muszą być empirycznie powiązane czy też skorelowane, (2) zmienna będąca przyczyną musi występować wcześniej od zmiennej, na którą jakoby wpływa oraz (3) obserwowany skutek nie może być wyjaśniony jako skutek działania innej zmiennej.
- Sam związek czy korelacja nie stanowi jeszcze zależności przyczynowej. Pozorna zależność przyczynowa jest to związek, który w rzeczywistości jest spowodowany przez jedną lub wiele innych zmiennych.
- Doskonała zależność statystyczna między dwiema zmiennymi nie jest właściwym kryterium ustalania związku przyczynowego w badaniach społecznych. Możemy więc powiedzieć, że między *X* a *Y* istnieje związek przyczynowy, nawet jeżeli *X* nie jest wyłączną przyczyną *Y*.
- Większość wyjaśniających badań społecznych stosuje probabilistyczny model przyczynowości. O *X* można rzec, że jest przyczyną *Y*, jeżeli widać, że ma pewien wpływ na *Y*.
- Istnieją dwa główne typy przyczyn: konieczna i wystarczająca. *X* jest konieczną przyczyną *Y*, jeżeli *Y* nie może zdarzyć się bez *X*. *X* jest wystarczającą przyczyną *Y*, jeżeli *Y* zdarza się zawsze wtedy, gdy zdarza się *X*. Najbardziej z naukowego punktu widzenia satysfakcjonującym odkryciem jest przyczyna jednocześnie konieczna i wystarczająca.
- Badacze w rozumowaniu dotyczącym przyczyn muszą się strzec kilku powszechnych rodzajów błędów, do których należą: prowincjonalizm, pochopne wnioski, ustalenie wątpliwych przyczyn, przemilczanie dowodów i fałszywe dylematy.
- Choć dedukcyjny model nauki w swej wyidealizowanej postaci sugeruje, że definiowanie i pomiar zmiennych są oddzielone od określania związków między nimi, to owocne okazuje się dostrzeganie bliskich związków między tymi procesami.
- Zanim zaakceptujemy wnioski dotyczące związku przyczynowego, powinniśmy prześledzić, jak zmienne zostały zoperacjonalizowane, jaki jest logiczny związek między nimi, jak silna jest między nimi zależność oraz czy sformułowane wnioski dokładnie odzwierciedlają to, na co pozwalają wyniki badania.

Kluczowe pojęcia

Oto terminy, które w tym rozdziale zostały zdefiniowane w odpowiednich kontekstach, i które można też znaleźć w słowniczku na końcu książki:

korelacja

pozorna zależność

Pytania i ćwiczenia

1. Opisz warunki, które pozwalałyby ci wywnioskować, że wykształcenie jest:
   a) konieczną przyczyną ograniczenia uprzedzeń;
   b) wystarczającą przyczyną ograniczenia uprzedzeń;

c) konieczną i wystarczającą przyczyną ograniczenia uprzedzeń.

2. Dlaczego wybrałeś właśnie tę uczelnię, na której teraz studiujesz? Stwórz idiograficzne wyjaśnienie, podając możliwie wiele szczegółowych czynników, które doprowadziły do twojej decyzji.
3. Z idiograficznego wyjaśnienia twojego wyboru uczelni wybierz dwa lub trzy czynniki, które według ciebie będą najodpowiedniejsze do wyjaśnienia wyboru dokonanego przez wielu innych studentów. Przedyskutuj, co zyskuje się, a co traci przy wyjaśnianiu nomotetycznym (wyboru uczelni przez studentów) opartym na tych czynnikach, w porównaniu z twoim pierwotnym wyjaśnieniem.
4. W społeczeństwie amerykańskim kobiety zarabiają z grubsza 70% tego, co zarabiają mężczyźni. Jak przypuszczasz, co jest „przyczyną" tej różnicy? Opisz procedury, za pomocą których mógłbyś sprawdzić swoje domysły.
5. Wybierz temat, który cię interesuje i korzystając z sieci znajdź przynajmniej dwa popularyzatorskie sprawozdania z badań naukowych dotyczących związków przyczynowych (na przykład w gazetach codziennych, czasopismach, na stronach poświęconych zdrowiu). Sprawozdania takie są publikowane w mediach prawie każdego dnia. Zbadaj, czego możesz, a czego nie możesz się dowiedzieć z tych sprawozdań o tym, jak zmienne zostały zoperacjonalizowane i przetestowane oraz czy sprawozdania te właściwie przedstawiają uprawnione wnioski wynikające z danych. Czy zaakceptowałbyś te wnioski w takiej postaci? Jeśli nie, jakich informacji potrzebujesz, aby je rozsądnie ocenić? Czy widzisz jakieś wady w rozumowaniu lub w sposobie, w jaki media informują o wynikach badań?

Dodatkowe lektury

Beck E. M., Tolney Stewart E., *The Killing Fields of the Deep South: The Market for Cotton and the Lynching of Blacks, 1882–1930*, „American Sociological Review" 1990, 55, s. 526–539. Analiza strukturalnych przyczyn samosądów na Południu USA. Badacze społeczni często podejmują tego typu analizę związków przyczynowych. Niektóre z badanych zmiennych to inflacja, cena bawełny oraz procent czarnoskórych w populacji.

Davis James A., *The Logic of Causal Order*, Sage, Beverly Hills, Cal. 1985. Davis analizuje wymiar logiczny i statystyczny przyczynowości w naukach społecznych.

Hirschi Travis, Selvin Hanan, *Principles of Survey Analysis*, Free Press, New York 1973 (patrz szczególnie część II). Doskonałe twierdzenia dotyczące związku przyczynowego ujęte w praktyczne ramy. Nie znam lepszego omówienia problemu związku przyczynowego, który umieszczony byłby w kontekście wyników konkretnych badań. Książka dobrze się czyta, jest stymulująca, a ogólnie – po prostu znakomita.

Lazarsfeld Paul, *Foreword*, w: Herbert Hyman, *Survey Design and Analysis*, Free Press, New York 1955. Klasyczne i wciąż aktualne twierdzenia na temat przyczynowości w naukach społecznych. W kontekście modelu analizy rozbudowanej Lazarsfeld dostarcza jasno sformułowanych kryteriów ustalania związku przyczynowego.

Shaver Kelly G., *The Attribution of Blame: Causality, Responsibility and Blameworthiness*, Springer-Verlag, New York 1985. Shaver omawia wiele aspektów związków przyczynowych, przedstawionych w tym rozdziale i pokazuje, w jaki sposób wiążą się one z pojęciami odpowiedzialności i winy.

Wallace William A., *Causality and Scientific Explanation*, University of Michigan Press, Ann Arbor 1972. Gdybyś zainteresował się poważniej zagadnieniem przyczynowości, ta dwutomowa praca dostarcza kompletnego opisu historii tego pojęcia w obrębie nauki, od średniowiecza do czasów obecnych.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodat-

kowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Część
2

Struktura procesu badawczego

4. Plan badań

5. Konceptualizacja, operacjonalizacja i pomiar

6. Indeksy, skale i typologie

7. Logika doboru próby

Odpowiednie postawienie pytania jest często trudniejsze niż sama odpowiedź. W rzeczywistości odpowiednio zadane pytanie często zawiera w sobie część odpowiedzi. Możesz dotrzeć do odpowiedzi, gdy tłumaczysz komuś, o co w pytaniu naprawdę chodzi.

Część 2 jest poświęcona strukturze procesu badawczego, czyli odpowiedniemu zadawaniu pytań naukowych. Część 3 opisuje niektóre metody dokonywania obserwacji naukowych w badaniach społecznych.

Rozdział 4 dotyczy początków badań. Omawia on niektóre cele badań, jednostki analizy i obszary szczególnego zainteresowania badaczy społecznych, a także przyczyny, dla których uczeni podejmują projekty badawcze.

W rozdziale 5 zajmiemy się określeniem przedmiotu pomiaru, czyli procesami konceptualizacji oraz operacjonalizacji. Przyjrzymy się niektórym pojęciom, których używamy dość często w życiu codziennym – uprzedzenia, liberalizm, szczęście itd. – i pokażemy, jak ważne jest wyjaśnienie, co naprawdę mamy na myśli, stosując takie pojęcia w badaniach. Ten proces wyjaśniania i uściślania to *konceptualizacja*.

Gdy już określimy, co rozumiemy pod pewnymi pojęciami, możemy przystąpić do pomiaru zjawisk, do których się one odnoszą. Proces określania kolejnych kroków czy działań (operacji), służących zmierzeniu przedmiotu badań, nazywany jest *operacjonalizacją*. Rozdział 5 poświęcony jest ogólnie operacjonalizacji, ze zwróceniem szczególnej uwagi na formułowanie pytań, stosowanych w wywiadach i kwestionariuszach.

Kończąc wprowadzenie do zagadnień pomiaru, rozdział 6 zrywa z chronologicznym omawianiem sposobu prowadzenia badań. W rozdziale tym zajmiemy się technikami pomiaru zmiennych w badaniach ilościowych przy użyciu zestawu kilku wskaźników: indeksów, skal i typologii. Możemy na przykład zadać respondentom w sondażu pięć pytań dotyczących ich postaw wobec równości płci, a potem zebrać odpowiedzi na wszystkie pięć pytań w jeden złożony miernik postaw równościowych w odniesieniu do płci. Choć takie złożone mierniki konstruowane są w czasie analizy danych (patrz część 4), to jednak zebranie surowych danych, potrzebnych do ich stworzenia, musi być zapewnione w planie badań i podczas ich faktycznej realizacji.

Na końcu zobaczymy, jak przedstawiciele nauk społecznych dobierają ludzi (albo rzeczy) do obserwacji. W rozdziale 7 zajmiemy się doborem próby i fundamentalnym dla nauki zagadnieniem *uogólniania*. Jak się przekonasz, można wybrać do obserwacji kilka osób lub rzeczy, a wyniki obserwacji odnieść potem do znacznie szerszej grupy ludzi lub rzeczy. Na przykład pytając w sondażu 1000 obywateli USA, na kogo będą głosować w wyborach prezydenckich, możemy poprawnie przewidzieć, jak będą głosowały dziesiątki milionów. Rozdział ten poświęcony jest technikom zwiększającym *możliwości uogólniania* wyników obserwacji.

Dzięki temu, czego nauczysz się z części 2, osiągniesz pełnię wiedzy na temat prowadzenia kontrolowanych obserwacji badawczych. Część 3 pokaże, jak wykonać następny krok. ■

Rozdział 4

Plan badań

Przegląd holograficzny

Badacze zjawisk społecznych mają do dyspozycji bogaty wybór schematów badań. Zaplanowanie badań obejmuje dokładne określenie, kto lub co ma być badane, kiedy, jak oraz w jakim celu.

Wprowadzenie

Trzy cele badań
Eksploracja
Opis
Wyjaśnianie

Jednostki analizy
Jedostki indywidualne
Grupy
Organizacje
Wytwory społeczne
Przegląd jednostek analizy
Błędne wnioskowanie o jednostkach analizy: błąd ekologiczny oraz redukcjonizm

Wymiar czasowy
Badania przekrojowe
Badania dynamiczne
Badania zbliżone do dynamicznych
Przykłady strategii badawczych

Jak zaplanować projekt badawczy
Zaczynamy
Konceptualizacja
Wybór metody badawczej
Operacjonalizacja
Populacja i dobór próby
Obserwacje
Przetwarzanie danych
Analiza
Zastosowanie
Przegląd procesu planowania badań

Oferta badawcza
Elementy oferty badawczej

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Nauka jest przedsięwzięciem, którego celem jest „odkrywanie". Ale bez względu na to, co chce się odkryć, jest wiele dróg, którymi można podążać. Tak jak w życiu. Załóżmy, że chcesz się dowiedzieć, czy konkretny samochód, powiedzmy nowy Burpo-Blasto, będzie dla ciebie dobrym autem. Można oczywiście dokonać zakupu i sprawdzić na własnej skórze. Można też porozmawiać z wieloma użytkownikami tego typu samochodu albo ludźmi, którzy rozważali jego zakup, ale tego nie uczynili. Mógłbyś przejrzeć ogłoszenia i sprawdzić, czy nie ma wśród nich dużej liczby tanio sprzedawanych B-B. Mógłbyś przeczytać ocenę Burpo-Blasto w jakimś czasopiśmie konsumenckim. Mógłbyś też skorzystać z kilku sposobów na raz. Podobna sytuacja zachodzi w przypadku badań naukowych.

Ostatecznie badania naukowe sprowadzają się do prowadzenia obserwacji i interpretowania ich wyników, co jest tematem części 3 oraz 4 tej książki. Jednak przed przystąpieniem do obserwacji i analizy należy przygotować plan. Trzeba ustalić, co, dlaczego oraz jak będzie obserwowane i analizowane. Właśnie o to chodzi w planie badań.

Mimo że mogą wystąpić różnice w szczegółach w zależności od tego, co jest obiektem badań, w każdym schemacie badań staniesz wobec dwóch głównych zadań. Po pierwsze, musisz określić tak dokładnie, jak to tylko możliwe, czego chcesz się dowiedzieć. Po drugie, musisz określić, w jaki sposób najlepiej to zrobić. Co ciekawe, jeśli sprawnie sobie poradzisz z pierwszą częścią, to prawdopodobnie w tym samym procesie rozstrzygniesz drugą. Jak mawiają matematycy, właściwie sformułowane pytanie zawiera w sobie odpowiedź.

Powiedzmy, że jesteś zainteresowany zbadaniem korupcji w rządzie. Jest to z pewnością właściwy i godny temat badań. Co jednak masz na myśli używając słowa „korupcja"? Dokładnie, jaki typ zachowania to korupcja? Co znaczy „rząd"? Kogo chcesz badać – wszystkich pracowników państwowych? Osoby na wybieralnych stanowiskach publicznych? Urzędników służby cywilnej? Ostatecznie zaś, jaki jest cel badań? Czy chcesz ustalić rozmiary korupcji? Czy może chcesz się dowiedzieć, dlaczego korupcja występuje? To właśnie na tego typu pytania trzeba dać odpowiedź podczas planowania badań.

Rozdział ten stanowi ogólne wprowadzenie do planowania badań, natomiast inne rozdziały w części 2 dokładniej omawiają konkretne aspekty. W praktyce okazuje się, iż wszystkie te aspekty planowania badań są wzajemnie powiązane. Podczas lektury kolejnych rozdziałów z części 2 zależności te będą stawały się jaśniejsze.

Zaczniemy od krótkiego przeglądu trzech głównych celów badań społecznych. Pomoże to nam określić, jaki typ badań podjąć. Potem omówimy jednostki analizy, czyli co lub kogo chcemy badać. Następnie zaś przyjrzymy się różnym sposobom radzenia sobie z czasem w badaniach społecznych oraz ze sposobami badania zmieniającego się w czasie obiektu.

Po omówieniu tych idei powrócimy do tego, jak zaplanować projekt badawczy. Przegląd procesu badawczego ma dwa cele – poza opisem, jak zabrać się do zaplanowania badań, jest to także mapa dalszych rozdziałów książki.

Na końcu przyjrzymy się elementom oferty badawczej. Często bowiem przeprowadzenie badań jest uzależnione od zebrania odpowiednich funduszy na większe badania bądź też od aprobaty prowadzącego zajęcia, tak więc musi ono zostać poprzedzone dokładnym opisem intencji badacza. Przekonasz się, że sporządzenie oferty badawczej daje doskonałą okazję do rozważenia jeszcze przed rozpoczęciem badań wszystkich ich aspektów.

Trzy cele badań

Badania społeczne mogą służyć wielu celom. Trzy najczęstsze i najbardziej użyteczne cele to eksploracja, opis i wyjaśnianie. Mimo że konkretne badanie może mieć, i najczęściej ma, więcej niż jeden cel, to korzystne jest ich oddzielne omówienie, gdyż każdy z nich pociąga za sobą inne skutki dla pozostałych aspektów planowania badań.

Eksploracja

Większość badań społecznych jest prowadzona po to, aby rzucić światło na jakiś temat bądź też oswoić badacza z jakąś tematyką. Takie podejście zdarza się najczęściej, gdy badacz wchodzi w nową dla siebie dziedzinę zainteresowań lub gdy przedmiot badań sam w sobie jest stosunkowo nowy.

Dla przykładu załóżmy, że masowe niezadowolenie z rządu przeradza się w rewoltę płatników podatków. Ludzie zaczynają odmawiać płacenia podatków i organizują się wokół tej kwestii. Mógłbyś chcieć dowiedzieć się czegoś więcej o tym ruchu. Jaki jest jego zasięg? Jakim cieszy się poparciem społecznym? Jak jest zorganizowany? Jacy ludzie są w nim najbardziej aktywni?

Badania eksploracyjne mogłyby ci pomóc w uzyskaniu przynajmniej przybliżonych odpowiedzi na niektóre z tych pytań. Mógłbyś na przykład sprawdzić dane statystyczne w urzędzie skarbowym, zgromadzić literaturę dotyczącą tego ruchu i zapoznać się z nią, uczestniczyć w spotkaniach czy przeprowadzić wywiady z przywódcami.

Badania eksploracyjne są odpowiednie także dla trwalszych zjawisk. Załóżmy, iż jesteś niezadowolony z formalnych, uniwersyteckich wymogów dotyczących prac magisterskich i chcesz przyczynić się do ich zmiany. Możesz zatem zgłębić historię tych wymagań na twojej uczelni oraz spotkać się z władzami uczelni, by dowiedzieć się, dlaczego obowiązują takie standardy. Możesz porozmawiać z paroma studentami, by wysondować, jakie są ich odczucia w tej sprawie. Mimo że to ostatnie działanie nie przyniesie najprawdopodobniej dokładnego obrazu opinii studentów, może ono pokazać, jakie mogłyby być wyniki szerszych badań.

Niekiedy po badaniach eksploracyjnych kontynuuje się badania metodą grup fokusowych, czy inaczej moderowanych dyskusji w małych grupach. Ta technika jest często stosowana w badaniach rynkowych. Przyjrzymy się im dokładniej w rozdziale 10.

Z reguły badania eksploracyjne mają trzy cele: (1) zaspokojenie ciekawości badacza i jego pragnienia lepszego zrozumienia przedmiotu, (2) zbadanie możliwości podjęcia szerszych badań, (3) wypracowanie metod, które zostaną użyte w dalszych badaniach.

Na przykład jakiś czas temu zdałem sobie sprawę ze wzrastającej popularności pewnego rodzaju „mediów". Osoba, która określa się jako medium, wchodzi w stan transu i zaczyna przemawiać głosem, który, jak twierdzi, pochodzi z zewnątrz. Niektóre z głosów twierdzą, że pochodzą ze świata duchów osób umarłych; inne, iż są z obcych planet; a jeszcze inne utrzymują, że istnieją w innych wymiarach rzeczywistości trudnych do opisania za pomocą zwykłego języka ludzi.

Głosy, określane czasami mianem „bytów", używały niekiedy metafory radia czy telewizji, aby przybliżyć fenomen, który reprezentują. Pewnego razu w trakcie wywiadu jeden z bytów powiedział mi: „Gdy oglądasz wiadomości, to nie wierzysz przecież, że spiker jest naprawdę wewnątrz telewizora. Tak samo jest ze mną, używam ciała medium w taki sam sposób, w jaki spiker używa telewizora".

Problem ten zainteresował mnie z wielu punktów widzenia, wśród których jednym z ważniejszych był problem metodologiczny. Jak bowiem naukowo badać coś, co tak zdecydowanie zrywa z tym, co uważamy za oczywiste, na przykład przestrzenią, czasem, przyczynowością oraz jednostkowością?

Jako że nie miałem ścisłej teorii ani konkretnych oczekiwań, postanowiłem jedynie dowiedzieć się czegoś więcej. Używając niektórych z technik badań terenowych, które są opisane w dalszej części książki, rozpocząłem zbieranie informacji i tworzenie kategorii, tak aby gromadzone przeze mnie obserwacje miały jakikolwiek sens. Czytałem książki oraz artykuły dotyczące tego zjawiska, rozmawiałem z ludźmi, którzy brali udział w tego typu sesjach. Potem sam brałem udział w seansach, obserwując zarówno ludzi, którzy uczestniczyli w nich wraz ze mną, jak i samo medium. Następnie przeprowadziłem wiele wywiadów bezpośrednich z osobami, które były mediami.

W większości wywiadów, poza typowymi pytaniami biograficznymi, pytałem badanych o to, jak to się stało, że zaczęli występować jako medium, jak to było i dlaczego zdecydowali się

nadal to robić. Potem medium wchodziło w stan transu, a ja kontynuowałem wywiad z bytem. Pytałem na przykład: „Kim jesteś?", „Skąd pochodzisz?", „Dlaczego to robisz?", „Skąd mogę wiedzieć, czy jesteś prawdziwy, czy udajesz?". Mimo że przystępując do wywiadów miałem przygotowane z góry kilka pytań, rozmowy prowadziłem w kierunku, który wydawał się właściwy w świetle otrzymywanych odpowiedzi.

Ten przykład badań eksploracyjnych pokazuje, w jaki sposób często zaczynają się badania społeczne. Badacze wychodzący od teorii dedukcyjnych przed przystąpieniem do badań mają przygotowane kluczowe zmienne, natomiast jednym z pierwszych z moich zadań było zidentyfikowanie zmiennych, które mogły odpowiadać za określenie typu ludzi, jacy z dużym prawdopodobieństwem mogą uczestniczyć w seansach. Na przykład w odniesieniu do medium zwracałem uwagę na płeć, wiek, wykształcenie, wyznanie, pochodzenie z określonego regionu, wcześniejszy udział w doświadczeniach metafizycznych. Wybrałem większość z tych zmiennych, gdyż na ogół mają one wpływ na zachowanie.

Notowałem także różnice w okolicznościach towarzyszących takim sesjom. Niektóre media mówiły, że muszą wejść w głęboki trans, inne, iż wystarczający jest lekki trans, a jeszcze inne pozostawały całkowicie świadome. Większość z nich siedziała w czasie seansu, ale niektóre stały lub chodziły. Niektórym do transu wystarczyły zupełnie zwykłe warunki, ale inne potrzebowały specjalnych metafizycznych rekwizytów, takich jak przyciemnione światła, kadzidełka i mantry. Jednak z większości z tych różnic zdałem sobie sprawę dopiero w trakcie pierwszych obserwacji.

W przypadku bytów byłem zainteresowany sklasyfikowaniem ich ze względu na podawane przez nie miejsce pochodzenia. W trakcie wywiadów wypracowałem zestaw pytań dotyczących specyficznych aspektów „rzeczywistości", jako próbę klasyfikacji pojawiających się odpowiedzi. Podobnie uczyniłem w kwestii przewidywania przyszłości.

W miarę postępu badań moja analiza poszczególnych zagadnień skupiała się coraz bardziej wokół pewnych kwestii, ponieważ określiłem zmienne, które wydawały się godne uwagi, na przykład płeć, wykształcenie, wyznanie. Warto zauważyć, że badania rozpoczynałem ze względnie nie zapisaną kartą.

Badania eksploracyjne mają dużą wartość w naukach społecznych. Są kluczowe, gdy badacz wkracza na nowy grunt, i zawsze pomagają się rozeznać w temacie badań. Badania eksploracyjne są także źródłem dla teorii ugruntowanej, co pokazano w rozdziale 2.

Główną wadą badań eksploracyjnych jest to, że zazwyczaj nie dają one satysfakcjonujących odpowiedzi na pytania badawcze, choć mogą wskazywać na takie odpowiedzi i sugerować, jakimi metodami można je osiągnąć. Powód, dla którego badania eksploracyjne rzadko mają charakter ostateczny, wiąże się z reprezentatywnością. Badani ludzie mogą bowiem nie być typowi dla większej populacji, którą faktycznie jesteśmy zainteresowani. Gdy zrozumiesz kwestię reprezentatywności, będziesz w stanie stwierdzić, czy badanie, które właśnie przeprowadziłeś, dało rzeczywiście odpowiedź na pytanie badawcze, czy tylko wskazało drogę do odpowiedzi. (Reprezentatywność jest szerzej omówiona w rozdziale 7).

Opis

Głównym celem wielu badań społecznych jest opis jakiejś sytuacji lub wydarzenia. Badacz obserwuje, a potem opisuje to, co zaobserwował. Ponieważ obserwacja naukowa jest staranna i przemyślana, opis naukowy jest z reguły trafniejszy i dokładniejszy niż zwykły opis.

Spis ludności USA jest doskonałym przykładem opisowych badań społecznych. Jego celem jest trafny i dokładny opis obszernego zestawu różnych cech ludności USA, a także ludności zamieszkującej mniejsze jednostki terytorialne, takie jak poszczególne stany czy hrabstwa. Inne przykłady badań opisowych to profile demograficzne populacji z uwzględnieniem wieku i płci, wykonywane przez demografów, statystyki przestępczości w różnych miastach, badania marketingowe opisujące ludzi, którzy używają lub mogliby używać jakiegoś produktu. Badacz, który prowadzi staranny zapis chronologiczny zdarzeń zachodzących w miejscu pikiety związków zawodowych, kie-

ruje się celami opisowymi, a w każdym razie realizuje takie cele. Inny badacz, który zlicza i sprawozdaje, ile razy poszczególni parlamentarzyści głosowali za związkami zawodowymi lub przeciw nim, także wypełnia funkcje opisowe.

Opis był pierwotnym celem wielu badań jakościowych. Na przykład etnografia antropologiczna może dążyć do szczegółowego opisu kultury jakiejś społeczności przedpiśmiennej. Badania takie zarazem rzadko kiedy ograniczają się do czysto opisowych celów. Badaczy interesuje na ogół dalsze wyjaśnienie, *dlaczego* istnieją zaobserwowane prawidłowości i jakie są ich konsekwencje.

Wyjaśnianie

Trzecim z głównych celów naukowych badań społecznych jest wyjaśnianie. Badania opisowe odpowiadają na pytania, co, kiedy, gdzie oraz jak, wyjaśniające natomiast – dlaczego. Gdy zatem William Sanders (1994) rozpoczął opis różnych przypadków stosowania przemocy przez gangi, chciał także zrekonstruować proces, który prowadził do przemocy wśród gangów wywodzących się z różnych grup etnicznych.

Odnotowanie częstości chodzenia do kościoła jest zajęciem opisowym, ale odnotowanie, dlaczego jedni ludzie chodzą do kościoła, a inni nie, jest już wyjaśnianiem. Publikowanie raportów o częstości popełniania przestępstw w różnych miastach jest przypadkiem opisu, ale określenie zmiennych, które pomagają zrozumieć, dlaczego w jednych miastach wskaźnik przestępczości jest wyższy niż w innych, jest już wyjaśnianiem. Badacz, który zamierza zbadać, dlaczego demonstracja antyaborcyjna zakończyła się brutalnym starciem z policją, zamiast po prostu opisać, co się wydarzyło, ma za cel wyjaśnianie.

Podobnie jest z sondażem na temat postaw wobec legalizacji marihuany. Może on służyć celom opisowym, po prostu odnotowując postawy różnych segmentów populacji. Ale badacz może także spróbować odkryć czynniki kształtujące postawy ludzi wobec tej kwestii, to znaczy zadać pytanie *dlaczego* („Dlaczego ludzie mają określone postawy wobec marihuany?"). Wyjaśniająca analiza danych Generalnego Sondażu Społecznego (GSS General Social Survey) z 1993 roku wykazuje, że 28% mężczyzn oraz 15% kobiet stwierdziło, iż marihuana powinna zostać zalegalizowana. Choć ogromna większość zarówno mężczyzn, jak i kobiet była przeciwna legalizacji, mężczyźni byli prawie dwa razy częściej niż kobiety skłonni do zalegalizowania marihuany, co oznacza, że płeć okazuje się jedną ze zmiennych wpływających na postawy wobec marihuany.

Nie jest zaskakujące, że także poglądy polityczne są skorelowane z postawami wobec legalizacji marihuany. Wśród liberałów 37% opowiedziało się za legalizacją marihuany, w porównaniu z 15% wśród umiarkowanych i 15% wśród konserwatystów. W kategoriach sympatii partyjnych, legalizację poparło 39% demokratów, 18% niezależnych i 15% republikanów.

Mimo że wyodrębnienie trzech celów badań jest użyteczne, warto podkreślić, iż większość badań zawiera elementy wszystkich trzech celów. Załóżmy, że masz ocenić skuteczność nowej formy psychoterapii. Twoje badania będą zawierać elementy eksploracji, gdyż będziesz rozpoznawał zmienne mogące mieć znaczenie oraz zarysujesz oddziaływanie terapii. Będziesz chciał opisać takie zjawiska, jak częstość powrotu do zdrowia. Bez wątpienia spróbujesz też wyjaśnić, dlaczego ta terapia skuteczniej działa na niektóre typy ludzi niż na inne, lub też dlaczego terapia ta dla niektórych typów ludzi (lub problemów) jest skuteczniejsza niż inne formy leczenia.

Zobaczysz, jak w praktyce realizuje się te cele, gdy przejdziemy do omówienia innych aspektów planowania badań. A teraz zastanówmy się, kogo lub co chcesz odkrywać, opisywać i wyjaśniać.

Jednostki analizy

W naukowych badaniach społecznych nie istnieją prawie żadne bariery dotyczące **jednostek analizy**, czyli tego, kto lub co ma być poddawane badaniu. Zagadnienie to odnosi się do wszystkich

form badań społecznych, choć konsekwencje są najbardziej wyraźne w przypadku badań ilościowych o charakterze nomotetycznym.

Koncepcja jednostek analizy może się na początku wydawać dość ryzykowna, gdyż badania – szczególnie nomotetyczne – operują zazwyczaj na dużych zbiorach ludzi bądź rzeczy, lub też na ich agregatach. Ważne jest rozróżnienie jednostek analizy oraz zbiorowości, na temat których będziemy formułować uogólnienia. Na przykład badacz może zajmować się grupą ludzi, taką jak demokraci, studenci, Afroamerykanki poniżej 30 roku życia, albo jakąkolwiek inną. Jednak jeśli jest zainteresowany eksploracją, opisem lub wyjaśnianiem tego, jak członkowie tych różnych grup ludzkich zachowują się jako *jednostki*, to elementem analizy jest jednostka, a nie grupa. Dzieje się tak, mimo że później badacz przechodzi do uogólnień na temat zbiorów jednostek, tak jak w przypadku stwierdzenia, że więcej demokratów niż republikanów popiera legalizację marihuany. Pomyśl o tym w ten sposób – określona postawa w sprawie marihuany jest cechą, którą można przypisać tylko jednostce, nie zaś grupie. Nie ma czegoś takiego jak „umysł" zbiorowy, który może się skłaniać ku określonej postawie. Zatem nawet gdy dokonujemy uogólnień na temat demokratów, stwierdzamy coś o własnościach, które mają oni jako jednostki.

W przeciwieństwie do tego możemy czasem badać grupy pojmowane jako jednostkowe byty czy „aktorzy", którzy mają cechy *jako grupa*. Na przykład możemy zechcieć porównać cechy różnych typów gangów ulicznych. W tym przypadku naszymi jednostkami analizy będą gangi (a nie ich członkowie) i możemy przystąpić do formułowania uogólnień dotyczących różnych typów gangów.

Badacze społeczni najczęściej wybierają jako jednostki analizy pojedynczych ludzi. Mogą oni odnotować cechy tych ludzi – płeć, wiek, miejsce urodzenia, postawy itd. Następnie mogą łączyć opisy, aby przedstawić złożony obraz grupy, którą jednostki reprezentują; wszystko jedno czy jest to gang uliczny, czy też całe społeczeństwo.

Na przykład możesz odnotować wiek i płeć każdego studenta zapisanego na wykład z nauk politycznych. Potem zaś możesz scharakteryzować opisowo tę grupę jako zbiorowość o średniej wieku wynoszącej 18,6 roku, składającą się w 53% z mężczyzn i 47% z kobiet. Mimo że ostateczny opis będzie dotyczył całej grupy studentów, bazuje on na cechach członków grupy jako jednostek.

Takie samo rozróżnienie jednostek analizy i zbiorowości ma miejsce w badaniach wyjaśniających. Załóżmy, że chciałbyś dowiedzieć się, czy studenci, którzy mają dobre wzorce uczenia się, otrzymali lepsze oceny z nauk politycznych niż studenci, którzy mają złe wzorce. Zmienną *wzorce uczenia się* zoperacjonalizowałbyś w kategoriach liczby godzin spędzanych tygodniowo nad nauką. Możesz potem połączyć razem studentów o dobrych wzorcach uczenia się i studentów o złych wzorcach i sprawdzić, która grupa dostała lepsze oceny z egzaminu. Celem tych badań byłoby wyjaśnienie, dlaczego pewne grupy studentów wypadają na egzaminie lepiej niż inne. Jednostkami analizy cały czas pozostają pojedynczy studenci.

Jednostki analizy w danych badaniach są zazwyczaj także jednostkami obserwacji. Aby zatem zgłębić uwarunkowania sukcesu na egzaminie z nauk politycznych, będziemy obserwować pojedynczych studentów. Jednak czasem „obserwujemy" nasze jednostki analizy pośrednio. Na przykład, załóżmy, że chcemy stwierdzić, czy różnice zdań w kwestii kary śmierci wpływają na częstość rozwodów. W tym przypadku, aby rozróżnić małżeństwa, które w tej sprawie mają zgodne oraz niezgodne poglądy, „obserwowalibyśmy" pojedynczych mężów i żony, pytając ich o ich postawy wobec kary śmierci. Naszymi jednostkami obserwacji są w tym przypadku pojedynczy ludzie: mężowie i żony, ale jednostkami analizy (czyli tym co chcemy badać) są małżeństwa.

Jednostki analizy są to zatem te przedmioty, które badamy, aby stworzyć ich syntetyczny opis oraz wyjaśnić różnice między nimi. W większości projektów badawczych będzie dla ciebie zapewne jasne, co jest jednostką analizy. Gdyby tak jednak nie było, konieczne jest jej określenie, gdyż w przeciwnym razie nie będziesz mógł zdecydować, kogo lub czego mają dotyczyć obserwacje.

Niektóre badania starają się opisać lub też wyjaśnić więcej niż jedną jednostkę analizy. W takich przypadkach badacz musi z góry przewidzieć, jakiego rodzaju wnioski chce wyciągnąć w odniesieniu do których jednostek analizy. Na przykład możemy chcieć się dowiedzieć, jaki typ studentów (jednostek) odnosi największe sukcesy w swojej karierze zawodowej. Możemy też chcieć sprawdzić, z jakiego typu uczelni (organizacji) wywodzą się absolwenci, którzy odnoszą największe sukcesy.

Aby nadać temu omówieniu bardziej konkretny charakter, przyjrzyjmy się kilku typowym jednostkom analizy w naukach społecznych.

Jednostki indywidualne

Jak widzieliśmy, indywidualne jednostki ludzkie są w badaniach społecznych najbardziej chyba typowymi jednostkami analizy. Badacze zazwyczaj opisują i wyjaśniają grupy społeczne i interakcje poprzez agregowanie i analizowanie opisów jednostek.

W badaniach społecznych każdy typ jednostki może stać się jednostką analizy i zazwyczaj badania zajmują się określonymi typami lub klasami jednostek. Ta uwaga jest o wiele bardziej istotna, niż może się wydawać na pierwszy rzut oka. Ponieważ nauki społeczne usiłują wyjaśniać ludzkie zachowania ogółem, można by sądzić, że wyniki badań naukowych mają największą wartość, gdy odnoszą się do wszystkich rodzajów ludzi. Jednak w praktyce przedstawiciele nauk społecznych rzadko zajmują się wszystkimi rodzajami ludzi. W najlepszym przypadku ich badania odnoszą się do ludzi zamieszkujących tylko jeden kraj; niektóre badania porównawcze przekraczają granice państw. Jednak często badania mają dość ograniczony zasięg.

Przykłady klas jednostek, które mogą być wybrane do badań, obejmują studentów, homoseksualistów i lesbijki, pracowników przemysłu samochodowego, wyborców amerykańskich, wiernych, samotnych rodziców. Zauważ, że każdy z tych terminów określa pewną populację pojedynczych osób. Badania opisowe, w których jednostkami analizy są pojedynczy ludzie, zazwyczaj stawiają sobie za cel opis populacji składającej się z tych ludzi, natomiast celem badań wyjaśniających jest odkrycie dynamiki społecznej wewnątrz tej populacji.

Jako elementy analizy jednostki ludzkie mogą być opisane w kategoriach ich przynależności do grup społecznych. Zatem jednostka może być opisana jako należąca do bogatej bądź biednej rodziny, albo jako mająca matkę z wyższym wykształceniem lub nie. W trakcie badań możemy ustalić czy ludzie, których matka miała wyższe wykształcenie, będą częściej studiować niż ci, których matka takiego wykształcenia nie miała. Możemy też dowiedzieć się, czy absolwenci szkół średnich z bogatych rodzin będą częściej studiować niż ci z biednych. W każdym przypadku jednostką analizy – tą „rzeczą", której cechy pragniemy opisać lub wyjaśnić – jest indywidualna jednostka. Później zaś agregujemy indywidualne jednostki i formułujemy uogólnienia na temat populacji, do której należą.

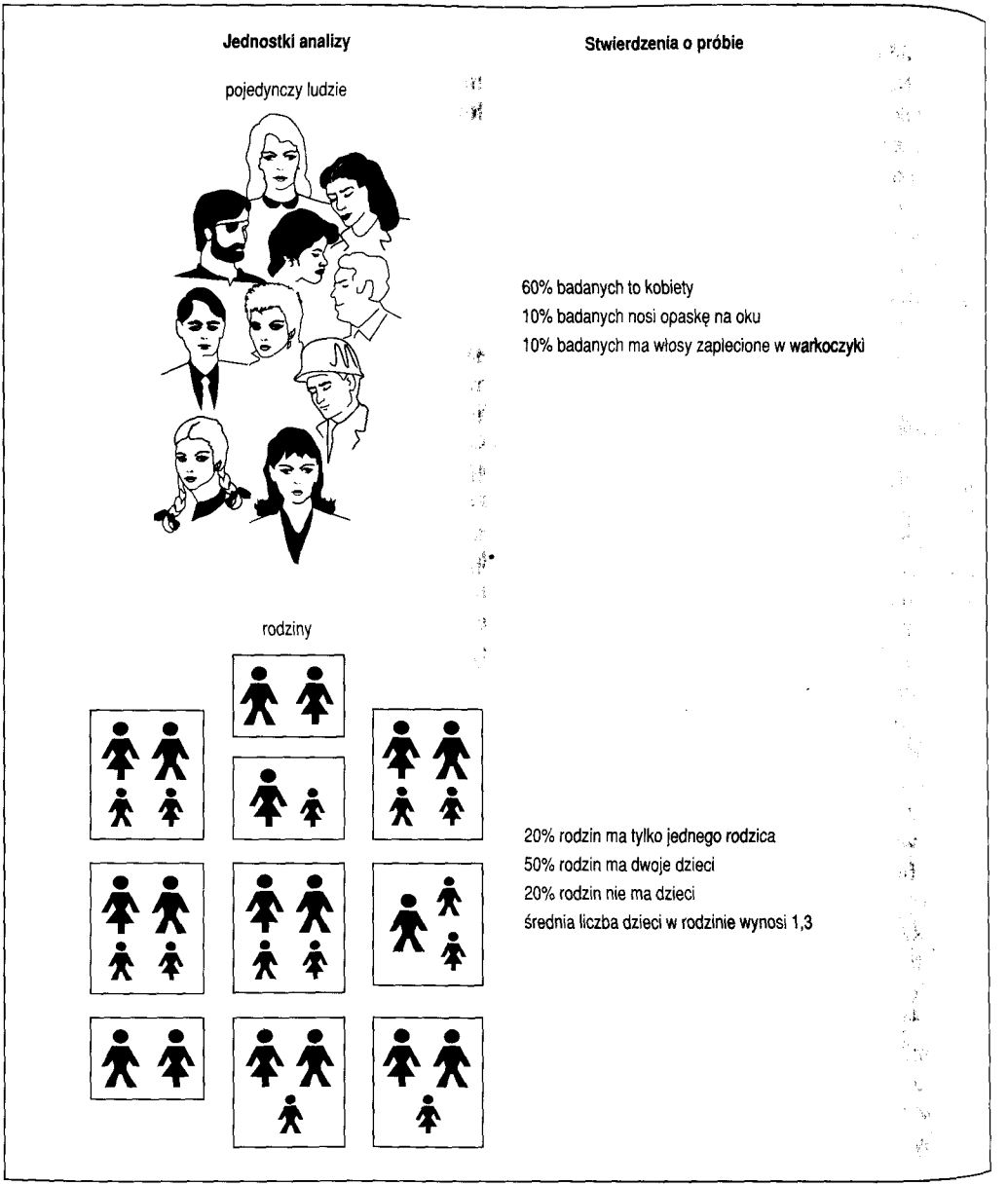
Grupy

W badaniach społecznych jednostkami analizy mogą być także grupy społeczne. Możemy być bowiem zainteresowani cechami przynależącymi do grupy, traktowanej jako pojedynczy byt. Gdybyś miał badać przestępców, analizując zachowania członków gangu, jednostką analizy byłaby pojedyncza jednostka (przestępca). Ale gdybyś badał wszystkie gangi w mieście, by na przykład zgłębić różnice między dużymi a małymi gangami, gangami z centrum i z przedmieść itd., to bardziej interesowałyby cię same gangi niż ich członkowie indywidualnie. W tym przypadku jednostką analizy byłby gang – grupa społeczna.

Oto inny przykład. Załóżmy, że jesteś zainteresowany dostępem do komputerów w różnych segmentach społeczeństwa. Mógłbyś opisywać rodziny ze względu na ich całkowity roczny dochód oraz fakt posiadania lub nieposiadania komputera. Później mógłbyś zagregować rodziny i opisać średni dochód oraz procent posiadających komputer. Byłbyś wtedy w stanie stwierdzić, czy bardziej prawdopo-

Rysunek 4.1

Przykłady jednostek analizy

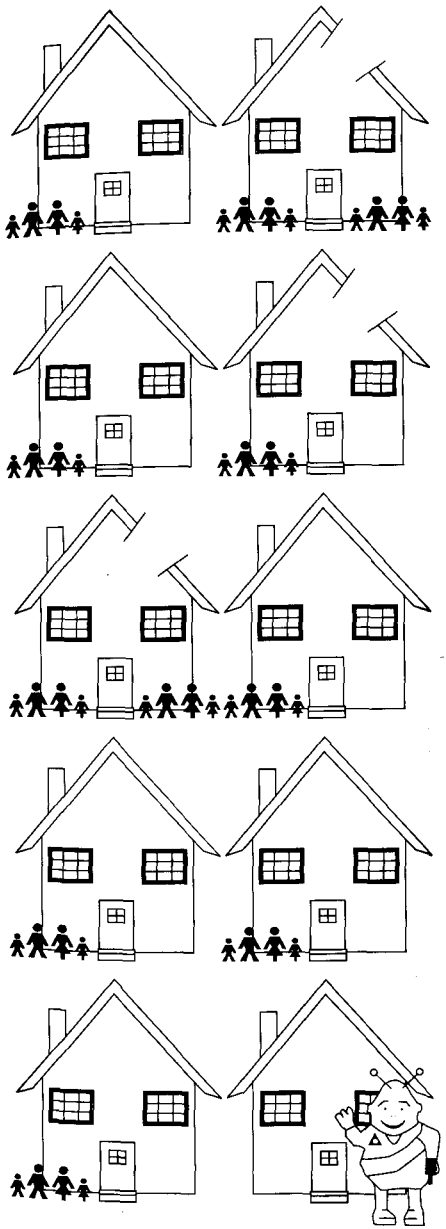


dobne jest posiadanie komputera wśród rodzin o wysokich dochodach, czy wśród rodzin o niskich dochodach. W tym przypadku jednostką analizy byłaby rodzina.

Tak jak w przypadku innych jednostek analizy, możemy wnioskować o cechach grup społecznych na podstawie cech ich członków. Dzięki temu możemy opisać daną rodzinę pod względem wieku, rasy, wykształcenia głowy rodziny. W badaniach opisowych możemy stwierdzić, jaki procent wszystkich rodzin stanowią te, w których głowa rodziny ma wyższe wykształce-

Rysunek 4.1
Przykłady jednostek analizy (ciąg dalszy)

Jednostki analizy
gospodarstwa domowe



Stwierdzenia o próbie

20% gospodarstw domowych jest zamieszkanych przez więcej niż jedną rodzinę

30% gospodarstw domowych ma dziurę w dachu

10% gospodarstw domowych jest zamieszkanych przez przybyszów z obcej planety

Zauważ, że 33% rodzin mieszka w wielorodzinnych *gospodarstwach domowych*, ale w tym *przypadku* jednostką analizy jest rodzina.

nie. W badaniach wyjaśniających możemy stwierdzić, czy takie rodziny mają przeciętnie więcej, czy mniej dzieci niż rodziny, w których głowa rodziny nie ma wyższego wykształcenia. W każdym z tych przykładów jednostką analizy jest rodzina. Jeśli jednak spytalibyśmy, czy ludzie z wyższym wykształceniem mają więcej czy mniej dzieci niż ludzie gorzej wykształceni, to jednostkami analizy byłyby pojedyncze osoby.

Innymi jednostkami analizy na poziomie grup mogą być grupy koleżeńskie, pary małżeńskie, bloki spisowe wyodrębnione do spisu powszechnego, miasta czy regiony geograficzne. Tak jak w przypadku pojedynczych ludzi, każdy z tych terminów oznacza jakąś populację. Na przykład termin gangi uliczne może oznaczać populację, w której skład wchodzą wszystkie gangi uliczne, powiedzmy, z terenu danego miasta. Można zatem opisać tę populację poprzez uogólnianie wyników badań dla pojedynczych gangów. Mógłbyś na przykład opisać geograficzne rozmieszczenie gangów w mieście. W badaniach wyjaśniających mógłbyś ustalić, czy duże gangi częściej niż małe angażują się w wojny między gangami. Doszedłbyś więc do konkluzji dotyczących populacji gangów w mieście, używając pojedynczych grup jako jednostek analizy.

Organizacje

Jednostkami analizy w naukowych badaniach społecznych mogą być także formalne organizacje społeczne. Na przykład badacz może badać przedsiębiorstwa, sugerując populację wszystkich przedsiębiorstw. Pojedyncze przedsiębiorstwa mogą być scharakteryzowane pod kątem liczby pracowników, rocznego zysku netto, aktywów brutto, liczby podpisanych kontraktów, odsetka pracowników reprezentujących mniejszości rasowe czy etniczne itd. Możemy określić, czy w dużych firmach jest procentowo mniej czy więcej pracowników reprezentujących mniejszości niż w małych. Inne formalne organizacje społeczne, które mogą być jednostkami analizy, to na przykład kongregacje kościelne, szkoły wyższe, dywizje wojskowe, wydziały uniwersyteckie, supermarkety. Rysunek 4.1 jest ilustracją graficzną kilku różnych jednostek analizy wraz z przykładami twierdzeń, jakie można o nich formułować.

Wytwory społeczne

Inną jednostką analizy są **wytwory społeczne** czyli inaczej wszelkie wytwory istot społecznych lub ich zachowań. Jedną z klas tych wytworów są konkretne przedmioty, takie jak książki, wiersze, obrazy, samochody, budynki, piosenki, ceramika, dowcipy, studenckie usprawiedliwienia nieobecności na egzaminach czy odkrycia naukowe. W badaniach Robina Wagnera-Pacifici (1995) dotyczących ataku policji filadelfijskiej na działaczy MOVE, opisanego w rozdziale 3, jednostką analizy były oficjalne oświadczenia władz.

Tak jak pojedynczy ludzie czy grupy społeczne odnoszą się do populacji, tak każdy przedmiot społeczny odnosi się do zbioru przedmiotów tej samej klasy: wszystkich książek, wszystkich powieści, wszystkich biografii, wszystkich podręczników do socjologii, wszystkich książek kucharskich. W badaniach, w których jednostkami analizy są książki, pojedynczą książkę można scharakteryzować przez jej wagę, długość, cenę, treść, liczbę rysunków, liczbę sprzedanych egzemplarzy, opis jej autora. Później można analizować populację wszystkich książek albo książek jakiegoś rodzaju w celu opisu lub wyjaśniania – na przykład jakie rodzaje książek sprzedają się najlepiej i dlaczego.

Podobnie, badacz społeczny mógłby analizować, czy najwięcej zainteresowania losem klasy robotniczej wykazuje malarstwo rosyjskie, chińskie czy amerykańskie. Jednostkami analizy byłyby obrazy, określone między innymi przez narodowość swego autora. Można też analizować artykuły redakcyjne w danej gazecie, dotyczące lokalnego uniwersytetu, aby opisywać albo może wyjaśniać zmiany w czasie stanowiska gazety względem uczelni. W tym przypadku jednostką analizy byłby pojedynczy artykuł redakcyjny.

Interakcje społeczne tworzą kolejną klasę wytworów społecznych, które mogą być badane. Możemy na przykład opisać małżeństwa

jako wyznaniowo lub rasowo mieszane (lub nie), świeckie lub religijne w swej ceremonii, kończące się rozwodem (lub nie). Można też opisać jednego lub oboje partnerów (na przykład „będąca/y wcześniej zamężna/żonaty", „kibic Oakland Raider", „poszukiwana/y przez FBI"). Kiedy badacz odnotowuje, że małżeństwa wyznaniowo mieszane są częściej niż małżeństwa między osobami tej samej wiary zawierane przed władzami świeckimi, wtedy jednostkami analizy stają się małżeństwa, a nie zaangażowane w nie jednostki.

Inne interakcje społeczne, które mogą stać się jednostkami analizy, to na przykład przyjaźnie, sprawy sądowe, wypadki drogowe, rozwody, bójki na pięści, wodowania statków, porwania samolotów, zamieszki rasowe, egzaminy końcowe, demonstracje studenckie, przesłuchania kongresowe. Przesłuchania w Kongresie mogą być na przykład charakteryzowane przez to, czy były przeprowadzane w czasie kampanii wyborczej, czy przewodniczący komisji kandydowali na jakieś urzędy, czy otrzymali darowizny na kampanię od zainteresowanych stron itd. Zauważ, że jeśli nawet scharakteryzujemy i porównamy przesłuchania ze względu na cechy przewodniczących komisji, to jednostką analizy pozostaną przesłuchania, a nie pojedynczy przewodniczący komisji.

Przegląd jednostek analizy

Przykłady z tego rozdziału wskazywałyby, iż w badaniach społecznych istnieje niemal nieskończona ilość różnych możliwych jednostek analizy. Mimo że zazwyczaj przedmiotem badań są pojedynczy ludzie, w przypadku wielu pytań badawczych właściwszą drogą poszukiwania odpowiedzi może być badanie innych jednostek analizy. Właściwie badacze społeczni mogą badać wszystko, co odnosi się do życia społecznego.

Ponadto wymienione w tym rozdziale typy jednostek analizy nie wyczerpują wszystkich możliwości. Na przykład Morris Rosenberg (1968, s. 234–248) mówi o indywidualnych, grupowych, organizacyjnych, instytucjonalnych, przestrzennych, kulturowych i społecznych jednostkach analizy. John oraz Lyn Loflandowie (1995, s. 103–113) wskazują na działania zwyczajowe i epizody, spotkania, role, związki, grupy, organizacje, miejsca zamieszkania, światy społeczne, style życia oraz subkultury – jako na stosowne jednostki badania. Ważną rzeczą jest uchwycenie logiki jednostek analizy. Potem twoje możliwości prowadzenia owocnych badań są ograniczone tylko wyobraźnią.

Kategoryzacja możliwych jednostek analizy może doprowadzić do tego, że pojęcie to wydaje się niepotrzebnie skomplikowane. To, jak nazwiesz jednostkę analizy – grupą, organizacją formalną czy wytworem społecznym – jest nieistotne. Kluczem jest jasność dotycząca tego, co nią jest. Gdy przystępujesz do projektu badawczego, musisz zdecydować się, czy badasz małżeństwa, czy partnerów z osobna, przestępstwa czy przestępców, przedsiębiorstwa czy ich dyrektorów. Jeśli tego nie zrobisz, to ryzykujesz tym, że wyciągniesz błędne wnioski, gdyż twoje twierdzenia dotyczące danej jednostki analizy będą się opierać faktycznie na badaniu innej. Zobaczymy tego przykład, gdy zajmiemy się błędem ekologicznym.

Błędne wnioskowanie o jednostkach analizy: błąd ekologiczny oraz redukcjonizm

W tym momencie naszych rozważań jest rzeczą właściwą wprowadzenie dwóch typów błędnego wnioskowania, których istnienia powinieneś być świadom: błędu ekologicznego oraz redukcjonizmu. Każdy to potencjalna pułapka związana z jednostkami analizy; do każdego z nich może dojść w czasie przeprowadzania badań i wyciągania wniosków z ich wyników.

Błąd ekologiczny

W tym kontekście „ekologiczny" odnosi się do grup, zbiorów lub systemów: czegoś większego niż jednostki. **Błąd ekologiczny rozumowania** jest to założenie, że to, czego dowiadujemy się o jednostce zbiorowej („ekologicznej"), mówi nam coś także o pojedynczych elementach, z których się ona składa. Rozważmy hipotetyczną ilustrację tego błędu.

Załóżmy, że jesteśmy zainteresowani strukturą poparcia dla pewnej kandydatki w ostatnich wyborach samorządowych. Załóżmy, że znamy wyniki z każdego obwodu głosowania, tak że potrafimy stwierdzić, który popierał ją najbardziej, a który najmniej. Załóżmy także, że mamy dane ze spisu powszechnego charakteryzujące te obwody. Analiza takich danych mogłaby wykazać, iż obwody zamieszkiwane przez względnie młodych wyborców oddały większy procent głosów na tę kandydatkę niż obwody zamieszkiwane przez starszych wyborców. Kusiłoby nas może, aby wnioskować na podstawie tych wyników, iż młodsi wyborcy chętniej niż starsi głosują na kobiety. Innymi słowy, że wiek wpływa na popieranie kobiet w polityce. Dochodząc do takiego wniosku, ryzykujemy popełnienie błędu ekologicznego, dlatego że to starsi wiekiem wyborcy zamieszkujący „młode" okręgi wyborcze mogli głosować na naszą kandydatkę. Nasz problem polega na tym, że badaliśmy *obwody głosowania* jako jednostki analizy, ale chcemy wyciągać wnioski dotyczące *wyborców*.

Podobny problem powstałby, gdybyśmy dowiedzieli się, że wskaźnik przestępczości jest wyższy w miastach zamieszkiwanych przez dużą liczbę Afroamerykanów niż w miastach, gdzie jest ich niewielu. Nie wiedzielibyśmy, czy przestępstwa są rzeczywiście popełniane przez Afroamerykanów. Podobnie, gdybyśmy odkryli, że wskaźnik samobójstw w krajach protestanckich jest wyższy niż w katolickich, cały czas nie wiedzielibyśmy na pewno, czy więcej protestantów niż katolików popełniło samobójstwo.

Pomimo tych niebezpieczeństw przedstawiciele nauk społecznych często nie mają wyboru i muszą odnieść się do konkretnego pytania badawczego przez analizę ekologiczną. Prawdopodobnie najwłaściwszy rodzaj danych jest po prostu nieosiągalny. Na przykład wyniki wyborcze z każdego obwodu głosowania oraz charakterystyka każdego obwodu mogą być łatwe do zdobycia, ale możemy nie mieć środków na przeprowadzenie powyborczego sondażu wśród indywidualnych wyborców. W takich przypadkach możemy formułować wstępne wnioski, rozpoznając i odnotowując ryzyko popełnienia błędu ekologicznego.

Powinieneś uważać, aby nie popełniać błędu ekologicznego, ale nie powinieneś z kolei popadać w coś, co można by nazwać błędem indywidualistycznym. Niektórzy ludzie, którzy po raz pierwszy mają do czynienia z badaniami społecznymi, mają problem z pogodzeniem ogólnych prawidłowości, dotyczących postaw i zachowań, z jednostkowymi wyjątkami. Tak jak pokazaliśmy w rozdziale 2, ważności uogólnień i stwierdzeń probabilistycznych nie podważa istnienie wyjątków. Na przykład fakt, iż ktoś zna bogatego demokratę, nie podważa reguły, że większość bogatych ludzi głosuje na republikanów – jest to ogólna prawidłowość. Podobnie istnienie osoby, która doszła do fortuny bez jakiegokolwiek formalnego wykształcenia, nie podważa ogólnej prawidłowości, że wyższe wykształcenie jest związane z wyższymi zarobkami.

Błąd ekologiczny dotyczy czegoś innego – pomieszania jednostek analizy w taki sposób, że wyciągamy wnioski dotyczące jednostek tylko na podstawie obserwacji grup. Mimo że prawidłowości dotyczące zmiennych na poziomie grup mogą być rzeczywiste, niebezpieczeństwo polega na wnioskowaniu na podstawie zaobserwowanych cech grup o cechach wchodzących w ich skład jednostek, gdy tymczasem nie obserwowaliśmy naprawdę jednostek.

Redukcjonizm

Drugim typem możliwego błędnego rozumowania związanego z jednostkami analizy jest redukcjonizm. Redukcjonizm oznacza postrzeganie i wyjaśnianie złożonych zjawisk w kategoriach jednego wąskiego pojęcia lub zbioru pojęć. A zatem „redukujemy" to, co w rzeczywistości jest złożone, do prostego wyjaśnienia.

Na przykład naukowcy z różnych dziedzin mają skłonność do poszukiwania pewnych typów odpowiedzi, a ignorowania innych. Socjologowie zazwyczaj biorą pod uwagę tylko zmienne socjologiczne (takie jak wartości, normy oraz role), ekonomiści tylko zmienne ekonomiczne (takie jak podaż, popyt i wartość krańcowa), a psychologowie tylko zmienne psychologiczne (takie jak typy osobowości, przeżycia traumatyczne). Wyjaśnianie wszyst-

kich czy większości ludzkich zachowań w kategoriach czynników ekonomicznych nazywamy redukcjonizmem ekonomicznym; wyjaśnianie wszystkich czy większości ludzkich zachowań w kategoriach psychologicznych to redukcjonizm psychologiczny itd. Zauważ, jak ta kwestia odnosi się do omówienia paradygmatów teoretycznych z rozdziału 2.

Innym przykładem może być pytanie o przyczyny rewolucji amerykańskiej. Czy było to podzielane przez ludzi przywiązanie do wartości, jaką jest wolność osobista? Czy może trudne położenie ekonomiczne kolonii w stosunku do Anglii? Może zaś megalomania ojców założycieli? Jeśli będziemy badać tylko *jedną* przyczynę, to narażamy się na redukcjonizm.

Redukcjonizm każdego rodzaju sugeruje zwykle, że pewne jednostki analizy lub zmienne są istotniejsze niż inne. Jeśli mielibyśmy uznać wspólnie podzielane wartości za przyczyny rewolucji amerykańskiej, to jednostką analizy byliby pojedynczy osadnicy. Jednak ekonomista mógłby jednostką analizy uczynić 13 kolonii i badać gospodarczą organizację i warunki w każdej z nich. Psycholog zaś mógłby wybrać jako jednostki analizy pojedynczych przywódców w celu badania ich osobowości.

Podobnie jak błąd ekologiczny, redukcjonizm może mieć miejsce, gdy użyjemy niewłaściwych jednostek analizy. Jednak wybór odpowiedniej dla danego pytania badawczego jednostki analizy nie zawsze jest sprawą jasną. Badacze społeczni, zwłaszcza z różnych dziedzin, często nad tym dyskutują.

Wymiar czasowy

Do tej pory w tym rozdziale zajmowaliśmy się planem badań jako procesem decyzyjnym dotyczącym tego, jakie aspekty kogo lub czego chcemy obserwować oraz w jakim celu. Teraz musimy rozważyć zestaw różnych możliwości związanych z wymiarem czasowym, które nakładają się na każde z naszych poprzednich rozważań. Możemy zdecydować się na zebranie obserwacji mniej więcej w tym samym czasie lub rozłożenie ich w dłuższym okresie.

Czas pełni różne funkcje w planowaniu oraz przeprowadzaniu badań, niezależnie od czasu potrzebnego na ich realizację. W rozdziale 3 stwierdziliśmy, że kolejność wydarzeń i sytuacji w czasie odgrywa kluczową rolę przy ustalaniu związków przyczynowych (wrócimy do tego w części 4). Czas wpływa także na możliwość uogólniania wyników badań. Czy opisy oraz wyjaśnienia, wynikające z konkretnych badań, trafnie przedstawiają sytuację sprzed 10 lat, sytuację za 10 lat, czy wyłącznie obecną sytuację? Generalnie badacze mają dwie możliwości potraktowania kwestii czasu w planie swoich badań: badania przekrojowe oraz badania dynamiczne.

Badania przekrojowe

Badania przekrojowe składają się z obserwacji próby czy przekroju jakiejś populacji lub zjawiska, zebranych w jednym punkcie w czasie. Badania eksploracyjne i opisowe są często badaniami przekrojowymi. Na przykład pojedynczy Amerykański Spis Powszechny jest badaniem nastawionym na opis populacji Stanów Zjednoczonych w danym czasie.

Wiele badań wyjaśniających to także badania przekrojowe. Badacz przeprowadzający duży ogólnokrajowy sondaż, w celu zbadania źródeł uprzedzeń rasowych i religijnych, będzie najprawdopodobniej zajmował się jednym okresem. To tak, jakby robił „zdjęcie" źródeł uprzedzeń w konkretnym momencie historii.

Przekrojowe badania wyjaśniające mają pewną naturalną słabość. Mimo że wnioski z tych badań bazują na obserwacjach przeprowadzonych tylko w jednym czasie, zazwyczaj ich celem jest wyjaśnienie procesów przyczynowych, zachodzących w czasie. Problem ten ma coś wspólnego z ustalaniem prędkości ruchomego obiektu na podstawie jednego kadru, który zatrzymuje nagle ruch tego obiektu.

Na przykład Yanjie Bian przeprowadził w Chinach badania sondażowe wśród robotników w Tianjin. Celem ich było badanie stratyfikacji społecznej we współczesnym chińskim mieście. Jednak przy podejmowaniu tych badań w 1988 roku był on świadomy wagi zmian

powodowanych przez serię ogólnokrajowych kampanii, takich jak Wielka Proletariacka Rewolucja Kulturalna, rozpoczętych Rewolucją Chińską z 1949 roku (która wyniosła chińskich komunistów do władzy) i trwających do dziś.

> „Kampanie te zmieniły polityczny klimat oraz wpłynęły na pracę oraz zajęcia ludzi w wolnym czasie. Z ich powodu trudno jest z przekrojowego sondażu (takiego jak prezentowany w tej książce) wyciągać wnioski na temat ogólnych prawidłowości w chińskich miejscach pracy oraz ich wpływu na Chińczyków. Wnioski takie mogą być ograniczone do jednego okresu i powinny być sprawdzane na podstawie danych z innych okresów".
>
> (1994, s. 19)

Problem uogólnień na temat życia społecznego formułowanych na podstawie pojedynczych „zdjęć" stale powraca w tej książce. Jedno z rozwiązań jest zasugerowane w ostatnim zdaniu Biana: o danych „z innych okresów". Badania społeczne często wiążą się z powrotem do pewnych zjawisk i z budowaniem na wynikach wcześniejszych badań.

Badania dynamiczne

W przeciwieństwie do badań przekrojowych **badania dynamiczne** (*longitudinal studies*) są tak zaplanowane, aby umożliwić obserwację tego samego zjawiska przez dłuższy czas. Na przykład badacz może wziąć udział w działaniach ludzi związanych z kultem UFO od ich początku aż do upadku, a także obserwować je. W innych badaniach dynamicznych do analizy zmian zachodzących w czasie wykorzystywane są zapisy lub wytwory. Jeśli na przykład analizujemy trendy w gazetowych artykułach redakcyjnych albo orzeczenia Sądu Najwyższego w czasie, to są to badania dynamiczne, niezależnie od tego czy badacz zbiera i analizuje obserwacje w jednym punkcie w czasie, czy w trakcie faktycznego zachodzenia badanych zdarzeń.

Wiele projektów badań terenowych, angażujących bezpośrednią obserwację, oraz niekiedy wywiady pogłębione to w naturalny sposób badania dynamiczne. Na przykład gdy Ramona Asher oraz Gary Fine (1991) badali doświadczenia życiowe żon alkoholików, byli oni w stanie przeanalizować ewolucję burzliwych stosunków małżeńskich tych kobiet w czasie, niekiedy włącznie z reakcjami badanych na samo badanie.

W klasycznym opracowaniu *When Prophecy Fails* (1956) Leon Festinger, Henry Reicker oraz Stanley Schachter byli szczególnie zainteresowani tym, co stało się z kultem latających spodków, po tym jak przepowiednie kultu, dotyczące spotkania z obcymi, nie spełniły się. Czy wyznawcy kultu będą rozwiązywać swe grupy, czy może staną się jeszcze bardziej oddani swym przekonaniom? Aby dać odpowiedź na to pytanie, potrzebne były badania dynamiczne. (Okazało się, że grupy zdwoiły wysiłki, aby pozyskać nowych członków).

Badania dynamiczne, takie jak ogólnokrajowe spisy, mogą być trudne w kwantyfikacji. Niemniej jednak są one często najlepszą metodą badania zmian w czasie. Istnieją trzy specjalne typy badań dynamicznych, o których trzeba wiedzieć: badania trendów, badania kohort demograficznych oraz badania panelowe.

Badania trendów

Badania trendów to typ badań dynamicznych, które koncentrują się na zmianach w czasie wewnątrz jakiejś populacji. Prosty przykład to porównanie danych Amerykańskiego Spisu Powszechnego z dziesiątek lat pokazujące przesunięcia w składzie ludności Stanów Zjednoczonych. Podobny użytek z danych archiwalnych uczynili Michael Carpini i Scott Keeter (1991), którzy chcieli wiedzieć, czy obecnie obywatele USA są lepiej, czy gorzej poinformowani o polityce niż poprzednie pokolenie. Aby to sprawdzić, porównali oni wyniki kilku sondaży Gallupa przeprowadzanych w latach 40. i 50. z sondażem z 1989 roku, w którym zadano kilka takich samych pytań, badających wiedzę polityczną.

Ogólnie analiza wskazywała, że obywatele obecnie są nieco lepiej poinformowani niż poprzednie pokolenia. W 1989 roku 74% badanych wiedziało, kto jest wiceprezydentem Stanów Zjednoczonych, w porównaniu do 67% w 1952 roku. Znacznie większy procent badanych potrafił wytłumaczyć prezydenckie weta

oraz motywy ich odrzucenia przez Kongres, niż miało to miejsce w 1947 roku. Jednak większy procent respondentów wiedział, kto jest jego politycznym przedstawicielem w 1947 roku (38%) niż w 1989 roku (29%).

Dokładniejsza analiza wskazuje jednak, że nieco większa wiedza polityczna jest skutkiem tego, iż badani w 1989 roku byli trochę lepiej wykształceni niż respondenci we wcześniejszych badaniach. Gdy uwzględniono wykształcenie, okazało się, że w poszczególnych grupach wykształcenia poziom wiedzy politycznej faktycznie spadł.

Badania kohort demograficznych

W **badaniach kohort** badacz analizuje zmiany zachodzące w poszczególnych subpopulacjach (kohortach demograficznych). Najczęściej kohorta to grupa wiekowa, na przykład ludzie urodzeni w latach 50. Może to być jednak także inna grupa zdefiniowana w kategoriach czasu, jak na przykład ludzie urodzeni w czasie wojny wietnamskiej, ludzie zawierający małżeństwa w 1994 roku itd. Przykładem badań kohort może być seria ogólnokrajowych sondaży, przeprowadzanych co około 20 lat w celu zbadania postaw kohorty ludzi, urodzonych w czasie II wojny światowej wobec zaangażowania się Stanów Zjednoczonych w politykę międzynarodową. W roku 1960 można było objąć badaniami próbę osób w wieku 15–20 lat, w roku 1980 inną próbę osób w wieku 35–40 lat, a w roku 2000 próbę osób w wieku 55–60 lat. Chociaż konkretny zbiór ludzi badanych w każdym z tych sondaży byłby różny, każda próba reprezentowałaby kohortę demograficzną ludzi urodzonych między 1940 a 1945 rokiem.

James Davis (1992) sięgnął po badania kohort, usiłując wyjaśnić zmiany poglądów politycznych w latach 70. oraz 80. w Stanach Zjednoczonych. Ogólnie potwierdził on istnienie trendu liberalnego w sprawach takich jak rasa, płeć, religia, polityka, przestępczość, wolność słowa. Ale czy ten trend oznaczał, że ludzie generalnie stają się nieco bardziej liberalni, czy tylko odzwierciedlał wypieranie starszych, bardziej konserwatywnych pokoleń przez młodsze, bardziej liberalne?

Aby odpowiedzieć na to pytanie, Davis przeanalizował ogólnokrajowe sondaże z czterech różnych okresów, odległych jeden od drugiego o pięć lat. W każdym z tych sondaży pogrupował respondentów w grupy wiekowe o rozpiętości także pięciu lat. Taka strategia pozwoliła mu porównać różne grupy wiekowe w dowolnym okresie, a także śledzić rozwój poglądów politycznych każdej grupy z osobna.

Jednym z analizowanych przez Davisa pytań było pytanie, czy osoba, która przyznaje się do bycia komunistą, powinna mieć prawo do przemawiania w społeczności respondenta. Młodsi respondenci konsekwentnie w każdym z okresów byli bardziej niż starsi skłonni do zezwolenia komuniście na przemawianie. Na przykład w pierwszym z badań 72% osób w wieku 20–40 lat zajęło takie liberalne stanowisko, w porównaniu do 27% w grupie wiekowej 80 lat i więcej. Wyniki badań Davisa dla najmłodszej kohorty są przedstawione w tabeli 4.1.

Taki wzór zachowań – lekki konserwatywny zwrot w latach 70. oraz następujące po nim liberalne odbicie w latach 80. charakteryzuje kilka kohort, które Davis analizował (1992, s. 269).

Badania panelowe

Mimo że są podobne do badań trendów oraz badań kohort demograficznych, **badania panelowe** dotyczą za każdym razem tego samego zbioru ludzi. Możemy na przykład przeprowadzać wywiady z tą samą próbą wyborców każdego

Tabela 4.1

Wiek a liberalizm polityczny

Daty przeprowadzania sondaży	1972–1974	1977–1980	1982–1984	1987–1989
Wiek kohorty	20–24	25–29	30–34	35–39
Procent osób, które pozwoliłyby przemawiać komuniście	72%	68%	73%	73%

miesiąca kampanii wyborczej, pytając, na kogo mają zamiar głosować. Dzięki takim badaniom będziemy w stanie nie tylko analizować ogólne trendy w preferencjach wyborczych dotyczących różnych kandydatów, ale także pokazać precyzyjne prawidłowości dotyczące stałości i zmienności zamiarów wyborczych. Na przykład badania trendów wskazujące, że kandydaci A oraz B mają 1 września oraz 1 października poparcie dokładnie połowy głosujących, mogą oznaczać, że nie nastąpiła żadna zmiana zamiarów wyborczych, że wszyscy wyborcy jednego kandydata przeszli na stronę drugiego, albo też coś pośredniego. Badania panelowe pozwoliłyby uniknąć takiego zamieszania, pokazując, oprócz wielu innych faktów, którzy wyborcy przenieśli swe poparcie z A na B, a którzy z B na A.

Joseph Veroff, Shirley Hatchett i Elizabeth Douvan (1992) byli zainteresowani procesem dostosowywania się małżonków wśród młodych małżeństw, a szczególnie różnicami między parami białymi a afroamerykańskimi. Uczestników do badań rekrutowano spośród par, które zgłosiły się po poświadczenie aktu małżeństwa w hrabstwie Wayne w stanie Michigan w okresie od kwietnia do czerwca 1996 roku.

Biorąc pod uwagę wpływ, jaki badania mogłyby mieć na procesy dostosowawcze w parach, które zgodziły się brać w nich udział, podzielono próbę losowo na pół – na grupę eksperymentalną oraz grupę kontrolną (pojęcia te omówimy dokładniej w rozdziale 8). Z parami z pierwszej grupy prowadzono intensywne wywiady przez cztery lata, z drugą grupą zaś kontaktowano się tylko krótko raz na rok.

Dzięki badaniu tych samych par w czasie badacze mogli śledzić konkretne pojawiające się problemy oraz sposób, w jaki sobie z nimi radzono. Produktem ubocznym badań było stwierdzenie, że małżonkowie z intensywnie badanych par nieco lepiej dopasowywali się do siebie. Badacze sądzili, że wywiady mogły zmusić pary do przedyskutowania spraw, które bez tego mogłyby być pomijane.

Porównanie trzech typów badań dynamicznych

Aby uwypuklić różnice między badaniami trendów, kohort oraz panelowymi, porównajmy te trzy schematy badań w odniesieniu do tej samej zmiennej: przynależności religijnej. Badania trendów mogłyby wykazywać przesunięcia w przynależności religijnej w USA w różnych okresach. Czyni tak regularnie Instytut Gallupa. Badania kohort mogłyby badać przesunięcia w przynależności religijnej wśród „pokolenia Wielkiego Kryzysu", konkretnie na przykład tych, którzy w 1932 roku mieli od 20 do 30 lat. Moglibyśmy badać próbę ludzi, którzy mieli od 30 do 40 lat w 1942 roku, nową próbę ludzi, którzy mieli od 40 do 50 lat w 1952 roku itd. Badania panelowe rozpoczęlibyśmy na próbie z całej populacji albo z jakiegoś specyficznego podzbioru. Później badalibyśmy te konkretne jednostki w czasie. Zauważ, że tylko badania panelowe dałyby pełny obraz przesunięć między różnymi religiami oraz przesunięć do i z kategorii „żadna". Badania trendów oraz kohort ujawniałyby tylko zmiany netto.

Generalnie badania dynamiczne mają oczywistą przewagę nad przekrojowymi. Dostarczają informacji opisujących proces w czasie. Niestety przewaga ta, szczególnie w przypadku szeroko zakrojonych badań sondażowych, jest osiągana bardzo dużymi nakładami zarówno czasu, jak i pieniędzy. Być może trzeba gromadzić obserwacje w czasie, gdy zachodzą określone wydarzenia, a metoda obserwacji może wymagać zaangażowania wielu ankieterów.

Badania panelowe, które dostarczają najpełniejszych danych o zmianach w czasie, stają przed szczególnym problemem: wyczerpania panelu. Niektórzy respondenci biorący udział w pierwszej fali badań mogą nie chcieć lub nie być w stanie uczestniczyć w późniejszych falach. Istnieje niebezpieczeństwo, że ci, którzy wypadli z badań, mogą nie być typowi, przez co wyniki badań zostaną zafałszowane. Gdy zatem Carol s. Aneshensel wraz z kolegami porównywali w badaniu panelowym zachowania seksualne młodzieży latynoskiej i nielatynoskiej za pomocą badań panelowych, szukali oni różnic między młodzieżą latynoską urodzoną w Meksyku a młodzieżą latynoską urodzoną w Stanach Zjednoczonych w grupie osób, które wypadły z panelu – i znaleźli takie różnice. Różnice tego typu należało brać pod uwagę, aby uniknąć mylnych wniosków o różnicach

Wymiar czasowy i starzenie się

Joseph J. Leon

Wydział Nauk Behawioralnych,
California State Politechnic University, Pomona

Jednym ze sposobów zrozumienia tego, jaki rodzaj wymiaru czasowego został zastosowany w danym badaniu, jest wyobrażenie sobie pewnej ilości różnych projektów dotyczących starzenia się w społeczeństwie amerykańskim. Jeśli badalibyśmy próbę ludzi w 1990 roku i porównywali różne grupy wiekowe, ten schemat badań nazywałby się *badaniami przekrojowymi*. Jeśli dobralibyśmy inną próbę ludzi w 2000 roku i porównali nowe dane z danymi z 1990 roku, ten schemat badań nazywałby się *badaniami trendów*.

Załóżmy, że chcielibyśmy badać tylko te osoby, które w 2000 roku miały od 51 do 60 lat, i porównać je z próbą osób, które w 1990 roku miały od 41 do 50 lat (kohorta demograficzna 41–50-latków). Ten schemat badań nazywałby się *badaniami kohort*. Można by też dokonać porównania dla kohort 51–60-latków oraz 61–70-latków. Gdybyśmy teraz chcieli przeprowadzić *badania panelowe* dotyczące starzenia się w Stanach Zjednoczonych, musielibyśmy dobrać w 1990 roku próbę osób i w 2000 roku powtórzyć badania na tej samej próbie. Pamiętaj, że w badaniach w 2000 roku weźmie udział mniej ludzi, gdyż wszyscy 41–50-latkowie z roku 1990 będą mieć w 2000 roku od 51 do 60 lat, w 2000 roku zaś nie będzie żadnych 41–50-latków. Ponadto niektóre z osób dobranych do próby w 1990 roku mogą już nie żyć w roku 2000.

Badania przekrojowe

1990
↑ 41–50
↓ 51–60
↕ 61–70
↕ 71–80

Badania kohort

1990		2000
41–50	⟷	41–50
51–60	⟷	51–60
61–70	⟷	61–70
71–80	⟷	71–80

Badania trendów

1990		2000
41–50	⟷	41–50
51–60	⟷	51–60
61–70	⟷	61–70
71–80	⟷	71–80

Badania panelowe

1990		2000
41–50*	⟷	41–50*
51–60*	⟷	51–60*
61–70*	⟷	61–70*
71–80*	⟷	71–80*
	⟷	+81*

⟷ Oznacza porównanie
* Oznacza te same jednostki

między Latynosami a nie-Latynosami (Aneshensel i in. 1989). Inne różnice między tymi trzema typami badań dynamicznych przedstawia ramka zatytułowana „Wymiar czasowy i starzenie się".

Badania zbliżone do dynamicznych

Badania dynamiczne nie zawsze są wykonalnym czy praktycznym sposobem badania procesów zachodzących w czasie. Na szczęście badacze często są w stanie wyciągać przybliżone wnioski na temat takich procesów nawet wtedy, gdy dostępne są jedynie dane przekrojowe. Oto kilka sposobów.

Czasami dane przekrojowe zakładają zachodzenie pewnych procesów w czasie na mocy prostej logiki. Na przykład w badaniach dotyczących zażywania narkotyków przez studentów, przeprowadzonych na Uniwersytecie Hawajskim, o których wspominam w rozdziale 2, pytano studentów, czy kiedykolwiek próbowali każdego z kilku zabronionych prawem narkotyków. Badanie wykazało, że niektórzy studenci próbowali zarówno marihuany, jak i LSD, nie-

którzy tylko jednego, a niektórzy nie próbowali żadnego. Dane były zbierane w jednym punkcie w czasie. Zapewne niektórzy studenci później także eksperymentowali z narkotykami. Wydaje się zatem, że niemożliwe jest stwierdzenie na podstawie takich badań, czy studenci są bardziej skłonni do tego, aby spróbować najpierw LSD, czy marihuany. Jednak dokładniejsza analiza danych wykazała, że mimo iż niektórzy studenci przyznawali się, że próbowali marihuany, ale nie próbowali LSD, nie było ani jednego, który próbowałby LSD, ale nie palił marihuany.

Z tego odkrycia badacze wywnioskowali – zgodnie ze zdrowym rozsądkiem – że używanie marihuany jest wcześniejsze niż zażywanie LSD. Gdyby proces eksperymentowania z narkotykami przebiegał w odwrotnej kolejności, badania w dowolnym momencie powinny wykazać istnienie studentów, którzy próbowali LSD, ale nie próbowali marihuany, oraz nieistnienie studentów, którzy próbowali tylko marihuany.

Badacze mogą również korzystać z logicznego wnioskowania za każdym razem, gdy jasny jest chronologiczny porządek zmiennych. Jeśli w przekrojowych badaniach studentów wyższych uczelni odkrylibyśmy, że studenci, którzy uczęszczali do prywatnych szkół średnich, osiągają lepsze wyniki niż studenci, którzy chodzili do szkół państwowych – wnioskowalibyśmy, że to typ ukończonej szkoły średniej wpływa na oceny na studiach, a nie na odwrót. Moglibyśmy zatem wyciągać wnioski na temat procesów zachodzących w czasie, mimo że nasze obserwacje zostałyby poczynione tylko w jednym punkcie czasowym.

Bardzo często różnice wiekowe zauważone w badaniach przekrojowych są podstawą wniosków dotyczących procesów zachodzących w czasie. Załóżmy, że jesteśmy zainteresowani prawidłowościami, dotyczącymi pogarszania się zdrowia w trakcie typowego cyklu życia. Można zająć się tym tematem poprzez analizę wyników rocznych badań medycznych w dużych szpitalach. Można by pogrupować karty zdrowia w zależności od wieku badanego oraz ocenić każdą z grup wiekowych ze względu na takie aspekty stanu zdrowia, jak: wzrok, słuch, ciśnienie krwi itd. Przez porównywanie każdego z tych aspektów dla różnych grup wiekowych otrzymalibyśmy coś zbliżonego do historii zdrowotnej pojedynczego człowieka. Dzięki temu można by wnioskować, że u przeciętnego człowieka problemy ze wzrokiem pojawiają się wcześniej niż problemy ze słuchem. Trzeba być ostrożnym w tego typu stwierdzeniach, gdyż różnice te mogą być odzwierciedleniem jakichś szerszych trendów. Może lepsze metody badania słuchu, wprowadzone w szkołach, wpłynęły tylko na młodszych ludzi objętych tymi badaniami.

Proszenie ludzi o opowiadanie zdarzeń z ich przeszłości jest inną powszechnie stosowaną metodą zbliżoną do obserwacji w czasie. Badacze używają tej metody, gdy pytają ludzi o miejsce urodzenia, o to, kiedy skończyli szkołę średnią albo na kogo głosowali w 1996 roku. Badacze jakościowi często przeprowadzają wywiady pogłębione na temat „historii życia". Na przykład C. Lynn Carr (1988) użyła tej techniki w badaniu dotyczącym bycia „chłopczycą". Respondentki w wieku od 25 do 40 lat proszone były o zrekonstruowanie pewnych aspektów życia, począwszy od dzieciństwa, w tym doświadczeń związanych z poczuwaniem się do bycia „chłopczycą".

Niebezpieczeństwa związane z tą techniką są oczywiste. Czasami ludzi zawodzi pamięć, a czasami po prostu kłamią. Gdy po wyborach pyta się respondentów, na kogo głosowali, okazuje się nieuchronnie, że więcej ludzi głosowało na zwycięzcę, niż to miało miejsce w rzeczywistości. Jako część serii wywiadów pogłębionych, takie wyniki mogłyby zostać zweryfikowane na podstawie innych szczegółów, jednak wyniki oparte na pojedynczym pytaniu w sondażu należy traktować ostrożnie.

Nasze omówienie tego, w jaki sposób czas może pojawiać się w badaniach społecznych, wskazuje, że w twoich własnych projektach badawczych powinieneś zadać sobie kilka pytań. Planując jakiekolwiek badania, upewnij się, że przeanalizowałeś zarówno jawne, jak i ukryte założenia dotyczące czasu. Czy jesteś zainteresowany opisem lub wyjaśnianiem zmian jakiegoś procesu, zachodzącego w czasie, czy też po prostu tym, co istnieje w danej chwili? Jeśli chcesz opisać pewien proces zachodzący w czasie, to czy będziesz w stanie dokonywać obserwacji w różnych jego fazach? Może będziesz

musiał stosować metodę zbliżoną do obserwacji w czasie przez wyciąganie logicznych wniosków z tego, co obserwujesz w chwili obecnej? Jeśli zdecydujesz się na badania dynamiczne, to która metoda będzie najbardziej odpowiednia dla twoich celów badawczych?

Przykłady strategii badawczych

Jak wynika z powyższego omówienia, w badaniach społecznych jest wiele ścieżek, którymi można podążać. Poniższe krótkie cytaty ilustrują to jeszcze dobitniej. Czytając je, zwróć uwagę zarówno na treść badań, jak i na metody zastosowane do badania wybranego tematu. Czy są to badania eksploracyjne, opisowe, czy wyjaśniające (a może są kombinacją tych rodzajów)? Jakie jest źródło danych w każdym przypadku? Czy możesz zidentyfikować jednostki analizy? Czy wymiar czasowy ma znaczenie? Jeśli tak, to jak będzie potraktowany?

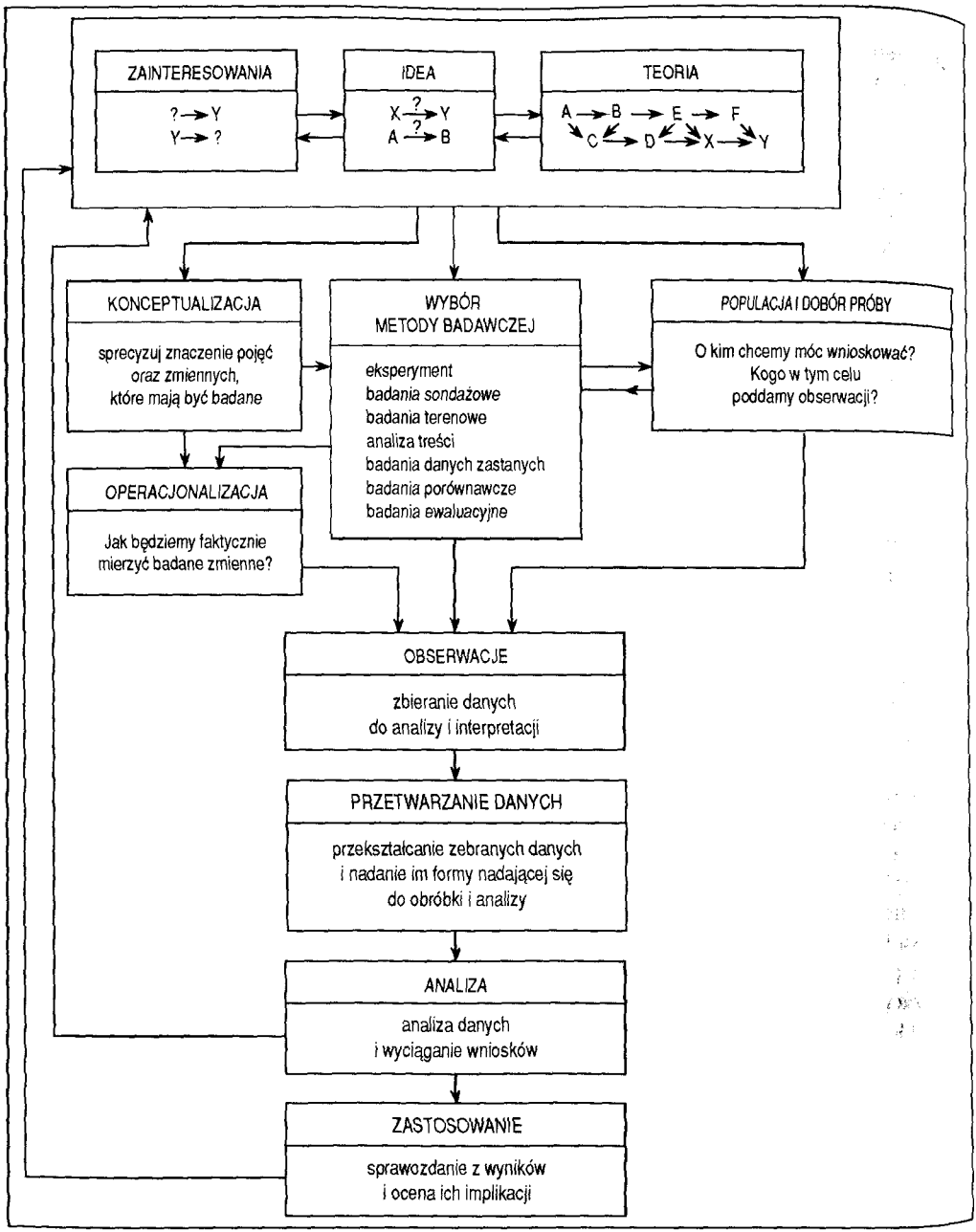
- „W niniejszym studium przypadku dotyczącym cichej mobilizacji, prowadzonej przez Południowokalifornijskie Centrum Kryzysowe dla Ofiar Gwałtu wykorzystano dane archiwalne, obserwacje oraz wywiady w celu zbadania, jak działała pewna organizacja feministyczna pracująca na rzecz zmian w policji, szkołach, sądownictwie oraz niektórych organizacjach stanowych i ogólnokrajowych między 1974 a 1994 rokiem" (Schmitt, Martin 1999, s. 364).
- „Niniejsze badania mają na celu zgłębienie, za pomocą wywiadów narracyjnych dotyczących historii życia, procesu dojrzewania 14 kobiet, które określały się jako chłopczyce" (Carr 1998, s. 528).
- „Wykorzystując wywiady z działaczami z byłej Estońskiej Socjalistycznej Republiki Radzieckiej, określamy okoliczności, dzięki którym istnieją subkultury dostosowawcze i opozycyjne, oraz sposób, w jaki przekształcają się one w ruchy społeczne" (Johnston, Snow 1998, s. 473).
- „Raport prezentuje wyniki badań etnograficznych dotyczących pewnych organizacji na rzecz walki z AIDS, działającej w małym mieście. Bazuje ono na połączeniu obserwacji uczestniczącej, wywiadów z uczestnikami oraz analizy wewnętrznych dokumentów organizacji" (Kilburn 1998, s. 89).
- „Na podstawie wywiadów przeprowadzonych w czasie badań terenowych w Palestynie w 1992, 1993 i 1994 roku oraz dokumentów archiwalnych i historycznych, twierdzę, że dyskurs palestyńskich feministek uformował się pod wpływem społeczno-politycznego kontekstu, w jakim palestyńskie kobiety działały i z jakim miały kontakt" (Abdulhadi 1998, s. 649).
- „Artykuł, na podstawie 51 wywiadów pogłębionych, przedstawia doświadczenia kobiet, dotyczące karmienia piersią w miejscach publicznych" (Stearns 1999, s. 308).
- „Na podstawie wywiadów oraz obserwacji z badań terenowych pokazuję, w jaki sposób system czasowego zatrudnienia w miejscu pracy nastawionym na współpracę zarówno wykorzystuje nowych pracowników, jak i kształtuje ich aspiracje i cele zawodowe" (Smith 1998, s. 411).
- „Zbierałem dane [dotyczące retoryki białych separatystów] z różnych środków przekazu, zajmujących się dyskursem publicznym, w tym z periodyków, książek, broszur, transkrypcji programów radiowych i telewizyjnych oraz doniesień w gazetach i czasopismach" (Berbrier 1998, s. 435).
- „W poniższej analizie badane będą nierówności dotyczące uzyskania zatrudnienia i emerytury wynikające z rasy oraz płci na podstawie ogólnokrajowej próby osób, które przeszły na emeryturę w latach 1980–1981" (Hogan, Perrucci 1998, s. 528).
- „Na podstawie wywiadów z dealerkami twardych narkotyków artykuł analizuje techniki, jakimi posługiwały się one, aby uniknąć aresztowania" (Jacobs, Miller 1998, s. 550).

Jak zaplanować projekt badawczy

Właśnie poznałeś niektóre możliwości planowania projektów badawczych, jakie mają do wyboru badacze społeczni. Złóżmy teraz

Rysunek 4.2

Proces badawczy



wszystko w całość i spójrzmy na rzeczywisty proces planowania projektu badawczego. Załóżmy, że miałbyś podjąć się badań. Jak zacząć? Co robić dalej?

Plan to początek badań, odzwierciedla on jednak wszystkie etapy, przez które przechodzi się następnie w czasie projektu. Zatem niniejsze omówienie zarówno dostarcza wska-

zówek, jak zacząć projekt badawczy, jak też stanowi przegląd tematów poruszanych w dalszych rozdziałach tej książki.

Rysunek 4.2 przedstawia schemat procesu badawczego w naukach społecznych. Przedstawiam ten diagram z oporami, dlatego że może on sugerować bardziej uporządkowaną kolejność działań w badaniach społecznych, niż zazwyczaj ma to miejsce w rzeczywistości. Mimo to taki wyidealizowany przegląd procesu badawczego stanowi kontekst dla szczegółów, dotyczących konkretnych składników badań społecznych. Zasadniczo jest to inny, bardziej szczegółowy obraz procesu naukowego, przedstawionego w rozdziale 2.

Na górze diagramu są zainteresowania, idea oraz teoria, jako możliwe punkty wyjścia dla badań. Litery (A, B, X, Y itd.) reprezentują zmienne albo pojęcia, takie jak uprzedzenia czy alienacja. Możesz być zatem ogólnie zainteresowany, dlaczego jedni ludzie są bardziej uprzedzeni niż inni. Lub też możesz chcieć poznać niektóre konsekwencje alienacji. Może też być inaczej: możesz rozpocząć badania od jakiejś określonej idei, dotyczącej tego, jak pewne rzeczy się dzieją. Na przykład możesz sądzić, że praca w usługach o niskim prestiżu (np. w barze szybkiej obsługi) jest przyczyną alienacji. Znaki zapytania na diagramie wskazują, że nie jesteś pewien, czy w rzeczywistości rzeczy dzieją się tak, jak podejrzewasz – i właśnie dlatego przeprowadzasz badania. Zauważ, że teoria jest reprezentowana jako zbiór złożonych zależności między wieloma zmiennymi.

Strzałki w obie strony między „zainteresowaniami", „ideą" i „teorią" mają oznaczać, że często występują przejścia od danego punktu wyjścia do innych. Początkowe zainteresowanie czymś może prowadzić do sformułowania jakiejś idei, która może pasować do szerszej teorii, a ta teoria może tworzyć nowe idee i pobudzać nowe zainteresowania.

Każdy z tych trzech powodów lub wszystkie naraz mogą wskazywać na potrzebę przeprowadzenia badań empirycznych. Ich celem może być chęć zgłębienia pewnego zagadnienia leżącego w polu naszych zainteresowań, przetestowanie określonej hipotezy lub sprawdzenie złożonej teorii. Bez względu na cel badacz musi podjąć wiele decyzji, jak wynika z pozostałych elementów diagramu.

Bądźmy bardziej konkretni i weźmy sobie za przykład konkretne badania. Załóżmy, że interesuje cię kwestia aborcji i chcesz się dowiedzieć, dlaczego niektórzy studenci popierają prawo do aborcji, a inni są mu przeciwni. Idąc krok dalej, powiedzmy, iż masz wrażenie, że studenci z wydziałów humanistycznych i społecznych są generalnie bardziej skłonni popierać ideę prawa do aborcji niż studenci z wydziałów przyrodniczych.

Jeśli ujmujemy to w kategoriach różnych możliwości, omówionych wcześniej w tym rozdziale, interesuje cię zapewne zarówno opis, jak i wyjaśnianie. Jaki procent studentów popiera prawo kobiety do aborcji (opis) i co sprawia, że jedni popierają to prawo, a inni się mu sprzeciwiają (wyjaśnianie)? Jednostką analizy są pojedynczy ludzie – studenci. Załóżmy, że zadowoliłbyś się ustaleniem, jak wygląda sytuacja w chwili obecnej. Mógłbyś więc zadecydować, że badania przekrojowe byłyby odpowiednie dla twoich celów. Choć ten typ badań nie dostarcza dowodów na istnienie procesu, możliwe jest przeprowadzenie analiz zbliżonych do dynamicznych przez prześledzenie zmian postaw studentów w czasie.

Zaczynamy

Na początku projektu będziesz miał zapewne cele eksploracyjne. W tym miejscu mógłbyś wybierać spośród kilku możliwych działań, prowadzących do zaspokojenia twojej ciekawości co do postaw studentów wobec prawa do aborcji. Na początek możesz przeczytać coś o tej kwestii. Jeśli masz przeczucie, że te postawy mogą mieć coś wspólnego z kierunkiem studiów, możesz sprawdzić, co w tej sprawie napisali inni badacze. Załącznik A pomoże ci skorzystać z uczelnianej biblioteki. Ponadto zechcesz zapewne porozmawiać z paroma ludźmi, którzy popierają prawo do aborcji, i z paroma, którzy są mu przeciwni. Mógłbyś wziąć udział w zebraniach grup anty- lub proaborcyjnych. Wszystkie te działania pomogłyby ci się przygotować do

podjęcia różnych decyzji dotyczących planu badań, który za chwilę omówimy.

Zanim zaczniesz planować badania, musisz określić ich cel. Jakiego rodzaju badań się podejmiesz – eksploracyjnych, opisowych czy wyjaśniających? Czy planujesz napisać raport z badań na potrzeby pracy magisterskiej lub w celu uzyskania zaliczenia? Czy może twoim celem jest zebranie informacji, które dostarczą ci argumentów za prawem do aborcji lub przeciw niemu? Czy chcesz napisać artykuł do gazety uniwersyteckiej lub pisma naukowego? Przy przeglądaniu już istniejącej literatury badawczej na ten temat powinieneś zwracać uwagę na decyzje innych badaczy, dotyczące wyboru określonego schematu badań. Zawsze zadawaj sobie przy tym pytanie, czy te same decyzje służyłyby realizacji twoich celów.

Zazwyczaj cel podejmowanych przez ciebie badań może zostać wyrażony w formie raportu. Załącznik C, umieszczony w ostatniej części książki, pomoże ci przygotować raport badawczy. Zalecam zaplanowanie takiego raportu jako pierwszy krok w planowaniu badań. Mimo że ostateczna wersja raportu może się różnić od pierwotnego wyobrażenia, ćwiczenie to pomoże ci porównać stosowność różnych schematów badawczych. Podczas takiego projektowania opisz jasno rodzaj twierdzeń, jakie chcesz sformułować w momencie zakończenia badań. Czyniąc to ułatwisz sobie wybór właściwych strategii badawczych. Oto kilka przykładów takich twierdzeń: „Studenci często wspominali o prawie do aborcji w kontekście omawiania spraw społecznych, które ich osobiście dotyczyły". „X procent studentów uniwersytetu stanowego opowiada się za prawem decydowania kobiet w sprawie aborcji". „Inżynierowie są (mniej/bardziej) skłonni niż socjologowie uznawać prawo do aborcji".

Konceptualizacja

Mając już dobrze zdefiniowany cel badań oraz jasny obraz wyników, jakie chcesz uzyskać, możesz przejść do następnego etapu w planowaniu badań – konceptualizacji. Często mówimy bez większego zastanowienia o takich pojęciach z zakresu nauk społecznych, jak uprzedzenia, alienacja, religijność, liberalizm. Aby jednak móc formułować na ich temat sensowne wnioski, musimy wyjaśnić, co przez te pojęcia rozumiemy. Konceptualizacja jest szerzej omówiona w rozdziale 5. Teraz zaś zobaczmy, z czym może się ona wiązać w naszym hipotetycznym przykładzie.

Jeśli masz zamiar badać, jak studenci odnoszą się do prawa do aborcji i dlaczego, pierwszą rzeczą, jaką należy uściślić, jest rozumienie terminu „prawo do aborcji". Jako że poparcie dla aborcji prawdopodobnie będzie różne w zależności od okoliczności, będziesz chciał zwrócić uwagę na warunki, pod jakimi ludzie mogliby zaakceptować aborcję lub nie: na przykład w sytuacji zagrożenia życia kobiety, w przypadku gwałtu lub kazirodztwa czy po prostu w zależności od osobistego wyboru.

Podobnie będziesz musiał sprecyzować znaczenie wszystkich innych pojęć, które planujesz badać. Jeśli chcesz badać zależności między opinią na temat aborcji a kierunkiem studiów, musisz zdecydować, czy brać pod uwagę oficjalny uniwersytecki podział, czy może uwzględnić także plany studentów. A co zrobisz z tymi, którzy nie wybrali jeszcze określonego kierunku?

W przypadku sondaży oraz eksperymentów konieczne jest sprecyzowanie takich pojęć z góry. W mniej ustrukturyzowanych badaniach, takich jak wywiad swobodny, ważną częścią badań może być odkrywanie różnych wymiarów, aspektów lub też niuansów zawartych w pojęciach. W takich przypadkach same badania mogą ujawnić i zarejestrować pewne aspekty życia społecznego, które nie były oczywiste na początku projektu.

Wybór metody badawczej

Jak pokazujemy w części 3, każda metoda badawcza ma swoje mocne i słabe strony, a do badania niektórych pojęć pewne metody nadają się bardziej niż inne. W naszych badaniach dotyczących postaw wobec aborcji najbardziej odpowiednią metodą mógłby być sondaż: przeprowadzenie wywiadów ze studentami lub poproszenie ich o wypełnienie kwestionariusza.

Sondaże szczególnie nadają się do badania opinii publicznej. Nie oznacza to bynajmniej, iż nie można z powodzeniem użyć innych metod prezentowanych w części 3. Można na przykład zastosować analizę treści i badać listy do redakcji, analizując różne wyobrażenia autorów listów na temat aborcji. Badania terenowe otwarłyby drogę do zrozumienia, jak ludzie kontaktują się ze sobą w sprawie aborcji, jak o niej rozmawiają i jak zmieniają swoje stanowisko. Inne metody badawcze wprowadzone w części 3 także mogą zostać zastosowane do badania tego tematu. Zazwyczaj najlepsze są schematy badawcze, które wykorzystują więcej niż jedną metodę i odnoszą korzyści z ich różnych zalet. Jeśli cofniesz się o kilka stron do przykładów rzeczywistych badań, zobaczysz, iż w kilku z nich skorzystano z więcej niż jednej metody badawczej.

Operacjonalizacja

Następnym krokiem po sprecyzowaniu pojęć będących przedmiotem badania oraz wybraniu metody badawczej jest operacjonalizacja, czyli wybór technik pomiaru (jest ona omówiona szerzej w rozdziałach 5 i 6). Jak widzieliśmy w rozdziale 3, w badaniach znaczenie zmiennych jest determinowane przez sposób ich pomiaru. Część zadania polega tutaj na podjęciu decyzji co do sposobu zbierania potrzebnych danych: poprzez bezpośrednią obserwację, analizę oficjalnych dokumentów, kwestionariusz czy za pomocą jakiejś innej techniki.

Jeśli do badania postaw wobec prawa do aborcji zdecydowałeś się na zastosowanie sondażu, to częścią operacjonalizacji jest ustalenie brzmienia pytań kwestionariusza. Możesz na przykład zoperacjonalizować swoją główną zmienną, pytając respondentów, czy uznaliby prawo kobiety do aborcji pod każdym z kolejnych warunków, które określiłeś w procesie konceptualizacji: w przypadku gwałtu lub kazirodztwa, gdy ciąża zagraża życiu kobiety itd. Kwestionariusz zaprojektowałbyś tak, aby pytać respondenta o akceptację lub odrzucenie prawa do aborcji w każdej z tych sytuacji. Na podobnej zasadzie określiłbyś dokładnie sposób, w jaki respondenci podawaliby swój kierunek studiów oraz co mogliby zrobić ci, którzy jeszcze nie wybrali kierunku studiów.

Populacja i dobór próby

Po dopracowaniu pojęć oraz sposobu ich pomiaru czas zdecydować, kogo lub co badać. Populacją w badaniach jest grupa (zazwyczaj ludzi), na temat której będziemy formułować wnioski. Jednak prawie nigdy nie jesteśmy w stanie przebadać wszystkich członków interesującej nas populacji, ani też przeprowadzić wszystkich możliwych obserwacji. W każdym zatem przypadku dobieramy próbę spośród danych, które w idealnych badaniach mogłyby zostać zebrane i przeanalizowane. Dobór próby informacji odbywa się oczywiście w życiu codziennym i często prowadzi do obciążonych obserwacji (przypomnij sobie z rozdziału 1 omówienie „wybiórczej obserwacji"). Badacze społeczni z większym rozmysłem dobierają próbę tego, co będzie podlegać obserwacji.

Rozdział 7 opisuje takie metody doboru próby, które dają adekwatne odzwierciedlenie całej interesującej nas populacji. Zauważ, iż na rysunku 4.2 decyzje dotyczące populacji i doboru próby są związane z decyzjami dotyczącymi metod badawczych. Techniki probabilistycznego doboru próby są odpowiednie dla szeroko zakrojonych badań sondażowych lub dla analizy treści, badacz terenowy natomiast musi niekiedy dobrać tylko takich informatorów, którzy zapewnią uzyskanie zrównoważonego obrazu badanej sytuacji. Eksperymentator zaś mógłby przypisać badanych do grupy eksperymentalnej i kontrolnej w sposób zapewniający porównywalność.

W hipotetycznym badaniu postaw studentów wobec aborcji właściwą populację będą stanowić studenci twojej uczelni. (Jeśli chcesz wnioskować w ogóle o wszystkich studentach, w rozdziale 7 dowiesz się, jak to zrobić). Jednak dobór próby do badań, co zostanie w dalszej części książki pokazane, wymaga od nas jeszcze większej precyzji. Czy włączać do niej tylko studentów studiów dziennych, czy także zaocznych i wieczorowych? Licencjackich czy magisterskich? Czy brać pod uwagę studentów

zagranicznych? Czy uwzględniać studia podyplomowe i doktoranckie? Jest wiele takich pytań i na każde z nich trzeba dać odpowiedź uwzględniającą cel badań. Na przykład jeśli twoim celem jest przewidywanie wyników lokalnego referendum w sprawie aborcji, możesz chcieć ograniczyć populację tylko do osób mających prawo głosu i deklarujących udział w głosowaniu.

Obserwacje

Po decyzji co badać, wśród kogo i jaką metodą, jesteś gotowy do rozpoczęcia obserwacji – zebrania danych empirycznych. W rozdziałach części 3, które opisują różne metody badawcze, znajdziesz odpowiednie techniki obserwacji dla każdej z tych metod.

By przeprowadzić sondaż na temat aborcji, można wydrukować kwestionariusze i wysłać je do próby dobranej spośród studentów. Można też zorganizować zespół ankieterów i przeprowadzić sondaż telefoniczny. Względne wady i zalety tych i innych metod omówione są w rozdziale 9.

Przetwarzanie danych

W zależności od wybranej metody badawczej zgromadziłeś dużą ilość informacji w formie, która prawdopodobnie nie będzie od razu nadawała się do interpretacji. Jeśli spędzisz miesiąc, bezpośrednio obserwując gang uliczny, będziesz miał wystarczającą ilością notatek, aby zapełnić nimi książkę. W badaniach historycznych na temat etnicznej różnorodności w twojej szkole mógłbyś zgromadzić tomy oficjalnych dokumentów, wywiadów z pracownikami szkolnej administracji oraz innymi ludźmi itd. Rozdział 14 opisuje niektóre sposoby przetwarzania i przekształcania danych w naukach społecznych, służące analizie ilościowej lub jakościowej.

W przypadku sondażu „surowe" obserwacje występują na ogół w formie kwestionariuszy z zakreślonymi kratkami, odpowiedziami wpisanymi w pozostawione miejsce itd. Faza przetwarzania danych sondażowych z reguły obejmuje klasyfikację (kodowanie) wpisanych odpowiedzi oraz przeniesienie wszystkich informacji do komputera.

Analiza

Gdy zgromadzone dane mają już odpowiednią formę, można przystąpić do ich interpretacji w celu wyciągania wniosków dotyczących zainteresowań, idei oraz teorii, które zainspirowały badania. Rozdziały od 13 do 17 opisują niektóre z licznych sposobów analizy danych. Zauważ, że na rysunku 4.2 wyniki analiz dają sprzężenie zwrotne w kierunku początkowych zainteresowań, idei lub teorii. Często sprzężenie to stanowi początek kolejnego cyklu badań.

W sondażu dotyczącym postaw studentów wobec prawa do aborcji faza analizy mogła służyć zarówno celom opisowym, jak i wyjaśnianiu. Można by zacząć od obliczenia odsetka studentów, którzy są za każdą z różnych wersji prawa do aborcji lub przeciw niej. W ujęciu całościowym tych kilka wyników procentowych dałoby dobry obraz opinii studentów w tej sprawie.

Wychodząc poza prosty opis można by opisać opinie w ramach pewnych podzbiorów studentów, takich jak studenci różnych kierunków. Jeśli plan twoich badań wymagał także uchwycenia innych informacji o respondentach, możesz również porównywać mężczyzn i kobiety, studentów różnych lat albo innych uwzględnionych przez ciebie kategorii. Opisy podgrup mogą zaprowadzić cię do analizy wyjaśniającej.

Zastosowanie

Ostatnim etapem procesu badawczego jest wykorzystanie przeprowadzonych badań oraz wypływających z nich wniosków. Na początku zechcesz zapewne ogłosić swoje wyniki, tak aby inni dowiedzieli się, co ustaliłeś. Wskazane może być przygotowanie, a nawet publikacja pisemnego raportu. Może zorganizujesz ustną prezentację, podobną do profesjonalnych czy

naukowych referatów, wygłaszanych na zebraniach. Inni studenci także mogą być zainteresowani tym, czego dowiedziałeś się na ich temat.

Może zechcesz wyjść poza proste złożenie raportu z wyników badań i zaproponujesz dyskusję na temat ich praktycznych implikacji. Czy wyniki wskazują na jakieś działania, które mogłyby zostać podjęte, aby osiągnąć pewne cele? Zarówno zwolennicy, jak i przeciwnicy aborcji byliby tym zainteresowani.

Na koniec upewnij się, że nie pominąłeś rozważań na temat wskazań, jakie wynikają z twoich badań dla dalszych badań tego tematu. Jakie błędy powinny zostać poprawione w przyszłych badaniach? Jakimi drogami, które otwierają się w wyniku twoich badań, powinny podążać przyszłe projekty? Ten aspekt procesu badawczego będzie omówiony w rozdziale 19.

Przegląd procesu planowania badań

Jak wynika z tego przeglądu, zaplanowanie badań wymaga podjęcia wielu decyzji dotyczących tematu, populacji, metod badawczych oraz celu badań. Mimo że będziesz chciał wziąć pod uwagę wiele sposobów studiowania zagadnienia – i wykorzystać zarówno swoją wyobraźnię, jak i wiedzę na temat różnych metod badawczych – planowanie badań jest procesem zawężania możliwości i koncentrowania się na celach konkretnych badań.

Jeśli przygotowujesz projekt badawczy na jakieś zajęcia, wiele aspektów schematu badań może być z góry określonych, włącznie z metodą (na przykład eksperyment) czy tematem (jeśli zajęcia dotyczą określonego tematu, na przykład uprzedzeń). W poniższym podsumowaniu zakładam, że masz wolny wybór w zakresie tematu oraz strategii badawczej.

Planując projekt badawczy, warto zacząć od oceny trzech rzeczy: swoich zainteresowań, umiejętności oraz dostępnych środków. Każda z tych ocen określa wiele możliwości badań.

Wyobraźmy sobie, że rozpoczynamy dość konwencjonalny projekt badawczy. Zadaj sobie pytanie, co cię interesuje, co chciałbyś zrozumieć. Na pewno masz kilka pytań dotyczących zachowań społecznych oraz postaw ludzi. Dlaczego niektórzy są politycznymi liberałami, a inni konserwatystami? Dlaczego jedni ludzie są bardziej religijni niż inni? Dlaczego ludzie przystępują do organizacji paramilitarnych? Czy wyższe uczelnie i uniwersytety nadal dyskryminują pracowników, będących przedstawicielami jakichś mniejszości? Dlaczego kobiety pozostają w związkach, w których są wykorzystywane? Spędź trochę czasu nad pytaniami, które cię interesują i dotyczą.

Gdy już masz kilka pytań, na które chciałbyś znać odpowiedzi, zastanów się, jaki rodzaj informacji jest potrzebny do udzielenia tych odpowiedzi. Jakie jednostki analizy zapewniłyby zdobycie najbardziej odpowiednich informacji: studenci, młode kobiety, sąsiedzi, miejscowości czy przedsiębiorstwa? W twoich myślach pytanie to będzie prawdopodobnie niemożliwe do oddzielenia od pytania o temat badań. Zadaj sobie później pytanie o to, jakie aspekty jednostek analizy zapewniłyby zdobycie informacji potrzebnych do udzielenia odpowiedzi na pytanie badawcze.

Gdy będziesz już miał jakieś wyobrażenie na temat potrzebnego dla twoich celów rodzaju informacji, zapytaj siebie, co możesz zrobić, aby takie informacje uzyskać. Czy jest możliwe, że odpowiednie dane są już gdzieś dostępne (na przykład w oficjalnych dokumentach), czy będziesz musiał sam je zbierać? Jeśli musisz zebrać je sam, jak się do tego zabierzesz? Czy będziesz musiał przeprowadzić badania sondażowe na dużej próbie, czy pogłębione wywiady z niewielką liczbą osób? Czy możesz dowiedzieć się czegoś przydatnego poprzez udział w spotkaniach jakichś grup? Czy można uzyskać potrzebne dane z dostępnych w bibliotece książek?

W miarę udzielania odpowiedzi na te pytania, znajdziesz się w środku procesu planowania badań. Pamiętaj o własnych umiejętnościach badawczych oraz dostępnych środkach. Planowanie doskonałych badań nie ma sensu, jeśli w praktyce nie można ich przeprowadzić. Może zechcesz wypróbować metodę badawczą, której jeszcze nie stosowałeś, ale uważaj, aby nie stało się to zbyt dużym obciążeniem.

Gdy już masz pomysł, co oraz jak chcesz badać, dokładnie przejrzyj raporty z wcześniej-

szych badań w czasopismach i książkach, by przekonać się, jak podchodzili do tego tematu inni badacze i czego się dowiedzieli. Przegląd literatury może skłonić cię do modyfikacji planu badań; być może zdecydujesz się skorzystać z metody innego badacza albo nawet powtórzyć jakieś wcześniejsze badania. Niezależne powtarzanie projektów badawczych jest standardową procedurą w naukach przyrodniczych i jest równie ważne w naukach społecznych, mimo że przedstawiciele tych ostatnich mają skłonność do pomijania tego faktu. Możesz wyjść dalej poza powtórzenie i objąć badaniami jakiś aspekt tematu, który wydaje ci się pominięty przez poprzednich badaczy.

Oto inna droga, którą możesz obrać. Załóżmy, że jakiś temat był wcześniej badany metodami badań terenowych. Czy można zaplanować eksperyment, który sprawdzałby wyniki, uzyskane przez wcześniejszych badaczy? A może przychodzą ci do głowy gotowe dane statystyczne, które można by zastosować do testowania wniosków z wcześniejszych badań? Czy jakieś duże badania sondażowe przyniosły wyniki, które chciałbyś dokładniej przeanalizować za pomocą obserwacji bezpośredniej i wywiadów pogłębionych? Użycie kilku różnych metod badawczych do sprawdzenia tego samego wyniku jest czasem nazywane *triangulacją*. Powinieneś zawsze mieć ją na uwadze jako wartościową strategię badawczą. Każda z metod badawczych ma właściwe sobie mocne i słabe strony, zatem zawsze istnieje niebezpieczeństwo, że wyniki badań będą odzwierciedlać, przynajmniej po części, charakter użytej metody. W idealnych badaniach powinno się uwzględniać więcej niż jedną metodę zmierzenia się z tematem.

Oferta badawcza

Dość często w trakcie planowania badań będziesz musiał przedstawić komuś szczegóły badań w celu ich zatwierdzenia i/lub oceny. Na przykład w przypadku zajęć dydaktycznych prowadzący najprawdopodobniej zechce zobaczyć „ofertę", zanim zabierzesz się do pracy. Jeśli w swojej późniejszej karierze będziesz chciał podjąć się przeprowadzenia poważniejszego projektu badawczego, będziesz musiał zapewne uzyskać finansowanie od jakiejś fundacji czy agendy rządowej. Będą one zdecydowanie żądały przedstawienia szczegółowej propozycji, opisującej, w jaki sposób zamierzasz wydać ich pieniądze.

Rozdział ten kończy się krótkim omówieniem sposobu przygotowania takiej oferty. Jest to jeszcze jedno podsumowanie procesu badawczego, który jest szczegółowo opisany w dalszych rozdziałach.

Elementy oferty badawczej

Mimo że każda z zamawiających badanie instytucji (lub każdy z prowadzących zajęcia) może mieć specyficzne wymagania dotyczące elementów czy struktury oferty badawczej, poniżej znajdziesz podstawowe elementy, które powinna ona zawierać.

Problem lub cel

Co dokładnie chcesz badać? Dlaczego jest to warte zbadania? Czy proponowane badania będą miały znaczenie praktyczne? Czy przyczynią się do rozwoju teorii społecznych?

Przegląd literatury

Co inni napisali na ten temat? Jakie teorie się do niego odnoszą i co w tej sprawie mówią? Jakie badania były wcześniej przeprowadzane? Czy ich wyniki są ze sobą spójne, czy też wykluczają się nawzajem? Czy dotychczasowe badania mają jakieś niedoskonałości, którym twoje badania mogłyby zaradzić?

Obiekt badań

Kogo lub co będziesz badał, aby zebrać dane? Najpierw określ badanych w ogólnych, teoretycznych kategoriach. Potem opisz ich w kategoriach szczegółowych, bardziej konkretnych, określając dostępność tych osób i sposób dotarcia do nich. Czy właściwą rzeczą będzie dobranie próby? Jeśli tak, jak ją dobierzesz? Czy

może się zdarzyć, że twoje badania będą mieć wpływ na badane osoby? Jeśli tak, to jak zapewnić, że nie wyrządzą im one krzywdy?

Pomiar

Jakie są kluczowe zmienne w twoich badaniach? Jak je zdefiniujesz i zmierzysz? Czy twoje definicje i metody pomiaru są powtórzeniem definicji i metod zastosowanych we wcześniejszych badaniach tego tematu, czy też różnią się od nich? Jeśli opracowałeś już narzędzie badawcze (na przykład kwestionariusz) lub też zdecydowałeś się na używanie czyjegoś, właściwą rzeczą byłoby załączenie go do oferty badawczej.

Metody zbierania danych

Jaki będzie faktyczny sposób zbierania danych do twoich badań? Czy przeprowadzisz eksperyment, czy sondaż? Czy przeprowadzisz badania terenowe, czy może skupisz się na wtórnej analizie danych statystycznych zebranych przez innych badaczy? Być może zastosujesz więcej niż jedną metodę.

Analiza

Wskaż, jaki rodzaj analiz planujesz przeprowadzić. Sformułuj ich cel oraz logikę. Czy jesteś zainteresowany dokładnym opisem? Czy zamierzasz wyjaśniać istniejący stan rzeczy? Czy planujesz wyjaśnić przyczyny pewnych różnic – na przykład dlaczego jedni studenci są bardziej liberalni niż inni? Jakie możliwe zmienne wyjaśniające uwzględnisz w swojej analizie i skąd będziesz wiedzieć, że trafnie wyjaśniłeś różnice?

Harmonogram

Często przed przystąpieniem do badań warto określić harmonogram poszczególnych etapów badań. Nawet jeśli nie włączysz go do oferty badawczej, zrób to na własny użytek. Jeśli nie nakreślisz ram czasowych dla kolejnych etapów badań i bieżącego rejestrowania ich realizacji, możesz znaleźć się w tarapatach.

Budżet

Gdy prosi się kogoś o pokrycie kosztów badań, konieczne jest przedstawienie budżetu, który określa sposób wydatkowania środków. Duże, drogie projekty badawcze zawierają w swych budżetach takie kategorie, jak: personel, sprzęt, materiały biurowe, telefony, przesyłki pocztowe. Nawet w przypadku projektu, którego koszty pokrywa się z własnej kieszeni, warto spędzić trochę czasu i oszacować wydatki na: materiały biurowe, ksero, dyskietki, rozmowy telefoniczne, przejazdy itd.

Jak widzisz, jeśli chciałbyś zrealizować projekt badawczy z zakresu nauk społecznych, dobrze byłoby przygotować propozycję badań dla własnego użytku, nawet gdyby nie było takiego wymogu ze strony prowadzącego zajęcia czy finansującej badania instytucji. Jeśli zamierzasz zainwestować w taki projekt swój czas i energię, powinieneś zrobić co się da, aby zapewnić, że ta inwestycja się opłaci.

Teraz, gdy masz już za sobą obszerny przegląd badań społecznych, przejdźmy do dalszych rozdziałów tej książki, by dowiedzieć się, jak konkretnie planować i realizować każdy etap badań. Jeśli znalazłeś już temat badań, który naprawdę cię interesuje, miej go na uwadze w miarę zdobywania wiedzy o tym, jak możesz się zabrać do zbadania go.

Główne zagadnienia

- Głównymi celami badań społecznych są eksploracja, opis oraz wyjaśnianie. Badania często mają więcej niż jeden cel.
- Eksploracja to próba wstępnego, ogólnego zrozumienia jakiegoś zjawiska.
- Opis to dokładne sprawozdanie i/lub pomiar cech badanej populacji albo zjawiska.
- Wyjaśnianie jest odkrywaniem i opisywaniem związków między różnymi aspektami badanego zjawiska. Badania opisowe odpowiadają na pytanie „jak jest?", badania wyjaśniające natomiast odpowiadają na pytanie „dlaczego?".
- Jednostkami analizy są ludzie lub rzeczy, których cechy badacz społeczny obserwu-

je, opisuje i wyjaśnia. Najczęściej w badaniach społecznych jednostką analizy jest pojedynczy człowiek, ale może być to także grupa społeczna, organizacja formalna, wytwór społeczny albo jakieś inne zjawisko, takie jak style życia czy interakcje społeczne.

- Błąd ekologiczny polega na wyciąganiu wniosków z analizy cech grup (na przykład osiedli) i odnoszeniu ich do jednostek (na przykład mieszkańców tych osiedli).
- Redukcjonizm jest próbą wyjaśnienia złożonego zjawiska w kategoriach wąskiego zestawu pojęć; przykładem jest próba wyjaśniania rewolucji amerykańskiej wyłącznie w kategoriach ekonomicznych (albo psychologicznych, albo w kategoriach idealizmu politycznego).
- Badania procesów zachodzących w czasie mogą być zrealizowane w postaci badań przekrojowych lub dynamicznych.
- Badania przekrojowe są oparte na obserwacjach prowadzonych w jednym punkcie w czasie. Mimo że cecha ta stanowi ograniczenie tych badań, badacze mogą czasami wyciągać z nich wnioski dotyczące procesów zachodzących w czasie.
- W badaniach dynamicznych przeprowadza się obserwacje w różnych okresach. Takie obserwacje mogą być prowadzone na próbach dobranych z populacji generalnej (badania trendów), na próbach dobranych z pewnej liczby określonych subpopulacji (badania kohort demograficznych) albo za każdym razem na tej samej próbie (badania panelowe).
- Planowanie badań zaczyna się od początkowych zainteresowań badacza, idei lub oczekiwań teoretycznych i zmierza poprzez serię wzajemnie powiązanych etapów do zawężenia zakresu badań. Dzieje się tak poprzez ścisłe zdefiniowanie pojęć, metod oraz procedur badawczych. Dobry plan badań z góry uwzględnia wszystkie te kroki.
- Na samym początku badacz określa znaczenie pojęć lub zmiennych, które ma zamiar studiować (konceptualizacja), wybiera metodę lub metody badawcze (np. eksperyment, a nie sondaż), określa badaną populację oraz, tam gdzie ma to zastosowanie, sposób doboru próby.
- Badacz operacjonalizuje badane pojęcia poprzez precyzyjne określenie, jak będą mierzone zmienne. Badania przechodzą do fazy przeprowadzania obserwacji, przetwarzania danych, analizy i zastosowania w postaci opisania wyników i określenia ich implikacji.
- Oferta badawcza zawiera uzasadnienie, dlaczego badania zostaną podjęte, oraz skrócony opis procedury badawczej. Często potrzebne jest zezwolenie na realizację projektu badawczego, bądź trzeba zdobyć niezbędne środki. Nawet gdy nie jest wymagane przedstawienie oferty, jest ona poręcznym narzędziem planowania badań.

Kluczowe pojęcia

jednostki analizy
wytwory społeczne
błąd ekologiczny rozumowania
redukcjonizm
badania przekrojowe
badania dynamiczne
badania trendów
badania kohort demograficznych
badania panelowe

Pytania i ćwiczenia

1. Za pomocą programu Infotrac lub w bibliotece znajdź raport z badań, który jest przykładem eksploracji, opisu lub wyjaśniania. Określ, któremu z trzech rodzajów celów służyły omówione w raporcie badania, i krótko uzasadnij, dlaczego tak uważasz.
2. Oto kilka przykładów rzeczywistych tematów badań. Określ dla każdego z nich jednostkę analizy. (Odpowiedzi na końcu rozdziału).
   a. „Kobiety oglądają telewizję częściej niż mężczyźni, dlatego że na pracę poza domem przeznaczają na ogół mniej godzin niż mężczyźni (...). Czarnoskórzy oglądają dziennie telewizję średnio o 45 minut dłużej niż biali" (Hughes 1980, s. 290).

b. „W 1960 roku 126 spośród 130 amerykańskich miast liczących powyżej 100 tysięcy mieszkańców miało co najmniej dwa ogólne ambulatoria publiczne akredytowane przez Amerykański Związek Szpitali" (Turk 1980, s. 317).

c. „Pierwsze organizacje TM [medytacji transcendentalnej] były małe i nieformalne. Grupa z Los Angeles rozpoczęła działalność w czerwcu 1959 roku, spotykając się w domu jednego z członków, gdzie, zbiegiem okoliczności, mieszkał Maharishi" (Johnston 1980, s. 337).

d. „Mimo że wydaje się, że na decyzje o zmianach w systemie opieki medycznej duży wpływ mają pielęgniarki, proces podejmowania decyzji jest zdominowany przez administrację oraz wyższy personel medyczny" (Comstock 1980, s. 77).

e. „Mimo że 667 tysięcy spośród 2 milionów farmerów w USA to kobiety, historycznie nie są one postrzegane jako farmerzy, ale raczej jako żony farmerów" (Votaw 1979, s. 8).

f. „Analiza sprzeciwu społeczności lokalnych wobec wspólnych domów dla osób upośledzonych umysłowo wskazuje, że osiedla o pogarszającym się standardzie najczęściej tworzą zorganizowane formy sprzeciwu, lecz osiedla zamieszkiwane przez wyższą klasę średnią wykorzystują prywatne kontakty z lokalnymi urzędnikami" (Graham, Hogan 1990, s. 513).

g. „Niektórzy analitycy w latach 60. przewidywali, że wzrost ambicji gospodarczych i przebojowości politycznej czarnej ludności spowoduje niezadowolenie z „uduchowionych" kościołów, zdominowanych przez czarnych" (Ellison, Sherkat 1990, s. 551).

h. „W trakcie analizy badamy, czy twierdzenia oraz empiryczne odkrycia współczesnych teorii organizacji odnoszą się bezpośrednio zarówno do prywatnych przedsiębiorstw produkcyjnych, jak i do instytucji świadczących usługi publiczne dla ludności" (Schiflett, Zey 1990, s. 569).

i. „W artykule badamy zróżnicowania w strukturze nazw stanowisk pracy dla różnych ról zawodowych. Analizując 3173 nazwy stanowisk pracy w kalifornijskim systemie służby cywilnej w 1985 roku badamy, jak oraz dlaczego poszczególne piony różnią się pod względem ilości kategorii stanowisk, co wpływa na zróżnicowanie pozycji w hierarchii, funkcji oraz miejsca w strukturze organizacyjnej" (Strang, Baron 1990, s. 479).

3. Przejrzyj jakieś czasopismo naukowe, szukając przykładów przynajmniej trzech różnych jednostek analizy. Określ każdą z nich i przedstaw cytaty, by uzasadnić swoje wnioski.

4. Wymyśl przykład badań – inny niż omówione w tekście – który pokazuje, jak badacz wpada w pułapkę błędu ekologicznego. Potem zmodyfikuj przykład tak, aby uniknąć tej pułapki.

5. Zajrzyj na stronę internetową Russell Sage Foundation (http://www.russellsage.org) i przejrzyj dostępne tam publikacje. Wybierz po jednym przykładzie badań przekrojowych, badań trendów, badań kohort oraz badań panelowych. Uzasadnij swój wybór.

Dodatkowe lektury

Bart Pauline, Frankel Linda, *The Student Sociologist's Handbook*, General Learning Press, Morristown, N.J. 1986. Małe, poręczne kompendium, które pomoże ci rozpocząć projekt badawczy. Napisane z perspektywy pisania prac semestralnych, jest wyjątkowo dobrym przewodnikiem po zawartej w czasopismach literaturze z zakresu nauk społecznych, która powinna być dostępna w dobrej bibliotece.

Casley D. J., Lury D. A., *Data Collection in Developing Countries*, Clarendon Press, Oxford 1987. Książka omawia specyficzne problemy związane z prowadzeniem badań w krajach rozwijających się.

Cooper Harris M., *Integrating Research: A Guide for Literature Reviews*, Sage, Newbury Park,

Cal. 1989. Autor krok po kroku prowadzi przez proces przeglądania literatury.

Hunt Morton, *Profiles of Social Research: The Scientific Study of Human Interactions*, Basic Books, New York 1985. Zajmujący i pouczający zbiór opisów projektów. Przedstawia badania Jamesa Colemana dotyczące segregacji w szkołach, a także kilka innych ważnych projektów, które ilustrują elementy badań społecznych w praktyce.

Iversen Gudmund R., *Contextual Analysis*, Sage, Newbury Park, Cal. 1991. Analiza kontekstualna badająca wpływ czynników otoczenia społecznego na zachowanie się jednostki. Badania samobójstw Durkheima są tego dobrym przykładem, gdyż określają one społeczny kontekst wpływający na prawdopodobieństwo dokonania aktu samozniszczenia.

Maxwell Joseph A., *Qualitative Research Design: An Interactive Approach*, Sage, Newbury Park, Cal. 1996. Maxwell omawia wiele tematów zawartych w tym rozdziale, ale koncentruje się głównie na badaniach jakościowych.

Menard Scott, *Longitudinal Research*, Sage, Newbury Park, Cal. 1991. Autor rozpoczyna od wyjaśnienia, dlaczego przeprowadzane są badania dynamiczne, później zaś szczegółowo opisuje różne schematy badawcze, a także daje wskazówki dotyczące analizy danych dynamicznych.

Miller Delbert, *Handbook of Research Design and Social Measurement*, Sage, Newbury Park, Cal. 1991. Przydatne kompendium, ułatwiające wprowadzenie lub powtórzenie licznych kwestii, związanych z planowaniem badań i pomiarem. Dodatkowo znajdziesz w tej książce mnóstwo praktycznych informacji dotyczących fundacji, czasopism naukowych i stowarzyszeń zawodowych.

Odpowiedzi do pytań i ćwiczeń, punkt 2

a. mężczyźni i kobiety, biali i czarnoskórzy (jednostki)
b. uwzględnione amerykańskie miasta (grupy)
c. organizacje medytacji transcendentalnej (grupy)
d. personel pielęgniarski (grupy)
e. farmerzy (jednostki)
f. osiedla (grupy)
g. czarni (jednostki)
h. prywatne przedsiębiorstwa produkcyjne oraz instytucje świadczące usługi publiczne dla ludności (organizacje formalne)
i. nazwy stanowisk pracy (wytwory)

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 5

Konceptualizacja, operacjonalizacja i pomiar

Przegląd holograficzny

Powiązane ze sobą procesy konceptualizacji, operacjonalizacji i pomiaru pozwalają badaczowi przekształcić ogólną ideę tematu badań w użyteczny i trafny pomiar zjawisk świata rzeczywistego. Zasadniczym elementem tego procesu jest przekład stosunkowo niejednoznacznych pojęć języka potocznego na precyzyjne przedmioty badania o jasno zdefiniowanym i mierzalnym znaczeniu.

Wprowadzenie

Mierzenie wszystkiego, co istnieje
Koncepcje, pojęcia i rzeczywistość
Pojęcia jako konstrukty

Konceptualizacja
Wskaźniki i wymiary
Wzajemna wymienialność wskaźników
Definicje realne, nominalne i operacyjne
Tworzenie ładu pojęciowego
Przykład konceptualizacji: pojęcie anomii

Definicje w badaniach opisowych i wyjaśniających

Wybory dokonywane w procesie operacjonalizacji
Zakres zmienności
Zmienność między skrajnościami
Uwaga na temat wymiarów
Definiowanie zmiennych i wartości
Poziomy pomiaru
Implikacje poziomów pomiaru
Wskaźniki proste i złożone
Kilka przykładów możliwych wyborów w procesie operacjonalizacji
Operacjonalizacja trwa

Kryteria jakości pomiaru
Precyzja i poprawność
Rzetelność
Trafność
Kto decyduje, co jest trafne?
Napięcie między rzetelnością a trafnością

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Ten i następny rozdział dotyczą procesu, w toku którego badacze przechodzą od ogólnej idei tego, co chcieliby badać, do skutecznego i dobrze zdefiniowanego pomiaru w świecie rzeczywistym. Poniższy rozdział zawiera omówienie wzajemnie ze sobą powiązanych procesów konceptualizacji, operacjonalizacji i pomiaru. W rozdziale 6 opieram się na tej podstawie, wyjaśniając bardziej złożone typy pomiaru.

Rozpoczniemy od zastanowienia się nad żywionymi niekiedy nieświadomymi przekonaniami co do możliwości mierzenia treści ludzkiego życia: miłości, nienawiści, uprzedzeń, religijności, radykalizmu, alienacji. Odpowiedź na pytanie, czy pomiar taki jest możliwy, brzmi: tak, ale potrzeba kilku stronic, by wyjaśnić, w jaki sposób jest on możliwy. Skoro już ustalimy, że badacze mogą mierzyć wszystko, co istnieje, przejdziemy do etapów, które składają się na ten proces.

Mierzenie wszystkiego, co istnieje

Napisałem wcześniej, że jednym z dwóch fundamentów nauki jest obserwacja. Uczeni często używają terminu *pomiar* zamiast tego słowa, które może się wydawać związane z potoczną, bierną formą zachowania. Mają wówczas na myśli staranną, przemyślaną obserwację świata rzeczywistego, której celem jest dokonanie opisu przedmiotów i zdarzeń w kategoriach wartości (cech) składających się na zmienną.

Być może jesteś nastawiony nieco sceptycznie wobec możliwości naukowego pomiaru naprawdę istotnych aspektów ludzkiej egzystencji społecznej. Jeśli czytałeś raporty z badań dotyczących czegoś w rodzaju liberalizmu, religijności lub uprzedzeń, poczułeś się być może zawiedziony tym, w jaki sposób badacze dokonywali pomiaru przedmiotu swoich badań. Być może wydawało ci się, że ich podejście było zbyt powierzchowne, że zagubili najbardziej znaczące aspekty sprawy. Być może religijność mierzyli poprzez to, ile razy dana osoba poszła do kościoła, albo liberalizm poprzez to, jak ktoś głosował w jednych wyborach. Twoje niezadowolenie wzrosłoby niezawodnie, gdybyś stwierdził, że sam zostałeś błędnie zaklasyfikowany przez dany system pomiaru.

Twoje uczucie niezadowolenia jest odbiciem istotnej prawdy o badaniach społecznych: większość zmiennych, które chcemy badać, naprawdę nie istnieje w taki sposób, w jaki istnieją skały. W istocie są one wymyślone. Co więcej, rzadko tylko miewają one jedno, niedwuznaczne znaczenie.

Aby dobrze zrozumieć, co mam na myśli, wyobraź sobie, że chcemy badać sympatie partyjne. Aby zmierzyć tę zmienną, moglibyśmy przejrzeć rejestry głosujących, by sprawdzić, ilu badanych zarejestrowało się jako demokraci lub republikanie, przyjmując to jako miernik ich sympatii partyjnych. Moglibyśmy jednak również po prostu zapytać, z jaką partią dany badany się identyfikuje, i przyjąć te odpowiedzi jako miernik. Zauważ, że te dwie różne możliwości pomiaru opierają się na dwóch cokolwiek odmiennych definicjach „sympatii partyjnych". Nawet uzyskane wyniki mogłyby się różnić: ktoś zarejestrował się wiele lat temu jako demokrata, lecz z biegiem czasu skłaniał się coraz bardziej ku filozofii republikańskiej. Osoba, która nigdy się nie zarejestrowała, mogłaby – zapytana o swoje sympatie – podać partię, z którą się najbardziej identyfikuje.

To samo rozumowanie można zastosować do *wyznania*. Niekiedy odnosimy tę zmienną do członkostwa w określonym związku wyznaniowym. W innych przypadkach oznacza to po prostu religię, z którą się identyfikujemy, jeśli taka istnieje. Być może dla ciebie oznacza to jeszcze coś innego, na przykład chodzenie do kościoła.

Prawda jest taka, że ani „sympatie partyjne", ani też „wyznanie" nie mają żadnego *prawdziwego* znaczenia, jeśli przez „prawdziwe" rozumiemy: odpowiadające jakiemuś obiektywnemu aspektowi rzeczywistości. Zmienne te nie istnieją w naturze, są tylko pojęciami, które stworzyliśmy, i którym przypisaliśmy konkretne znaczenie w określonym celu, na przykład w celu przeprowadzenia badań społecznych.

Ale przecież, mógłbyś powiedzieć, „sympatie" czy „wyznanie", podobnie jak cała masa in-

nych rzeczy, którymi zajmują się badacze społeczni, na przykład uprzedzenia lub współczucie – są w *pewien sposób* realne. Wygłaszamy w końcu sądy na ich temat, na przykład: „W Happytown 55% dorosłych identyfikuje się z partią republikańską, a 45% z nich to wyznawcy Kościoła Episkopalnego. Ogółem mieszkańcy Happytown nie są uprzedzeni i wykazują się wysokim poziomem współczucia". Wiadomo, że tego typu sądy wygłaszają nie tylko badacze społeczni, ale i zwykli ludzie. Jeśli rzeczy, o których mówią, nie istnieją w rzeczywistości, to co naprawdę mierzymy i o czym mówimy?

Właśnie, o czym? Zastanówmy się nad tym przez chwilę, biorąc za przykład zmienną, którą zajmuje się wielu badaczy społecznych (a także wielu innych ludzi), a mianowicie uprzedzenia.

Koncepcje, pojęcia i rzeczywistość

Wędrując drogą życia, ty i ja zaobserwowaliśmy wiele rzeczy i na podstawie naszych obserwacji nabraliśmy przekonania, że są one prawdziwe, słyszeliśmy też relacje innych ludzi, które sprawiały wrażenie prawdziwych. Na przykład:

- Słyszeliśmy na własne uszy ludzi mówiących brzydkie rzeczy o członkach mniejszości etnicznych.
- Słyszeliśmy, że kobiety są gorsze od mężczyzn.
- Czytaliśmy o linczach dokonywanych na Afroamerykanach.
- Czytaliśmy, że kobiety i przedstawiciele mniejszości są gorzej wynagradzani za tę samą pracę.
- Dowiedzieliśmy się o „czystkach etnicznych" i wojnach, w których jedna grupa etniczna usiłowała zetrzeć inną z powierzchni ziemi.

Dysponując już pewnym doświadczeniem, zauważyliśmy kolejne fakty. Ludzie uczestniczący w linczach bardzo często również brzydko mówili o Afroamerykanach. Co więcej, wielu z nich było zdania, że kobieta powinna „znać swoje miejsce". W końcu olśniło nas, że tendencje do tych kilku wymienionych powyżej zachowań współwystępują u pewnych ludzi i mają jakiś wspólny element. W którymś momencie ktoś wpadł na świetny pomysł: „Nazywajmy tego typu ludzi w skrócie *uprzedzonymi*. Możemy stosować to pojęcie, nawet jeśli dana osoba nie przejawia wszystkich tego typu zachowań – jeśli tylko jest dość podobna do innych uprzedzonych".

Jako że jesteśmy co do zasady zgodnie zainteresowani maksymalizacją skuteczności, przystaliśmy na tę propozycję. Oto jak powstało pojęcie „uprzedzeń". Nigdy ich nie obserwowaliśmy. Zgodziliśmy się używać tego terminu jako uproszczenia, jako nazwy oznaczającej zbiór, jak się zdaje, powiązanych zjawisk, które każdy z nas obserwował w swoim życiu. Mówiąc w skrócie, wymyśliliśmy uprzedzenia.

Oto kolejny argument za tym, że uprzedzenia nie istnieją poza naszym ogólnym porozumieniem co do określonego sposobu stosowania tego terminu. Każdy z nas stworzył sobie w umyśle pewien obraz tego, co ogólnie oznacza ten zestaw zjawisk, które zaobserwowaliśmy, oraz tego, co je ze sobą łączy. Kiedy wypowiadam słowo *uprzedzenia*, wywołuje ono pewien obraz mentalny zarówno w twoim, jak i w moim umyśle. To tak jakby szuflady z danymi w naszych głowach zawierały tysiące kart, oznaczonych etykietą w prawym górnym rogu. Jedna z kart w moim i twoim umyśle oznaczona została słowem „uprzedzenia". Na twojej karcie jest wszystko, co słyszałeś o uprzedzeniach i wszystko, co zaobserwowałeś, a co uznałeś za przykład tego zjawiska. Na mojej karcie jest to, co ja słyszałem i zaobserwowałem jako przykład uprzedzeń – a zawartość mojej karty jest inna niż twojej.

Terminem technicznym określającym te obrazy mentalne, te karty w naszych mentalnych szufladach z danymi, jest *koncepcja*. Tak więc i ty, i ja mamy pewną koncepcję uprzedzeń. Nie możemy bezpośrednio komunikować naszych obrazów mentalnych, używamy więc terminów wpisanych w prawym górnym rogu naszych mentalnych kart jako sposobu przekazania naszych koncepcji i obserwacji z nimi powiązanych. Terminy te umożliwiają nam komunikację i w końcu porozumienie co do ich konkretnego znaczenia. W badaniach społecznych proces dochodzenia do porozumienia co do znaczenia terminów nazywamy

konceptualizacją, jego wynik zaś – *pojęciem* (*concept*).

Weźmy inny przykład koncepcji. Załóżmy, że mam spotkać znaną ci osobę o imieniu Patrycja. Pytam cię więc, jaka jest Patrycja. Załóżmy teraz, że widziałeś, jak Patrycja pomaga zagubionym dzieciom odnaleźć rodziców i wkłada małego ptaszka z powrotem do jego gniazdka. Patrycja skłoniła cię, byś nosił ubogim rodzinom indyki na Święto Dziękczynienia i odwiedził w Boże Narodzenie szpital dziecięcy. Widziałeś, jak Patrycja płacze na filmie o matce walczącej z losem, ażeby chronić swoje dziecko. Przeszukując swoje obrazy mentalne, znajdziesz prawdopodobnie wszystkie te zjawiska – albo ich większość – zapisane na jednej karcie z napisem „współczucie". Przejrzysz inne dane na tej karcie i stwierdzisz, że stanowią one trafny opis Patrycji. Powiesz mi więc: „Patrycja jest pełna współczucia".

Teraz z kolei ja zaczynam przeszukiwać moje szuflady z obrazami mentalnymi, póki nie znajdę karty z napisem „współczucie". Przeglądam, co jest zapisane na moich kartach i mówię: „O, świetnie". Czuję, że wiem już, jaka jest Pat, ale moje oczekiwania stanowią odbicie moich danych, nie twoich. Potem, kiedy już spotkam Patrycję, stwierdzę prawdopodobnie, że moje doświadczenia są zgodne z treściami zapisanymi na mojej karcie „współczucie" i powiem, że miałeś rację. Załóżmy jednak, że moje obserwacje okażą się sprzeczne z danymi na mojej karcie. Powiem ci więc, że nie sądzę, by Patrycja była osobą szczególnie pełną współczucia – i zaczniemy porównywać nasze zapisy.

Mówisz więc: „Widziałem kiedyś, jak Patrycja płakała na filmie o matce walczącej z losem, ażeby chronić swoje dziecko". Przeglądam moją „kartę współczucia" i nie znajduję niczego podobnego. Gdzie indziej w moim zbiorze danych odnajduję ten rodzaj zjawisk na karcie oznaczonej „sentymentalizm". Odpowiadam więc: „To nie jest współczucie. To sentymentalizm".

Ażeby wzmocnić moją argumentację, mówię ci, że widziałem, jak Patrycja odmówiła ofiarowania datku na organizację zajmującą się ochroną wielorybów przed wyginięciem. „To oznacza brak współczucia", twierdzę. Przeszukując twoje dane, znajdujesz wieloryby na dwóch kartach: „ochrona środowiska" oraz „randki międzygatunkowe" i mówisz mi to. W końcu zaczynamy porównywać dane na naszych kartach oznaczonych „współczucie". Odkrywamy wówczas, że mamy wiele odmiennych obrazów mentalnych odpowiadających temu terminowi.

W szerszej perspektywie język i komunikacja są skuteczne tylko wówczas, gdy ty i ja mamy w znacznym stopniu zbieżne rodzaje danych na odpowiednich kartach mentalnych. Podobieństwa tych stron odpowiadają konwencjom, przyjętym w naszym społeczeństwie. W miarę dorastania dowiadujemy się mniej więcej tych samych rzeczy przy pierwszym zetknięciu się z danym terminem. Słowniki formalizują owe konwencje. Każdy z nas kształtuje następnie swoje obrazy mentalne w taki sposób, by odpowiadały takim konwencjom. Ponieważ jednak każdy z nas ma różne doświadczenia i dokonuje różnych obserwacji, dwie osoby nigdy nie dochodzą do dokładnie tego samego zbioru danych na żadnej karcie w swoim systemie danych. Jeśli chcemy mierzyć „uprzedzenia" lub „współczucie", musimy najpierw ustalić, co dokładnie uznajemy dla naszych celów za uprzedzenia lub współczucie.

Wracając do mojego twierdzenia na początku tego rozdziału: możemy mierzyć wszystko, co istnieje. Możemy na przykład mierzyć, czy Patrycja naprawdę wkłada ptaszki do gniazdek, odwiedza szpitale w Boże Narodzenie, płacze w kinie czy też odmawia przyczynienia się do ochrony wielorybów. Wszystkie te zachowania istnieją, możemy więc je mierzyć. Ale czy Patrycja naprawdę jest pełna współczucia? Na to pytanie nie możemy odpowiedzieć. Nie możemy mierzyć współczucia w żadnym sensie obiektywnym, ponieważ nie istnieje ono w taki sposób, jak opisane przed chwilą zachowania. Współczucie istnieje jedynie w formie konwencji co do sposobów stosowania tego terminu w procesach komunikacji dotyczącej rzeczy istniejących naprawdę.

Pojęcia jako konstrukty

Jeśli pamiętasz omówienie postmodernizmu z rozdziału 2, zapewne przyznasz, że pewni ludzi sprzeciwialiby się zakresowi „realności",

którą przypisałem powyżej pewnym zjawiskom. Czy Patrycja „naprawdę" odwiedza szpitale w Boże Narodzenie? Czy „naprawdę" istnieją szpitale? A Boże Narodzenie? Nie zamierzam co prawda przyjmować w tym rozdziale radykalnie postmodernistycznego punktu widzenia, ale sądzę, że rozumiesz wagę tego trudnego problemu intelektualnego: co jest, a co nie jest prawdziwe?

W tym kontekście Abraham Kaplan (1964) wprowadził rozróżnienie trzech klas rzeczy, które mierzą uczeni. Pierwsza klasa to przedmioty bezpośrednio obserwowalne: możemy je obserwować łatwo i bezpośrednio, na przykład kolor jabłka albo krzyżyk postawiony w kwestionariuszu. Druga klasa obejmuje przedmioty obserwowalne pośrednio, wymagające „stosunkowo bardziej wyrafinowanych, złożonych i pośrednich obserwacji" (1964, s. 55). Zauważamy, że respondentka postawiła krzyżyk przy słowie „kobieta" w kwestionariuszu i pośrednio obserwujemy płeć tej osoby. Książki historyczne i protokoły z zebrań rad nadzorczych spółek umożliwiają pośrednie obserwacje minionych działań społecznych. Trzecią wreszcie klasę przedmiotów obserwowalnych stanowią konstrukty – twory teoretyczne, oparte na obserwacji, lecz nieobserwowalne ani bezpośrednio, ani pośrednio. Dobrym przykładem jest współczynnik inteligencji, IQ. Został on matematycznie *skonstruowany* na podstawie obserwacji odpowiedzi na wiele pytań wchodzących w skład testu na inteligencję. Nikt nie jest w stanie bezpośrednio obserwować IQ. Nie jest to ani trochę bardziej „realna" cecha ludzi niż współczucie lub bycie uprzedzonym.

Kaplan (1964, s. 49) definiuje *pojęcie* jako „rodzinę koncepcji". Pojęcie jest więc, jak stwierdza Kaplan, konstruktem, czymś, co tworzymy. Pojęcia takie jak współczucie i uprzedzenia to konstrukty powstałe z twoich koncepcji, moich koncepcji i koncepcji wszystkich ludzi, którzy kiedykolwiek używali tych terminów. Nie mogą one być obserwowane bezpośrednio lub pośrednio, ponieważ nie istnieją – wymyśliliśmy je.

Podsumowując: pojęcia są konstruktami wywiedzionymi z procesu wzajemnego uzgadniania obrazów mentalnych (koncepcji). Nasze koncepcje stanowią podsumowanie nagromadzonych obserwacji i doświadczeń, które sprawiają wrażenie wzajemnie powiązanych. Mimo że te obserwacje i doświadczenia są, przynajmniej subiektywnie, realne, koncepcje i wywiedzione z nich pojęcia są jedynie tworami mentalnymi. Terminy powiązane z koncepcjami są tylko narzędziami stworzonymi dla celów systematyzacji i komunikacji. Termin taki jak *uprzedzenia* jest, z obiektywnego punktu widzenia, tylko zbiorem liter. Nie ma poza tym żadnej wewnętrznej realności. Ma tylko takie znaczenie, jakie zgodzimy się mu nadać.

Często jednak wpadamy w pułapkę wiary, że terminy oznaczające konstrukty mają jakieś wewnętrzne znaczenie, że nazywają jakieś prawdziwe byty istniejące w rzeczywistości. Niebezpieczeństwo rośnie, gdy zaczynamy traktować te terminy poważnie i usiłujemy je precyzować. Niebezpieczeństwo staje się jeszcze większe, gdy pojawiają się eksperci, którzy, jak się wydaje, wiedzą więcej od nas o prawdziwym znaczeniu terminu: w takiej sytuacji bardzo łatwo jest poddać się autorytetowi.

Skoro już przyjmiemy, że pojęcia takie, jak *uprzedzenia* czy *współczucie* mają prawdziwe znaczenie, zaczyna się męka odkrywania tych prawdziwych znaczeń i tego, co stanowi ich prawdziwy pomiar. Postrzeganie pojęć jako przedmiotów rzeczywistych nazywamy *reifikacją*. Reifikacja pojęć w życiu codziennym jest dość powszechna. W nauce staramy się powiedzieć jasno, czym jest to, co w danej chwili mierzymy.

Czy z powyższych rozważań wynika, że *współczucie*, *uprzedzenia* i inne podobne konstrukty nie mogą być mierzone? Co ciekawe, odpowiedź brzmi: nie (i dobrze, bo inaczej większość badaczy społecznych, takich jak my, byłaby bezrobotna). Powiedziałem, że możemy mierzyć wszystko, co istnieje. Konstrukty nie istnieją naprawdę w taki sam sposób, jak drzewa, ale mają inną ważną zaletę: są użyteczne. Oznacza to, że pomagają nam porządkować, komunikować i rozumieć rzeczy, które *są* prawdziwe. Pomagają nam przewidywać zjawiska rzeczywiste. Niektóre nasze przewidywania okazują się nawet prawdziwe. Konstrukty mogą funkcjonować w ten sposób, ponieważ,

mimo że same nie są obserwowalne, pozostają w określonej relacji do rzeczy, które *są* prawdziwe i obserwowalne. Pomostem między przedmiotami bezpośrednio i pośrednio obserwowalnymi a użytecznymi konstruktami jest proces zwany konceptualizacją.

Konceptualizacja

Jak się przekonaliśmy, codzienna komunikacja zachodzi z reguły za pomocą systemu chwiejnych i ogólnych uzgodnień co do sposobów użycia terminów. Mimo że ty i ja nie zgadzamy się całkowicie co do użycia terminu *współczucie*, mogę prawdopodobnie z całym spokojem przypuszczać, że Patrycja nie będzie wyrywać muszkom skrzydełek. Szeroki zakres nieporozumień i konfliktów – począwszy od międzyludzkich, a skończywszy na międzynarodowych – to cena, którą płacimy za nasz brak precyzji. Mimo to jednak jakoś sobie radzimy. Ale nauka chce czegoś więcej – nie może ona funkcjonować w kontekście takiego braku precyzji.

Proces, w toku którego określamy, co mamy na myśli używając danego terminu w badaniach, nazywamy **konceptualizacją**. Załóżmy na przykład, że chcemy się dowiedzieć, czy kobiety są bardziej współczujące niż mężczyźni. Sądzę, że wiele osób przypuszcza, iż tak się właśnie sprawy mają, interesujące byłoby jednak sprawdzenie, czy tak jest w istocie. Nie możemy w sposób sensowny zbadać tego zagadnienia, ani nawet zgodzić się co do odpowiedzi, bez pewnych roboczych ustaleń co do znaczenia słowa *współczucie*. Ustalenia te nazywam „roboczymi", ponieważ pozwalają nam one pracować nad danym zagadnieniem. Nie musimy się zgadzać – czy choćby udawać, że zgadzamy się – co do tego, że jakieś konkretne określenie znaczenia jest najlepsze z możliwych.

W toku konceptualizacji powstaje zatem szczególne, uzgodnione dla celów badawczych znaczenie pojęcia. Proces określania dokładnego znaczenia wymaga opisu wskaźników, których będziemy używać dla pomiaru naszego pojęcia oraz rozmaitych jego aspektów, zwanych wymiarami.

Wskaźniki i wymiary

Konceptualizacja nadaje jasne znaczenie pojęciu poprzez wyszczególnienie jednego lub większej liczby wskaźników tego, co mamy na myśli. **Wskaźnik** to znak obecności lub nieobecności badanego pojęcia. Oto przykład.

Moglibyśmy się zgodzić, że odwiedzanie szpitali dziecięcych podczas świąt Bożego Narodzenia czy Chanuki jest wskaźnikiem współczucia. Wkładanie małych ptaszków z powrotem do gniazd byłoby innym wskaźnikiem itd. Jeśli jednostką analizy w naszym badaniu jest pojedyncza osoba, możemy obserwować obecność lub nieobecność każdego ze wskaźników dla każdej badanej osoby. Następnie możemy dodać liczbę wskaźników współczucia zaobserwowanych dla każdej jednostki. Moglibyśmy uzgodnić na przykład dziesięć takich konkretnych wskaźników i odkryć, że w naszym badaniu w przypadku Patrycji wystąpiło ich sześć, Janka – cztery, Marysi – dziewięć itd.

Wracając do naszego pytania o to, czy kobiety są bardziej współczujące niż mężczyźni, moglibyśmy obliczyć, że badane kobiety uzyskały średnio 6,5 wskaźnika współczucia, mężczyźni zaś przeciętnie 3,2. Na podstawie analizy ilościowej różnic międzygrupowych moglibyśmy więc dojść do wniosku, że kobiety są ogólnie bardziej współczujące niż mężczyźni.

Zwykle jednak sprawa nie jest taka prosta. Wyobraź sobie, że interesuje cię zrozumienie niewielkiego fundamentalistycznego ruchu religijnego, a zwłaszcza wrogości członków tego ruchu wobec rozmaitych grup: gejów, niewierzących, feministek itp. Właściwie ludzie ci twierdzą, że każdy, kto odmawia przyłączenia się do ich ruchu i posłuszeństwa ich naukom, będzie się „smażył w piekle". W kontekście twojego zainteresowania współczuciem wydaje się, że nie ma go wśród nich zbyt wiele. Jednak w pismach grupy często mowa jest o współczuciu jej członków dla innych. Chciałbyś zbadać ten pozorny paradoks.

Aby zrealizować swój zamiar, mógłbyś na przykład starać się o kontakt z członkami grupy, usiłując ich poznać i dowiedzieć się więcej o ich poglądach. Mógłbyś im powiedzieć, że jesteś badaczem zainteresowanym zdobyciem

wiedzy o ich grupie, mógłbyś też po prostu wyrazić zainteresowanie, nie podając powodów.

W toku rozmów z członkami grupy, a może i uczestnicząc w ich ceremoniach religijnych, znalazłbyś się prawdopodobnie w okolicznościach, w których mógłbyś zacząć rozumieć, co oznacza dla nich słowo *współczucie*. Mógłbyś się na przykład dowiedzieć, że członkowie grupy są tak bardzo poruszeni losem smażących się w piekle grzeszników, że są gotowi do podjęcia agresywnych czy wręcz brutalnych działań, by nawrócić grzesznych na dobrą drogę. Zgodnie ze swoim paradygmatem członkowie grupy postrzegają zatem bicie gejów, prostytutek i pracowników klinik aborcyjnych jako akty współczucia.

Badacze społeczni koncentrują się na znaczeniach nadawanych słowom i działaniom przez samych badanych. Takie podejście pomaga często objaśnić obserwowane zachowania: w tej chwili rozumiesz przynajmniej, jak wyznawcy badanego kultu mogą uważać akty przemocy za wyraz współczucia. Z kolei poświęcanie uwagi znaczeniom nadawanym słowom i działaniom przez badanych niemal zawsze owocuje komplikacjami w zakresie pojęć, którymi interesują się badacze. (Wrócimy do tego zagadnienia pod koniec rozdziału, przy okazji omawiania trafności mierników).

Zawsze gdy traktujemy nasze pojęcia poważnie i zabieramy się do wyjaśnienia tego, co przez nie rozumiemy, odkrywamy niezgodności i nieścisłości. Nie tylko ty i ja się nie zgadzamy, ale też każdy z nas mógłby prawdopodobnie znaleźć wiele niejasności wśród swoich własnych obrazów mentalnych. Jeśli zastanowisz się przez chwilę, co rozumiesz przez współczucie, stwierdzisz prawdopodobnie, że twój obraz zawiera kilka rodzajów współczucia. Znaczy to, że dane na twoich kartach mentalnych mogą być łączone w grupy i podgrupy, na przykład współczucie względem przyjaciół, współwyznawców, ludzi i ptaków. Możesz też wynaleźć kilka różnych strategii grupowania danych. Mógłbyś na przykład pogrupować dane w kategoriach uczuć i działań.

Technicznym terminem na określenie tego typu grupowania jest **wymiar**: możliwy do wyszczególnienia aspekt pojęcia. Moglibyśmy na przykład mówić o uczuciowym wymiarze współczucia i behawioralnym wymiarze współczucia. W innym schemacie grupowania moglibyśmy rozróżnić współczucie dla ludzi i współczucie dla zwierząt. Moglibyśmy też postrzegać współczucie jako dopomaganie ludziom, aby mieli to, czego dla nich chcemy, w odróżnieniu od pomagania ludziom, aby mieli to, czego chcą oni sami. Moglibyśmy również odróżnić współczucie jako gotowość do przebaczenia od współczucia jako litości.

Tak więc moglibyśmy podzielić współczucie na kilka jasno zdefiniowanych wymiarów. Pełna konceptualizacja wymaga zarówno określenia wymiarów, jak i wskazania różnych wskaźników dla każdego z nich.

Wyszczególnienie różnych wymiarów pojęcia stanowi często pierwszy krok na drodze ku pełniejszemu zrozumieniu przedmiotu badań. Moglibyśmy na przykład zaobserwować, że kobiety są bardziej współczujące w sferze uczuciowej, mężczyźni zaś w sferze behawioralnej – czy też na odwrót. Niezależnie jednak od wyniku, nie moglibyśmy powiedzieć, czy to kobiety, czy mężczyźni są naprawdę bardziej współczujący. Nasze badania wykazałyby, że nie ma jednej odpowiedzi na to pytanie – co samo stanowi już postęp w naszym rozumieniu rzeczywistości.

Wzajemna wymienialność wskaźników

Pojęcie wskaźników może również w inny sposób wspomagać nasze wysiłki podejmowane w celu zrozumienia rzeczywistości za pomocą „nierzeczywistych" konstruktów. Załóżmy na chwilę, że ty i ja sporządziliśmy listę 100 wskaźników współczucia i jego różnych wymiarów. Załóżmy dalej, że nie zgadzamy się zasadniczo co do tego, które ze wskaźników dają najlepszy dowód na obecność lub brak współczucia w danym przypadku. Jeśli bylibyśmy dość zgodni co do pewnych wskaźników, moglibyśmy skoncentrować uwagę właśnie na nich, zgadzając się zarazem prawdopodobnie co do wyników, których dostarczają. Bylibyśmy wówczas w stanie stwierdzić, że pewni ludzie są bardziej współczujący niż inni – w pewnych

wymiarach. Załóżmy jednak, że nie zgadzamy się co do żadnego z tych możliwych wskaźników. Mimo to moglibyśmy jednak osiągnąć porozumienie co do tego, czy kobiety są bardziej współczujące niż mężczyźni. Jest to możliwe dzięki wzajemnej wymienialności wskaźników.

Zasada jest następująca: jeśli nie zgadzamy się w ogóle co do wartości tych wskaźników, jednym z rozwiązań byłoby zbadanie ich wszystkich. Załóżmy, że kobiety okazałyby się bardziej współczujące w kategoriach wszystkich 100 wskaźników – wszystkich preferowanych przeze mnie i wszystkich preferowanych przez ciebie. Moglibyśmy wówczas zgodzić się, że kobiety są bardziej współczujące niż mężczyźni nawet mimo braku zgody co do tego, co właściwie znaczy współczucie jako takie.

Wzajemna wymienialność wskaźników oznacza, że jeśli kilka różnych wskaźników reprezentuje w określonym stopniu to samo pojęcie, będą one wszystkie zachowywać się tak, jak zachowywałoby się owo pojęcie, gdyby było rzeczywiste i mogło być obserwowane. Tak więc zakładając podstawową zgodę co do znaczenia słowa „współczucie", jeśli kobiety są ogólnie bardziej współczujące niż mężczyźni, powinniśmy móc zaobserwować tę różnicę przy użyciu jakiegokolwiek sensownego miernika współczucia. Jeśli z kolei kobiety są bardziej współczujące w oparciu o niektóre wskaźniki, ale nie w oparciu o inne, powinniśmy sprawdzić, czy te dwa zestawy wskaźników odpowiadają różnym wymiarom współczucia.

Znasz już podstawową logikę konceptualizacji i pomiaru. Dalsze rozważania zawierają głównie rozwinięcie i uściślenie tego, o czym przed chwilą przeczytałeś. Zanim przejdziemy do technicznego opracowania pomiaru, musimy jednak uzupełnić obraz konceptualizacji o kilka sposobów zapewnienia znaczeniu pojęć takich cech, jak standardowość, spójność i powszechność.

Definicje realne, nominalne i operacyjne

Jak już widzieliśmy, zaplanowanie i przeprowadzenie badań społecznych wymaga od nas wyjaśnienia ewentualnych nieporozumień co do pojęć i rzeczywistości. Logicy i uczeni uznali w tym celu za stosowne rozróżnić trzy rodzaje definicji: realne, nominalne i operacyjne.

Pierwsza z nich stanowi wyraz reifikacji terminu. Jak ostrzega Carl Hempel:

> „Definicja «realna», wedle logiki klasycznej, nie jest wyrazem uzgodnienia znaczenia pewnego wyrażenia, lecz stwierdzeniem «zasadniczej natury» lub «cech istotnych» pewnego bytu. Pojęcie zasadniczej natury jest jednak tak niejasne, że określenie to staje się bezużyteczne dla celów rygorystycznego badania".
>
> (1952, s. 6)

Innymi słowy, próba określenia „realnego" znaczenia pojęcia wiedzie nas na manowce: bierzemy wówczas konstrukt za byt rzeczywisty.

Określanie pojęć w badaniach naukowych opiera się natomiast na definicjach nominalnych i operacyjnych. *Definicja nominalna* zostaje po prostu przypisana do danego terminu bez żadnych roszczeń co do tego, że stanowi ona odbicie „realnego" bytu. Definicje nominalne są arbitralne – gdybym chciał, mógłbym zdefiniować *współczucie* jako „wyrywanie piórek bezradnym ptaszkom" – mogą jednak być mniej lub bardziej użyteczne. Dla większości celów, zwłaszcza dla celów komunikacji, ta ostatnia definicja współczucia byłaby raczej bezużyteczna. Większość definicji nominalnych stanowi odbicie jakiegoś konsensusu, konwencji co do sposobu użycia danego terminu.

Definicja operacyjna, jak sobie być może przypominasz z poprzedniego rozdziału, precyzuje sposób pomiaru danego pojęcia – czyli działania (operacje), jakie będziemy w tym celu wykonywać. Definicja operacyjna jest raczej nominalna niż realna, ma jednak tę zaletę, że daje maksimum jasności co do znaczenia danego pojęcia w konkretnych badaniach. We mgle nieporozumień i niejasności co do „prawdziwego" znaczenia terminu możemy podać definicję roboczą na potrzeby naszego badania. Zajmując się w jakichś badaniach na przykład statusem społeczno-ekonomicznym (SSE), moglibyśmy po prostu stwierdzić, że be-

dziemy go traktować jako kombinację dochodu i poziomu wykształcenia. Podejmując taką decyzję, odcinamy się od innych możliwych aspektów SSE: pozycji zawodowej, oszczędności, własności, pochodzenia itp. Nasze wyniki będą interesujące tylko w takim zakresie, w jakim nasza definicja SSE okaże się użyteczna dla naszych celów.

Tworzenie ładu pojęciowego

Wyjaśnianie pojęć jest w naukach społecznych procesem nieskończonym. Catherine Marshall i Gretchen Rossman (1995, s. 18) piszą o „lejku pojęciowym", dzięki któremu zainteresowania badacza coraz bardziej się koncentrują. Tak więc ogólne zainteresowanie „działalnością społeczną" może się zawęzić do „jednostek oddanych sprawie podmiotowych uprawnień i zmiany społecznej", następnie zaś do „doświadczeń kształtujących rozwój jednostek oddanych działalności społecznej". Ten proces koncentracji powiązany jest nierozerwalnie z językiem, którego używamy.

W niektórych typach badań jakościowych wyjaśnianie pojęć jest kluczowym elementem w procesie zbierania danych. Załóżmy, że prowadziłbyś wywiady i obserwacje w radykalnej grupie politycznej zwalczającej ucisk w społeczeństwie amerykańskim. Wyobraź sobie, jak zmieniałoby się znaczenie słowa „ucisk" w miarę jak zagłębiałbyś się coraz bardziej w doświadczenia i światopogląd badanych. Mógłbyś na przykład początkowo postrzegać ucisk w kategoriach fizycznych i, być może, ekonomicznych. Im więcej jednak dowiadywałbyś się o badanej grupie, tym więcej uwagi zacząłbyś poświęcać możliwości ucisku o charakterze psychologicznym.

Ta sama uwaga stosuje się również do kontekstów, w których znaczenia wydają się bardziej ustalone. Na przykład w analizie materiałów tekstowych badacze społeczni mówią niekiedy o „kręgu hermeneutycznym", czyli cyklicznym procesie pogłębiania rozumienia.

> „Do rozumienia tekstu dochodzimy w drodze procesu, w którym znaczenie pojedynczych części jest uwarunkowane przez wyprzedzające rozumienie całego tekstu. Bliższe określenie znaczenia pojedynczych części może ostatecznie zmienić pierwotnie antycypowane znaczenie całości, które znów wpływa na znaczenie pojedynczych części itd."
>
> (Kvale 1996, s. 46)

Weźmy pojęcie „uprzedzeń". Załóżmy, że musiałbyś napisać definicję tego terminu. Mógłbyś pomyśleć najpierw o uprzedzeniach rasowych/etnicznych. Na pewnym etapie zdałbyś sobie zapewne sprawę, że w twojej definicji powinno się też znaleźć miejsce na uprzedzenia związane z płcią, religijne, wobec gejów i inne. Badanie każdego z tych szczególnych typów uprzedzeń wpływałoby na twoje rozumienie ogólnego pojęcia uprzedzeń. W miarę zmian tego ogólnego rozumienia postrzegałbyś jednak prawdopodobnie każdą z owych indywidualnych form nieco inaczej.

Ciągłe uściślanie pojęć występuje we wszystkich metodach badań społecznych. Nawet pisząc już ostateczny raport z badań, skonstatujesz zapewne niejednokrotnie, że uściślasz właśnie znaczenie ważnych pojęć.

Mimo że konceptualizacja to proces ciągły, szczególnie istotne jest, by poświęcić mu dostatecznie wiele uwagi na początku każdego planu badań, zwłaszcza jeśli chodzi o schematy badań o ściśle określonej strukturze, jak sondaż lub eksperyment. W badaniach sondażowych na przykład wynikiem operacjonalizacji jest ustalenie konkretnego zestawu pytań kwestionariusza, które odpowiadać będą badanym pojęciom. Bez ich ustalenia badania nie posuną się naprzód.

Nawet w mniej ustrukturyzowanych metodach badawczych ważne jest jednak, by przyjąć za punkt wyjścia wstępny zbiór antycypowanych znaczeń, które mogą być uściślane w toku gromadzenia i interpretacji danych. Nikt poważnie nie wierzy, że możemy obserwować życie bez pojęć wstępnych. Z tego też powodu obserwatorzy życia społecznego muszą być świadomi tych pojęciowych punktów wyjścia i zdawać sobie z nich jasno sprawę.

Zajmijmy się wstępną konceptualizacją w takim zakresie, w jakim stosuje się ją do

ustrukturyzowanych badań, takich jak badania sondażowe lub eksperymenty. Mimo że dzięki określeniu definicji nominalnych ukierunkowujemy nasze strategie obserwacyjne, nie pozwala nam to jeszcze na obserwację. W kolejnym kroku musimy dokładnie określić, co zamierzamy obserwować, jak będziemy to czynić, oraz jakie interpretacje zamierzamy nadać różnym możliwym obserwacjom. Wszystkie te dalsze uszczegółowienia składają się na definicję operacyjną pojęcia.

Biorąc jako przykład status społeczno-ekonomiczny, moglibyśmy postanowić, że zadamy respondentom dwa pytania, zgodnie z podjętą decyzją co do pomiaru SSE w kategoriach dochodu i poziomu wykształcenia:

1. Jaki był całkowity dochód twojej rodziny w ciągu ostatnich 12 miesięcy?
2. Jaki poziom wykształcenia osiągnąłeś?

Aby uporządkować nasze dane, postanowilibyśmy prawdopodobnie wypracować system kategoryzacji udzielanych odpowiedzi. W przypadku dochodu moglibyśmy użyć kategorii, takich jak „poniżej 5000 USD", „od 5000 do 10 000 USD" itp. Poziom wykształcenia można również ująć w kategorie: poniżej średniego, średnie, niepełne wyższe, wyższe. Moglibyśmy na koniec określić, w jaki sposób odpowiedzi respondenta na te dwa pytania byłyby łączone, by stworzyć miernik SSE.

W ten sposób stworzylibyśmy roboczą i praktyczną definicję SSE. Definicja ta, niezależnie od tego, czy inni zgadzają się z naszą konceptualizacją i operacjonalizacją, miałaby jedną zasadniczą z punktu widzenia nauki zaletę: byłaby doskonale konkretna i jednoznaczna. Gdyby nawet ktoś nie zgodził się z naszą definicją, osoba ta wiedziałaby dokładnie, jak interpretować nasze wyniki badań, ponieważ przyjęte przez nas znaczenie SSE – znajdujące odbicie w naszej analizie i wnioskach – byłoby jasne i precyzyjne.

Oto diagram obrazujący kolejne kroki pomiaru, począwszy od naszego ogólnego rozumienia znaczenia danego terminu aż po precyzyjny pomiar we w pełni ustrukturyzowanym badaniu socjologicznym:

Konceptualizacja
↓
Definicja nominalna
↓
Definicja operacyjna
↓
Pomiary w świecie rzeczywistym

Przykład konceptualizacji: pojęcie anomii

Dla podsumowania powyższych rozważań na temat konceptualizacji przyjrzyjmy się krótko historii pewnego konkretnego pojęcia socjologicznego. Badacze zamieszek miejskich często interesują się szczególnie rolą odgrywaną w tych zdarzeniach przez poczucie bezradności. Przedstawiciele nauk społecznych używają niekiedy w tym kontekście słowa *anomia*. Termin ten został po raz pierwszy użyty przez wielkiego socjologa francuskiego Émile'a Durkheima w jego klasycznej pracy z 1897 roku, zatytułowanej *Le Suicide*.

Posługując się jedynie rządowymi raportami na temat liczby samobójstw w różnych państwach i regionach, Durkheim stworzył dzieło geniuszu analitycznego. Aby określić wpływ religii na samobójstwo, porównał liczby samobójstw w krajach protestanckich z tym samym wskaźnikiem w krajach katolickich, następnie w protestanckich regionach krajów katolickich i w katolickich regionach krajów protestanckich itd. Aby określić możliwy wpływ pogody, zbadał liczby samobójstw w krajach i regionach północnych i południowych, zbadał też zróżnicowanie liczby samobójstw w różnych miesiącach i porach roku. Mógł dzięki temu wyciągnąć wnioski co do aktu w najwyższym stopniu osobistego i indywidualnego, nie dysponując żadnymi danymi na temat zaangażowanych weń jednostek.

Na poziomie bardziej ogólnym Durkheim stwierdził, że samobójstwo odzwierciedla też stopień jasności i stabilności reguł panujących w społeczeństwie. Zauważywszy, że czas wielkich przemian społecznych często sprawia, iż jednostki doświadczają poważnych rozterek

co do stawianych im oczekiwań, Durkheim uznał, że rozterki takie powodują poczucie zagubienia, lęku, a nawet autodestrukcję. Dla opisania tego społecznego stanu braku norm Durkheim wybrał termin *anomia* (*anomie*). Nie wymyślił tego słowa – używane zarówno w niemieckim, jak i francuskim, znaczy ono dosłownie: bezprawie. Angielski termin *anomy* stosowany był w ciągu co najmniej trzech stuleci przed Durkheimem na oznaczenie braku szacunku dla prawa boskiego. Durkheim jest jednak twórcą socjologicznego pojęcia anomii.

Po publikacji *Samobójstwa* przedstawiciele nauk społecznych uznali anomię za pojęcie użyteczne i wielu z nich rozszerzyło durkheimowskie zastosowanie tego terminu. Robert Merton w klasycznym artykule zatytułowanym *Struktura społeczna i anomia* (1938) doszedł do wniosku, że anomia jest skutkiem dysproporcji między celami a środkami uznanymi przez społeczeństwo. Sukces finansowy na przykład jest w naszym społeczeństwie powszechnie podzielanym celem, jednak nie wszystkie jednostki mają wystarczające zasoby, by osiągnąć go za pomocą akceptowanych środków. Nacisk na sam cel, jak podkreślał Merton, wytwarza stan chaosu normatywnego, ponieważ jednostki pozbawione dostępu do tradycyjnych dróg wiodących do bogactwa starają się je osiągnąć środkami bezprawnymi. Rozważania Mertona można więc uznać za krok ku dalszej konceptualizacji pojęcia anomii.

Mimo że Durkheim, a po nim Merton, stosowali pierwotnie pojęcie anomii jako określenie cechy społeczeństw, inni przedstawiciele nauk społecznych używali go do opisu jednostek. By sprecyzować to rozróżnienie, niektórzy uczeni postanowili używać anomii (*anomie*) w jej oryginalnym, społecznym znaczeniu, innego terminu zaś (*anomia*) w odniesieniu do cech jednostek. W danym społeczeństwie niektóre jednostki doświadczają zatem anomii (*anomia*), inne zaś nie. Elwin Powell, piszący 20 lat po Mertonie, zaproponował następującą konceptualizację anomii jako cechy jednostkowej (choć używał terminu *anomie*):

> „Kiedy cele działania stają się sprzeczne, niedostępne lub nieistotne, stan anomii narasta. Anomię, cechującą się ogólną utratą orientacji i towarzyszącym jej uczuciem «pustki» i apatii, można uznać po prostu za stan zaniku sensu".
> (1957, s. 132)

Powell twierdzi dalej, że istnieją dwa różne rodzaje anomii, i bada, jak wyrosły one z różnych doświadczeń zawodowych, by w niektórych przypadkach znaleźć finał w samobójstwie. Powell nie mierzył jednak w swoim studium anomii jako takiej – badał on zależność między samobójstwem a zawodem, wyciągając wnioski co do dwóch rodzajów anomii. Badanie to nie dostarczyło więc definicji operacyjnej *anomii*, lecz jedynie dalszą konceptualizację tego pojęcia.

Mimo że wielu badaczy proponowało definicje operacyjne *anomii*, jedno nazwisko warte jest obszerniejszej wzmianki. Dwa lata przed publikacją artykułu Powella Leo Srole opublikował (1956) zestaw pytań ankietowych, który jego zdaniem stanowił dobry miernik anomii doświadczanej przez jednostki. Składał się on z pięciu zdań, z którymi badany mógł się zgodzić lub nie zgodzić:

„1. Wbrew temu, co się często słyszy, wielu przeciętnym ludziom wiedzie się coraz gorzej.
2. Nie jest w porządku sprowadzać dzieci na świat teraz, gdy perspektywy na przyszłość są takie, a nie inne.
3. W naszych czasach trzeba żyć dniem dzisiejszym, a jutro niech się martwi samo o siebie.
4. Nie wiadomo, na kogo można liczyć w dzisiejszych czasach.
5. Nie ma większego sensu pisać do urzędników, bo przecież problemy przeciętnego człowieka tak naprawdę ich nie interesują".
(1956, s. 713)

W dziesięcioleciach, które nastąpiły po jej publikacji, skala Srole'a stała się probierzem dla badaczy społecznych. Odnajdziesz prawdopodobnie tę właśnie operacjonalizację anomii w wielu projektach badawczych opisywanych w czasopismach naukowych. Srole zwraca na to uwagę w ramce „Źródła anomii", którą zdążył przed śmiercią przygotować specjalnie do tej książki.

Źródła anomii

Leo Srole

Centrum Geriatrii i Gerontologii, Uniwersytet Columbia

Moje trwające przez całe życie zawodowe zainteresowanie anomią zaczęło się od lektury *Le Suicide* Durkheima na studiach w Harvardzie. Potem, jako doktorant w Chicago, studiowałem pod kierunkiem dwóch antropologów ze szkoły durkheimowskiej: Williama Lloyda Warnera i Alfreda Radcliffe-Browna. Radcliffe-Brown prowadził ożywioną korespondencję z Durkheimem, co czyni mnie ciotecznym „potomkiem" wielkiego francuskiego socjologa.

Wpływ pracy Durkheima na temat samobójstw był w moim przypadku permanentny, aczkolwiek niejednoznaczny. Z jednej strony, podchodziłem z głęboką rezerwą do jego przemyślnych, wytężonych i niejednokrotnie zagmatwanych wysiłków, których celem było wtłoczenie surowych biurokratycznych danych na temat liczby samobójstw w sztywne ramy jego jednokierunkowego determinizmu społecznego. Z drugiej strony, poruszało mnie konsekwentne zainteresowanie Durkheima siłą moralną relacji międzyludzkich, które wiążą nas z naszym czasem, miejscem i przeszłością. Poruszała mnie również głębia jego wizji tragicznych następstw, które są skutkiem kurczenia się i rozluźniania tych więzi.

Moje zainteresowanie anomią otrzymało impuls w postaci naocznej obserwacji, gdy w ostatnich miesiącach II wojny światowej służyłem w *United Nations Relief and Rehabilitation Administration* (UNRRA), pomagając odbudować rozdartą przez wojnę Europę. W nazistowskim obozie koncentracyjnym w Dachau widziałem na własne oczy głębię dehumanizacji, do której siły makrospołeczne, podobne do tych, którymi zajmował się Durkheim, doprowadziły u jednostek takich jak Hitler, Eichmann i inni im posłuszni na wszystkich szczeblach nazistowskich fabryk śmierci.

Po powrocie z mojej pracy w UNRRA czułem, że dawno już nadszedł czas, by wyjaśnić dynamikę leżącą u podstaw dezintegracji więzi społecznych. Musieliśmy pracować szybko, rezygnując z nacisku na rozwój makroteorii na rzecz bezpośrednich spotkań z jednostkami, przy zastosowaniu nowo opracowanej metodologii badań sondażowych. Badania tego typu powinny były moim zdaniem skoncentrować się na szerszym spektrum zachowań patologicznych niż tylko samobójstwa.

Moje wstępne badania obejmowały wiele zróżnicowanych działań. W 1950 roku na przykład przeprowadziłem wywiady z próbą 401 kierowców autobusów w Springfield w stanie Massachusetts. Cztery lata później, dzięki Studium Zdrowia Psychicznego Środkowego Manhattanu, uzyskałem dostęp do znacznie szerszej populacji. Te i inne badania dały mi możliwość rozszerzenia i ulepszenia moich pomiarów tej cechy u jednostek, która stanowi odzwierciedlenie cechy makrospołecznej nazwanej przez Durkheima *anomią* (*anomie*).

Kiedy zacząłem we własnej pracy stosować pojęcie Durkheima, postanowiłem wkrótce, że należy ograniczyć jego zastosowanie do jego znaczenia makrospołecznego i oddzielić je wyraźnie od jego jednostkowych przejawów. Dla tego ostatniego celu narzucało się wręcz użycie pokrewnego, lecz dotąd zapomnianego greckiego terminu anomia (*anomia*).

Po raz pierwszy opublikowałem moje omówienie konstruktu dotyczącego anomii w roku 1956 w artykule w „American Sociological Review"*, opisując, w jaki sposób zoperacjonalizowałem to pojęcie, oraz prezentując wyniki wstępnych badań terenowych z jego zastosowaniem. Do 1982 roku Indeks Cytowań Naukowych i Indeks Cytowań Nauk Społecznych zamieściły ponad 400 wzmianek o publikacjach w czasopismach z zakresu politologii, psychologii, polityki społecznej i socjologii, które opisywały zastosowanie narzędzi lub wyników przedstawionych w tym artykule, co sprawiło, że Amerykański Instytut Informacji Naukowej uznał ów artykuł za „klasyk bibliograficzny".

* Leo Srole, *Social Integration and Certain Corollaries: An Exploratory Study*, „American Sociological Review" 1956, 21, s. 709–716.

Powyższa skrócona historia anomii jako pojęcia socjologicznego stanowi ilustrację kilku faktów. Po pierwsze, stanowi ona dobry przykład procesu przechodzenia ogólnych pojęć w zoperacjonalizowane pomiary. Nie oznacza to, że problem operacjonalizacji anomii został rozwiązany raz na zawsze. Uczeni będą z pewnością w dalszym ciągu latami konceptualizować i operacjonalizować to pojęcie, poszukując coraz bardziej użytecznych mierników.

Skala Srole'a to przykład jeszcze innego ważnego zjawiska: pozostawienie konceptualizacji i operacjonalizacji jako czynności otwartych nie musi, wbrew oczekiwaniom niektórych, prowadzić do anarchii i chaosu. Często w ten sposób powstaje ład. Przede wszystkim, mimo że możemy zdefiniować anomię w dowolny sposób, na przykład poprzez rozmiar buta, jest bardziej prawdopodobne, że zdefiniujemy ją w sposób niezbyt odmienny od obrazów mentalnych innych ludzi. Gdybyś chciał zastosować definicję naprawdę niestandardową, inni ludzie najprawdopodobniej by cię zignorowali.

Kolejnym źródłem ładu jest fakt, że badacze przejmują na ogół jakąś konceptualizację i operacjonalizację danego pojęcia, bo przekonali się o jej użyteczności. Prowadzi to do standaryzacji definicji pojęć. Poza skalą Srole'a można tu podać przykład testów IQ i mierników demograficznych i ekonomicznych wypracowanych przez Biuro Spisowe USA. Stosowanie takich ustalonych mierników ma dwie zalety: były one już wielokrotnie testowane i poprawiane, a badania stosujące te same skale można porównywać. Jeśli ty i ja zrobimy oddzielne badania dwóch różnych grup z użyciem skali Srole'a, możemy następnie porównać nasze dwie grupy pod względem anomii.

Przedstawiciele nauk społecznych mogą więc mierzyć wszystko, co istnieje. Dzięki konceptualizacji i operacjonalizacji mogą nawet zrobić kawał dobrej roboty, mierząc coś, co nie istnieje. Przyjmując, że pojęcia takie jak status społeczno-ekonomiczny, uprzedzenia, współczucie czy anomia nie są ostatecznie realne, przedstawiciele nauk społecznych mogą stworzyć pewien ład w zakresie ich stosowania. Ład ten opiera się jednak na użyteczności, nie zaś na prawdzie ostatecznej.

Definicje w badaniach opisowych i wyjaśniających

Jak sobie przypominasz z rozdziału 4, opis i wyjaśnianie to dwa zasadnicze cele badań. Rozróżnienie ich ma istotne implikacje dla definiowania i pomiaru. Mogłoby się wydawać, że opis jest zadaniem łatwiejszym niż wyjaśnianie, możesz jednak poczuć się nieco zdziwiony, gdy przekonasz się, że definicje sprawiają znacznie więcej kłopotów w badaniach opisowych niż w wyjaśniających. Zanim przejdziemy do innych aspektów pomiaru, musisz w ogólnych zarysach zrozumieć, dlaczego tak właśnie jest (szerzej zajmiemy się tym tematem w części 4).

Nietrudno zrozumieć wagę jasnych i precyzyjnych definicji dla badań opisowych. Jeśli chcemy opisać i przedstawić stopę bezrobocia w danym mieście, nasza definicja bezrobotnego jest oczywiście decydująca. Definicja ta zależna jest od definicji innego terminu: ludności aktywnej zawodowo. Uznanie trzylatka za bezrobotnego jest zupełnie nonsensowne, ponieważ takie dziecko nie należy jeszcze do ludności aktywnej zawodowo. Moglibyśmy więc pójść za sugestią Biura Spisowego USA i wykluczyć wszystkie osoby poniżej 14 roku życia.

Ta konwencja sama w sobie nie daje jednak satysfakcjonującej definicji, ponieważ w gronie bezrobotnych mogliby się wówczas znaleźć studenci, emeryci, niepełnosprawni, gospodynie domowe. Moglibyśmy więc stosować dalej konwencję Biura, definiując ludność aktywną zawodowo jako „wszystkie osoby powyżej 14 roku życia, zatrudnione, poszukujące pracy, urlopowane lub oczekujące na powrót do pracy, z której zostały tymczasowo zwolnione". Jeśli student, gospodyni domowa lub emeryt nie poszukują pracy, nie zaliczylibyśmy ich do ludności aktywnej zawodowo. Bezrobotnymi byliby wówczas ludzie zaliczeni do ludności aktywnej zawodowo, zdefiniowanej wedle powyższych kryteriów, którzy nie są zatrudnieni.

Ale co znaczy „poszukiwać pracy"? Czy dana osoba musi być zarejestrowana w Urzędzie Pracy, czy też powinna chodzić od drzwi do drzwi

Dlaczego ważne są nazwy zmiennych?

Patricia Fisher

Podyplomowe Studium Planowania, Uniwersytet Tennessee

O operacjonalizacji łatwiej się mówi, niż ją robi. Łatwo wyjaśnić cel i znaczenie operacyjnych definicji zmiennych, łatwo jest nawet opisać, jak na ogół przebiega operacjonalizacja. Jednak dopóki samemu się nie spróbuje zoperacjonalizować jakiejś dość złożonej zmiennej, można nie doceniać pewnych ukrytych problemów, które się z tym wiążą. Szczególnie dużą wagę ma dla operacjonalizacji wybór konkretnej nazwy zmiennej. Rozważmy przykład z dziedziny planowania przestrzeni miejskiej.

Zmienną, która ma wielkie znaczenie dla planistów, jest *udział mieszkańców*. Planiści sądzą, że uczestnictwo mieszkańców w procesie planowania jest istotne dla udanego wprowadzenia planu w życie. Udział mieszkańców pomaga zrozumieć rzeczywiste i postrzegane potrzeby społeczności, poza tym zaś zaangażowanie mieszkańców wzmaga zwykle ich chęć współpracy i wsparcie dla wysiłków planistów. Mimo że wielu różnych planistów mogłoby zaproponować rozmaite definicje pojęć, nie doszłoby prawdopodobnie między nimi do większych nieporozumień co do znaczenia wyrażenia *udział mieszkańców*. Nazwa tej zmiennej wydaje się adekwatna.

Gdybyśmy jednak poprosili planistów, by przedstawili bardzo proste mierniki operacyjne udziału mieszkańców, odkrylibyśmy prawdopodobnie wielkie zróżnicowanie ich odpowiedzi, powodujące liczne niejasności. Jeden z planistów mógłby uznać za taką miarę liczbę osób prywatnych obecnych na zebraniach komisji miejskiej i innych zebraniach władz lokalnych; inny notowałby tematy poruszane przez osoby prywatne na tego typu zebraniach; jeszcze inny wziąłby pod uwagę liczbę obecnych na zebraniach władz lokalnych, listów, telefonów do burmistrza i innych urzędników oraz spotkań zorganizowanych przez poszczególne grupy interesów w danym czasie. Widzimy od razu, że każdy z planistów mierzyłby (w bardzo uproszczony sposób) inny wymiar udziału mieszkańców: jego zakres, tematy pobudzające mieszkańców do aktywności, a także formę ich udziału. Tak więc całkowicie satysfakcjonująca z pojęciowego punktu widzenia pierwotna *nazwa* naszej zmiennej, *udział mieszkańców*, okazała się nieodpowiednia dla celów operacjonalizacji.

Precyzyjne i jednoznaczne nazywanie zmiennych pełni ważną funkcję w badaniach. Jest ono zarówno warunkiem, jak i efektem dobrej operacjonalizacji. Nazwy zmiennych często ewoluują w toku powtarzanego procesu tworzenia definicji pojęciowych, potem definicji operacyjnych, potem zmiany nazwy pojęcia tak, by lepiej pasowała do tego, co ma być mierzone. Ten zapętlony proces trwa w nieskończoność (powyższy przykład przedstawia tylko jedno z powtórzeń), czego skutkiem jest stopniowe udoskonalanie nazw zmiennych i ich pomiarów, aż do osiągnięcia dostatecznego stopnia ścisłości. Niekiedy ostateczne pojęcie zmiennej różni się nieco od pierwotnego, ale przynajmniej mierzymy to, o czym mówimy – choćby z tego powodu, że mówimy o tym, co mierzymy!

prosząc o zajęcie? Czy może wystarczy pragnienie posiadania pracy czy też otwartość na propozycję ewentualnego zatrudnienia? Tradycyjnie „poszukiwanie pracy" definiuje się operacyjnie jako udzielenie odpowiedzi twierdzącej na pytanie ankietera: „Czy w ciągu ostatnich 7 dni poszukiwał(a) Pan(i) pracy?". (Siedem dni to najczęściej występujący okres, ale dla pewnych celów badawczych wskazane może się okazać jego skrócenie lub wydłużenie).

Jak widać, wynik badań opisowych na temat stopy bezrobocia zależy wprost od tego, jak rozwiązujemy każdy z problemów definicyjnych. Wydłużenie okresu, podczas którego dana osoba jest postrzegana jako poszukująca pracy, włącza więcej niezatrudnionych osób do ludności aktywnej zawodowo zdefiniowanej jak wyżej, przez co wzrasta też wyliczona w ten sposób stopa bezrobocia. Jeśli wybierzemy inną konwencję i zdecydujemy się mówić wyłącznie

o cywilnej ludności aktywnej zawodowo i bezrobotnych cywilach, wykluczamy personel wojskowy; w ten sposób również podnosimy wyliczoną stopę bezrobocia, ponieważ personel wojskowy jest – z definicji – zatrudniony. W ten sposób zdanie opisowe mówiące, że stopa bezrobocia w danym mieście wynosi 3% albo 9%, albo wszystko jedno ile, zależy wprost od zastosowanych definicji operacyjnych.

Powyższy przykład jest stosunkowo jasny, ponieważ istnieje kilka powszechnie uznanych konwencji dotyczących ludności aktywnej zawodowo i bezrobocia. Pomyśl jednak, jak trudno byłoby osiągnąć porozumienie co do definicji niezbędnych dla wygłoszenia zdania: „Czterdzieści pięć procent studentów tej uczelni to konserwatyści". Podobnie jak stopa bezrobocia, odsetek ten byłby bezpośrednio zależny od definicji przedmiotu pomiaru – w tym przypadku konserwatyzmu politycznego. Zastosowanie innej definicji mogłoby przynieść wynik w rodzaju: „Pięć procent studentów to konserwatyści".

Jak na ironię, definicje w badaniach wyjaśniających są mniej problematyczne. Załóżmy, że interesuje nas wyjaśnienie politycznego konserwatyzmu. Dlaczego jedni są, a inni nie są konserwatywni? Załóżmy dla uściślenia, że interesuje nas, czy skłonności konserwatywne wzrastają z wiekiem. Co by było, gdybyśmy mieli 25 różnych definicji *konserwatyzmu* i nie mogli uzgodnić, która z nich jest najlepsza? Jak już widzieliśmy w rozważaniach na temat wskaźników, nie jest to bynajmniej przeszkoda nie do pokonania w naszych badaniach. Załóżmy, że odkryliśmy, iż ludzie starzy są bardziej konserwatywni niż młodzi przy zastosowaniu każdej z 25 definicji. Najwyraźniej zatem dokładna definicja nie wniosłaby zbyt wiele do sprawy. Załóżmy, że okazało by się, iż starzy ludzie są bardziej konserwatywni w kategoriach każdej sensownej definicji konserwatyzmu, jaka nam przyjdzie do głowy. Nie ma więc znaczenia, jak brzmi nasza definicja. Moglibyśmy dojść do wniosku, że starzy ludzie są ogólnie bardziej konserwatywni niż młodzi, nawet nie mogąc dojść do porozumienia co do dokładnego znaczenia słowa *konserwatywni*.

Powyższy ogólny wzorzec rozumowania jest dość powszechny w prowadzonych badaniach, mimo że w praktyce badania wyjaśniające niezmiernie rzadko dają wyniki tak niedwuznaczne, jak to przedstawiono w przykładzie. W życiu społecznym istnieją spójne wzorce zależności, które owocują spójnością wyników badań. Spójność taka nie zawsze jednak wychodzi na jaw w kontekście opisu. Skutkiem zmieniających się definicji niemal zawsze są odmienne wnioski opisowe. W ramce „Dlaczego ważne są nazwy zmiennych?" zagadnienie to zostało szerzej omówione w związku ze zmienną *udział mieszkańców*.

Wybory dokonywane w procesie operacjonalizacji

Omawiając zagadnienie konceptualizacji, często odwoływałem się do operacjonalizacji, ponieważ oba te procesy są ze sobą ściśle powiązane. Podsumowując: konceptualizacja polega na uściśleniu i udoskonaleniu abstrakcyjnych pojęć, operacjonalizacja zaś na stworzeniu konkretnych procedur badawczych (operacji), które pozwolą na dokonanie empirycznych obserwacji odpowiadających tym pojęciom w świecie rzeczywistym.

Podobnie jak w przypadku metod gromadzenia danych, badacze społeczni dysponują wieloma możliwościami operacjonalizacji pojęcia. Mimo że niektóre z nich są ściśle ze sobą powiązane, rozdzieliłem je dla dobra wykładu. Musisz jednak zdawać sobie sprawę, że operacjonalizacja nie polega na ciągłym sprawdzaniu listy procedur.

Zakres zmienności

Operacjonalizując dane pojęcie, badacz musi zdawać sobie sprawę z tego, jaki zakres zmienności go interesuje. Pytanie brzmi: do jakiego stopnia skłonni jesteśmy łączyć wartości w ogólniejsze kategorie?

Załóżmy, że chcesz mierzyć dochód badanych, gromadząc informacje z dokumentów lub wywiadów. Najwyższy roczny dochód sięga milionów dolarów, ale osiąga go bardzo nie-

wiele osób. Jeśli nie zajmujesz się ludźmi bardzo bogatymi, prawdopodobnie niewiele zmieni w twoich badaniach wyodrębnienie bardzo wysokich progów dochodu. Zależnie od tego, kim są badani, ustalisz zapewne najwyższą kategorię dochodu na znacznie niższym poziomie, na przykład 100 000 USD i więcej. Decyzja ta sprawi oczywiście, że wrzucisz do jednego worka ludzi zarabiających miliardy dolarów rocznie i biedaków z marnymi stu tysiącami, ale oni to jakoś przeżyją, a i twoim badaniom nie zaszkodzi takie wymieszanie. Podobnego wyboru musisz dokonać na drugim krańcu skali zarobków. W amerykańskich badaniach ogólnokrajowych na ogół nieźle sprawdza się dolna kategoria dochodu w wysokości 5000 USD lub mniej.

W badaniach postaw i poglądów kwestia zakresu zmienności ma również inny wymiar – jeśli nie jesteś dość ostrożny, możesz skończyć, mierząc tylko połowę możliwych postaw, mimo że zupełnie nie o to ci chodziło. Oto przykład.

Załóżmy, że interesują cię postawy wobec upowszechnienia się elektrowni jądrowych. Mógłbyś przypuszczać, że niektórzy ludzie uważają energię jądrową za największy wynalazek od czasów koła, inni zaś w ogóle się tym nie interesują. Wobec tego typu przypuszczeń wydaje się, że miałoby sens zadanie badanym pytania, na ile popierają coraz powszechniejsze wykorzystanie energii jądrowej, oraz przedstawienie im zakresu odpowiedzi od „bardzo popieram" do „w ogóle nie popieram".

Taka operacjonalizacja pomija jednak połowę spektrum możliwych postaw wobec energii jądrowej. Wielu ludziom energia jądrowa nie jest po prostu obojętna: ludzie ci aktywnie, choć w mniejszym lub większym stopniu, sprzeciwiają się jej rozpowszechnianiu. W tym przypadku po lewej stronie zera mamy również do czynienia ze znacznym zróżnicowaniem postaw. Niektórzy sprzeciwiają się trochę, inni dość zdecydowanie, jeszcze inni bardzo mocno. Aby zmierzyć cały zakres zmienności, musiałbyś zatem zoperacjonalizować postawy wobec energii jądrowej na skali od bardzo mocnego poparcia, poprzez obojętność, aż po bardzo silny sprzeciw.

Powyższe uwagi znajdują zastosowanie do wielu zmiennych w naukach społecznych. Niemal każdy publicznie dyskutowany temat wywołuje zarówno poparcie, jak i sprzeciw, każde z nich w różnym stopniu nasilenia. Poglądy polityczne przyjmują zakres od skrajnie liberalnych do skrajnie konserwatywnych – zależnie od typu osób, które badasz, może ci zależeć na uwzględnieniu poglądów radykalnych na jednym lub obu końcach skali. Podobnie jest z religijnością – ludzie nie są tylko mniej lub bardziej religijni, niektórzy są zdecydowanie antyreligijni.

Chodzi o to, że nie musisz mierzyć całego zakresu zmienności w każdym przypadku. Powinieneś jednak brać pod uwagę to, co jest ci potrzebne ze względu na cel twoich badań. Jeśli różnica między zwolennikami a przeciwnikami religii nie ma znaczenia dla twoich celów badawczych, zrezygnuj z niej. Ktoś zdefiniował kiedyś pragmatyzm jako postawę, zgodnie z którą „różnica, która nie robi żadnej różnicy, nie jest różnicą". Bądź pragmatykiem.

Pamiętaj także, że decyzje dotyczące zakresu zmienności powinny uwzględniać oczekiwania co do rozkładu wartości zmiennej w grupie badanych. W badaniu postaw profesorów wyższych uczelni wobec wartości wyższego wykształcenia mógłbyś prawdopodobnie spokojnie zatrzymać się na braku wartości i nie martwić się o tych, którzy mogliby uważać, że wyższe wykształcenie zagraża zdrowiu studenta (gdybyś jednak badał studentów...).

Zmienność między skrajnościami

Stopień precyzji to kolejna troska badacza podczas operacjonalizacji zmiennych. Sprowadza się ona do tego, z jakim stopniem dokładności określić różnice między różnymi możliwymi wartościami składającymi się na daną zmienną. Czy ma dla naszego badania znaczenie, czy dana osoba ma 17, czy 18 lat, czy też można przeprowadzić badanie pakując je obie do kategorii oznaczonej „od 10 do 19 roku życia"? Nie odpowiadaj zbyt pospiesznie! Gdybyś chciał badać liczbę zarejestrowanych i głosujących wyborców, zdecydowanie chciałbyś się do-

wiedzieć, czy twoi badani byli już uprawnieni do głosowania. Ogólnie rzecz biorąc, przed pomiarem wieku musisz przyjrzeć się celom i procedurom twoich badań i zdecydować, czy duże lub niewielkie różnice wieku są dla ciebie istotne. W sondażu podjęcie tych decyzji jest konieczne dla zaprojektowania właściwego kwestionariusza. W przypadku wywiadów pogłębionych decyzje te będą wpływać na stopień dopytywania o szczegóły.

Te same uwagi stosują się także do innych zmiennych. Jeśli mierzysz przekonania polityczne, czy będzie dla ciebie istotne pytanie, czy dana osoba jest raczej liberalnym, czy też raczej konserwatywnym demokratą, czy też wystarczy ci znajomość jej sympatii partyjnych? Przy pomiarze przekonań religijnych – czy wystarczy ci informacja, że dana osoba jest protestantem, czy też powinieneś znać także wyznanie? Czy musisz wiedzieć tylko, czy dany respondent jest wolny, czy pozostaje w związku małżeńskim, czy też robi ci różnicę informacja, że nie brał on nigdy ślubu, pozostaje w separacji, owdowiał lub jest rozwiedziony?

Nie ma oczywiście ogólnej odpowiedzi na takie pytania. Odpowiedź wynika z celów danych badań czy pomiaru. Mogę ci jednak udzielić pożytecznej wskazówki. Jeśli tylko nie jesteś pewien, jak wiele szczegółów powinieneś uwzględnić w pomiarze, uwzględniaj ich raczej za dużo niż za mało. Jeśli respondentka w wywiadzie pogłębionym dobrowolnie zeznaje, że ma 37 lat, zanotuj „37 lat", a nie „po trzydziestce". Analizując dane możesz zawsze połączyć szczegółowe wartości w ogólniejsze kategorie, nigdy jednak nie zdołasz odzyskać różnorodności, które wrzuciłeś do jednego worka dokonując obserwacji i pomiaru.

Uwaga na temat wymiarów

Mówiliśmy już o wymiarach jako o cechach pojęć. Kiedy badacz zabiera się do tworzenia operacyjnych mierników zmiennych, często odkrywa on – albo, co gorsza, nawet nie zauważa – że nie ma jeszcze pełnej jasności co do tego, które wymiary zmiennej naprawdę go interesują. Oto przykład.

Załóżmy, że badasz postawy obywateli wobec rządu i chciałbyś włączyć w to badanie ich odczucia związane z korupcją. Oto tylko niektóre z wymiarów, które mógłbyś badać:

- Czy ludzie sądzą, że rząd jest skorumpowany?
- Na ile ich zdaniem jest on skorumpowany?
- Na ile są pewni swoich ocen dotyczących poziomu korupcji w rządzie?
- Co sądzą o korupcji w rządzie jako o problemie społecznym?
- Co ich zdaniem jest przyczyną korupcji?
- Czy sądzą, że jest ona nieunikniona?
- Co ich zdaniem należałoby w tej sprawie zrobić?
- Co byliby gotowi zrobić osobiście, by wyeliminować korupcję w rządzie?
- Na ile są pewni, że gotowi byliby zrobić to, co twierdzą, że by zrobili?

Listę tę można by ciągnąć dalej – odczucia ludzi wobec korupcji mają wiele wymiarów. Należy jednak absolutnie wyjaśnić sobie, które z nich są istotne w naszych badaniach – w przeciwnym razie możesz zmierzyć, co ludzie *czują*, gdy myślą o korupcji, podczas gdy naprawdę chciałeś się dowiedzieć, jak oceniają jej zasięg, lub odwrotnie.

Gdy już określiłeś, jak będziesz zbierał dane (na przykład poprzez sondaż czy badania terenowe), oraz zdecydowałeś, jaki poziom zmienności jest odpowiedni, jaki potrzebny stopień precyzji między skrajnościami i jakie konkretne wymiary zmiennych cię interesują, możesz stanąć przed kolejnym wyborem: matematyczno-logicznym. Oznacza to, że możesz być zmuszony wybrać poziom pomiaru. By omówić to zagadnienie, musimy powrócić do wartości i ich relacji względem zmiennych.

Definiowanie zmiennych i wartości

Wartość zmiennej, jak zapewne pamiętasz, jest cechą czegoś. Na przykład „kobieta", „stary" albo „student". Zmienne zaś to logiczne zbiory wartości. Tak więc *płeć* to zmienna złożona z dwóch wartości: kobieta i mężczyzna.

Procesy konceptualizacji i operacjonalizacji można postrzegać jako określenie zmiennych i tworzących je wartości. W kontekście badań dotyczących bezrobocia *status zatrudnienia* to zmienna o wartościach „zatrudniony" i „bezrobotny" – listę wartości można by wydłużyć, aby objęła także inne omawiane wcześniej możliwości, jak na przykład gospodyni domowa.

Każda zmienna musi mieć dwie istotne cechy: po pierwsze, składające się na nią wartości muszą być wyczerpujące. Aby zmienna mogła być w jakikolwiek sposób użyteczna w badaniu, musimy być w stanie sklasyfikować każdą obserwację w kategoriach wartości składających się na zmienną. Jeśli skonceptualizujemy zmienną *sympatie partyjne* tylko w kategoriach wartości „republikanin" i „demokrata", znajdziemy się w kłopocie, ponieważ część badanych będzie się identyfikowała z Zielonymi, część z Partią Reform czy z jakąkolwiek inną, część zaś (niekiedy znaczny odsetek) powie, że nie ma żadnych sympatii partyjnych. Aby sprawić, by lista wartości była wyczerpująca, moglibyśmy dodać kategorie „inna" i „żadna partia nie jest mi bliska". Niezależnie od tego, co zrobimy, musimy być w stanie zaklasyfikować każdą obserwację.

Wartości zmiennej muszą być także wzajemnie rozłączne. Każda obserwacja musi dać się zaklasyfikować w kategoriach jednej i tylko jednej wartości. Musimy na przykład zdefiniować „zatrudniony" i „bezrobotny" w taki sposób, by nikt nie był jednocześnie jednym i drugim. Oznacza to, że musimy być w stanie zaklasyfikować osobę, która pracuje, ale jednocześnie szuka pracy. (Moglibyśmy na przykład natknąć się na pełnoetatowego boksera wagi ciężkiej poszukującego splendoru i dreszczyku emocji w zawodzie socjologa). W takim przypadku moglibyśmy tak zdefiniować te wartości, by bycie zatrudnionym miało pierwszeństwo. Wówczas każdy, kto pracuje, jest osobą zatrudnioną, niezależnie od tego, czy poszukuje właśnie czegoś lepszego.

Poziomy pomiaru

Wartości zoperacjonalizowane jako rozłączne i wyczerpujące mogą być powiązane również w inny sposób. Na przykład wartości składające się na zmienne mogą reprezentować różne poziomy pomiaru. W tej części rozdziału zajmiemy się czterema poziomami pomiaru: nominalnym, porządkowym, interwałowym i ilorazowym.

Pomiar nominalny

Zmienne, których wartości są jedynie wzajemnie rozłączne i wyczerpujące, to **miary nominalne**. Wśród przykładów wymienić można płeć, wyznanie, sympatie partyjne, miejsce urodzenia, ukończoną uczelnię wyższą czy kolor włosów. Mimo że wartości tworzące każdą z tych zmiennych – jak na przykład „mężczyzna" i „kobieta" składające się na zmienną *płeć* – różnią się od siebie (i wyczerpują wszystkie możliwe rodzaje ludzkiej płci), nie posiadają one żadnej dodatkowej struktury. Mierniki nominalne stanowią jedynie zbiór nazw lub określeń cech.

Wyobraź sobie grupę ludzi opisaną w kategoriach takiej zmiennej i podzieloną fizycznie na grupy odpowiadające jej wartościom. Powiedzmy na przykład, że prosimy dużą grupę ludzi, by ustawili się w grupach wedle stanów, w których się urodzili: urodzeni w Vermont w jednej grupie, urodzeni w Kalifornii w innej itd. Zmienną jest *miejsce urodzenia*, wartości to urodzeni w Vermont, urodzeni w Kalifornii itd. Wszyscy ludzie stojący w danej grupie mają przynajmniej jedną cechę wspólną i różnią się pod tym właśnie względem od ludzi w innych grupach. Nieistotne jest, gdzie tworzą się poszczególne grupy, jak blisko siebie stoją i jak są rozmieszczone w sali. Znaczenie ma jedynie fakt, że członkowie danej grupy urodzili się w tym samym stanie i że poszczególne grupy różnią się między sobą stanem, w którym przyszli na świat ich członkowie. W kategoriach zmiennej nominalnej możemy o dwóch osobach powiedzieć tylko tyle, że są takie same lub różne.

Pomiar porządkowy

Zmienne o wartościach, które możemy logicznie uporządkować, to **miary porządkowe**. Różne wartości zmiennej porządkowej odpowia-

dają względnie niższemu lub wyższemu poziomowi zmiennej. Zmienne tego typu to *klasa społeczna, konserwatyzm, alienacja, uprzedzenia, poziom intelektualny* itp. W kategoriach zmiennej porządkowej możesz nie tylko określić, czy dwie osoby są takie same, czy też różne, lecz również stwierdzić, że jedna z nich jest „bardziej" niż druga – a więc bardziej konserwatywna, bardziej religijna, starsza itd.

W naukach przyrodniczych twardość jest najczęściej przywoływanym przykładem zmiennej porządkowej. Możemy powiedzieć, że jeden materiał (na przykład diament) jest twardszy od drugiego (powiedzmy szkła), jeśli pierwszy zarysowuje drugi, a nie na odwrót. Próbując zarysować rozmaite materiały za pomocą innych moglibyśmy w końcu ułożyć kilka materiałów w rzędzie od najbardziej miękkiego do najtwardszego. Nie możemy nigdy powiedzieć, jak twardy jest dany materiał w kategoriach absolutnych, możemy tylko powiedzieć, jak jest on twardy w kategoriach względnych – który materiał jest od niego twardszy, a który bardziej miękki.

Wróćmy do poprzedniego przykładu z grupowaniem ludzi. Tym razem wyobraź sobie, że prosimy o ustawienie się w jednej grupie osoby, które ukończyły wyższą uczelnię, w drugiej grupie stawiamy absolwentów szkół średnich, w trzeciej zaś osoby, które nie ukończyły szkoły średniej. Taki sposób grupowania badanych odpowiada przyjętym wcześniej wymaganiom dotyczącym wzajemnej rozłączności i wyczerpywania się wartości zmiennej. Moglibyśmy jednak dodatkowo uporządkować logicznie te trzy grupy według względnego poziomu wykształcenia (cechy wspólnej) członków każdej z nich. Moglibyśmy ustawić trzy grupy w rzędzie, od najniższego do najwyższego poziomu wykształcenia. Uporządkowanie takie stanowiłoby fizyczną ilustrację miary porządkowej. Wiedząc, w jakich grupach są dwie jednostki, moglibyśmy stwierdzić, że jedna z nich ma wyższe, niższe lub takie samo jak druga wykształcenie.

Zauważ, że w tym przykładzie nie ma znaczenia, jak blisko lub jak daleko od siebie rozmieszczone są poszczególne grupy wykształcenia. Grupy z wykształceniem wyższym i średnim mogą stać oddalone od siebie o dwa metry, podczas gdy grupa bez wykształcenia średniego może stać 200 metrów dalej. Faktyczne odległości nie mają tu żadnego znaczenia – po prostu grupa z wykształceniem średnim powinna się znaleźć między grupą bez wykształcenia średniego a grupą z wykształceniem wyższym, inaczej porządek ustawienia będzie niepoprawny.

Pomiar interwałowy

Dla wartości niektórych zmiennych faktyczne odległości między nimi mają jednak znaczenie. Zmienne takie to **miary interwałowe**. W tym przypadku odległości logiczne między wartościami mogą być wyrażane w postaci standardowych interwałów o określonym znaczeniu.

Na przykład na skali temperatury Fahrenheita różnica, czy inaczej odległość między 80 i 90 stopniami jest taka sama, jak między 40 i 50 stopniami. Jednak 80 stopni Fahrenheita nie oznacza, że dany przedmiot jest dwa razy bardziej gorący niż inny przedmiot o temperaturze 40 stopni, ponieważ punkt zerowy w skali Fahrenheita (jak i Celsjusza) został wybrany arbitralnie: zero stopni nie znaczy, że faktycznie nie ma żadnego ciepła. Podobnie 30 stopni poniżej zera na którejkolwiek z tych skal nie oznacza o 30 stopni mniej niż brak ciepła. (W przeciwieństwie do tych dwóch skal, skala Kelwina rozpoczyna się od zera absolutnego, co oznacza całkowity brak ciepła).

Mierniki konstruowane sztucznie, takie jak mniej lub bardziej akceptowane standaryzowane testy na inteligencję, to niemal jedyne mierniki interwałowe powszechnie stosowane w badaniach społecznych. Interwał dzielący 100 i 110 punktów możemy uważać za identyczny z tym dzielącym 110 i 120 punktów, a wynika to z rozkładu zaobserwowanych wyników wielu tysięcy ludzi, którzy rozwiązywali te testy na przestrzeni wielu lat. Byłoby jednak błędem wnioskować, że osoba z wynikiem IQ 150 jest o pięćdziesiąt procent inteligentniejsza niż osoba z wynikiem 100. (Nie możemy, ściśle rzecz biorąc, uważać za całkowicie pozbawioną inteligencji osoby z wynikiem 0 w standardowym teście IQ, aczkolwiek możemy ją uważać za niewłaściwego kandydata na

profesora czy choćby studenta wyższej uczelni. Ale może dziekan...?).

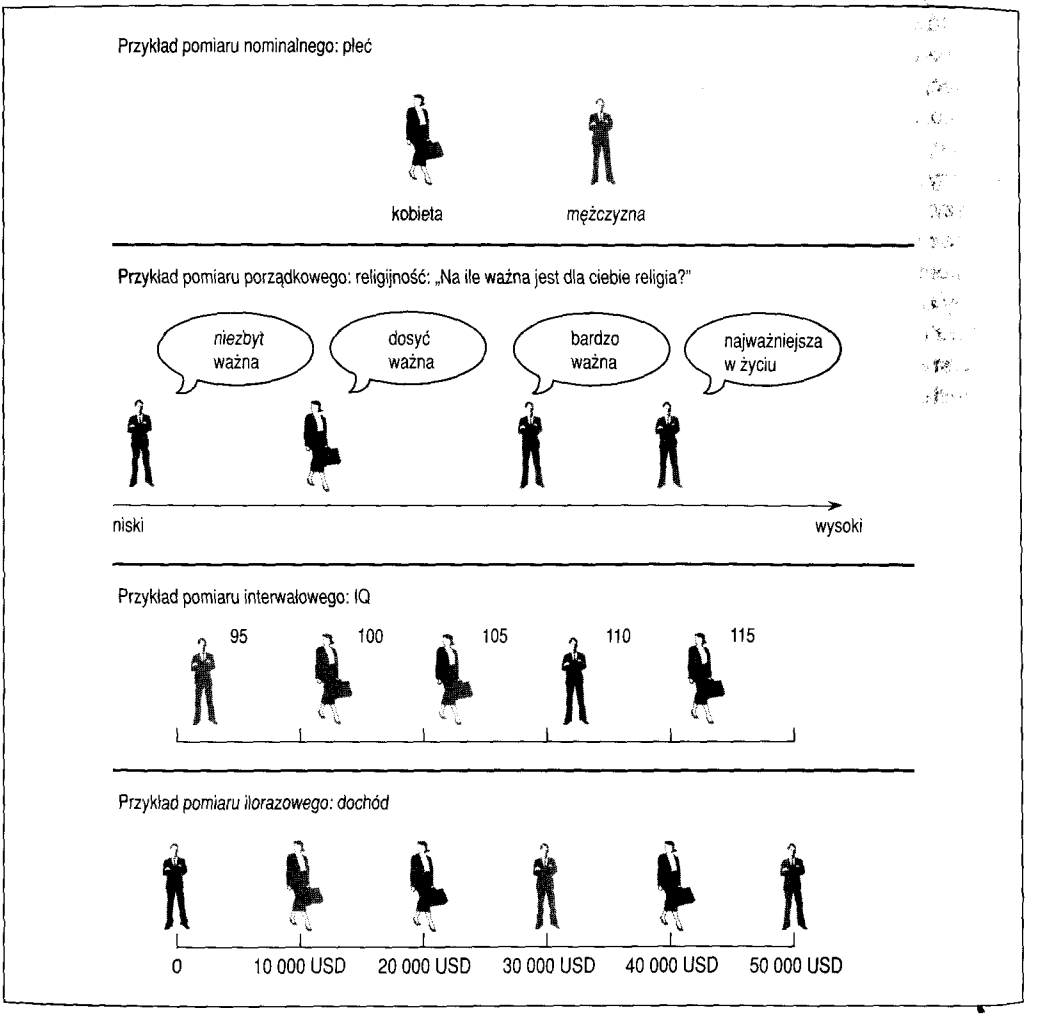
Porównując dwie osoby w kategoriach zmiennej interwałowej możemy stwierdzić, że są one od siebie różne (pomiar nominalny), oraz że jedna z nich jest „bardziej" niż druga (pomiar porządkowy). Oprócz tego zaś możemy jeszcze powiedzieć, „o ile" bardziej.

Pomiar ilorazowy

Większość stosowanych w naukach społecznych zmiennych, które spełniają *minimum* kryteriów pomiaru interwałowego, spełnia też kryteria pomiaru ilorazowego. W przypadku **pomiaru ilorazowego** wartości składające się na zmienną mają wszystkie powyżej opisane własności strukturalne, poza tym zaś oparte są na prawdziwym punkcie zerowym. Skala temperatury Kelvina jest jednym z tego typu mierników. Przykłady badań prowadzonych w naukach społecznych obejmują wiek, czas zamieszkiwania w określonym miejscu, liczbę organizacji, do których dana osoba należy, liczbę wizyt w kościele w danym okresie, liczbę zawartych małżeństw, liczbę przyjaciół narodowości arabskiej.

Rysunek 5.1

Poziomy pomiaru



Wracając do przykładu metodologicznej gry towarzyskiej, moglibyśmy poprosić zgromadzenie ludzi, ażeby pogrupowali się wedle wieku. Wszystkie jednolatki stałyby więc (ewentualnie siedziały lub leżały) razem, podobnie dwulatki, trzylatki itd. Fakt, że członkowie każdej z grup są w tym samym wieku oraz że każda grupa osób reprezentuje inny wiek, wystarcza dla spełnienia minimalnych wymagań stawianych zmiennym nominalnym. Uporządkowanie grup w szeregu od najmłodszej do najstarszej spełnia wymogi przewidziane dla zmiennych porządkowych i pozwala nam określić, czy dana osoba jest w tym samym wieku co inna, starsza od niej, czy też młodsza. *Jeśli rozmieścimy grupy w równych* odległościach, spełnimy dodatkowo warunek nałożony na zmienną interwałową, dzięki czemu będziemy w stanie powiedzieć, o ile lat jedna osoba jest starsza od drugiej. Wreszcie, ponieważ jedna z wartości zmiennej *wiek* odpowiada prawdziwemu zeru (dzieci noszone przez kobiety w ciąży tuż przed porodem), nasza grupa nieszczęsnych uczestników imprezy spełnia też wymogi pomiaru ilorazowego, co pozwala nam stwierdzić na przykład, że jedna osoba jest dwa razy starsza od drugiej. Innym przykładem miernika ilorazowego jest dochód, który obejmuje zakres od absolutnego zera właściwie do nieskończoności, jeśli przypadkiem jesteś założycielem firmy Microsoft.

Porównanie dwóch osób w kategoriach zmiennej ilorazowej pozwala nam więc stwierdzić, że (1) różnią się one lub są takie same, (2) jedna jest „bardziej" niż druga, (3) na ile się różnią i (4) jaki jest stosunek ilościowy jednej do drugiej (iloraz). Rysunek 5.1 przedstawia podsumowanie tej części rozważań w formie ilustracji graficznej czterech poziomów pomiaru.

Implikacje poziomów pomiaru

Jako że małe jest prawdopodobieństwo, by udało ci się dokonać dopiero co opisanych fizycznych przegrupowań ludzi (spróbuj tylko, jeśli nie chcesz być niepokojony zbyt licznymi zaproszeniami na imprezy), chciałbym zwrócić twoją uwagę na pewne praktyczne implikacje wprowadzonych powyżej rozróżnień. Implikacje te przejawiają się przede wszystkim w analizie danych (omówionej w części 4), musisz jednak przewidywać je już na etapie tworzenia struktury projektu badawczego.

Pewne techniki analizy ilościowej wymagają użycia zmiennych, które mają określony minimalny poziom pomiaru. Powinieneś zaplanować swoje techniki analityczne stosownie do poziomu pomiaru, do którego ograniczone są zmienne przeznaczone do analizy w projekcie badawczym, na przykład poziomu porządkowego. Dokładniej mówiąc, powinieneś przewidywać, jakie wnioski badawcze możesz wyciągnąć, dysponując zmiennymi o określonym poziomie pomiaru. Mógłbyś na przykład postanowić zbadać i opisać średni wiek badanej populacji (dodając wiek poszczególnych jednostek i dzieląc sumę przez ich liczbę), nie powinieneś jednak planować badania średniej przynależności religijnej, ponieważ jest to zmienna nominalna, średnia zaś wymaga danych na poziomie ilorazowym. (Mógłbyś jednak podać modalną – najczęstszą przynależność religijną).

Możesz przy tym traktować niektóre zmienne jako odpowiedniki różnych poziomów pomiaru. Mierniki ilorazowe to poziom najwyższy, z którego schodzimy, poprzez mierniki interwałowe i porządkowe, aż do najniższego poziomu pomiaru, czyli mierników nominalnych. Zmienną odpowiadającą wyższemu poziomowi pomiaru – powiedzmy, ilorazową – możemy traktować także jako zmienną na niższym poziomie pomiaru – na przykład jako zmienną porządkową. Przypomnij sobie na przykład wiek, który jest zmienną ilorazową. Gdyby interesowała cię jedynie relacja między wiekiem a jakąś zmienną na poziomie porządkowym – na przykład deklarowaną religijnością: duża, średnia i mała – mógłbyś zdecydować, że także wiek będziesz traktował jako zmienną porządkową. Mógłbyś opisywać badanych jako młodych, w średnim wieku i starych, precyzując, jaki zakres wieku należy do danej kategorii. Mógłbyś wreszcie, dla pewnych celów badawczych, używać wieku jako zmiennej nominalnej. Mógłbyś podzielić ludzi na urodzonych w czasie kryzysu lat 30. lub kiedy indziej. Innym pomiarem nominalnym, opartym nie tyle

na wieku, co na dacie urodzenia, mógłby być pomiar ludzi podzielonych na grupy według znaków zodiaku.

Poszukiwany poziom pomiaru zależy zatem od planowanych sposobów analitycznego wykorzystania danej zmiennej, z uwzględnieniem faktu, że pewne zmienne są nierozerwalnie związane z pewnymi poziomami pomiaru. Jeśli zmienna ma być używana na wiele sposobów, wymagających różnych poziomów pomiaru, badanie powinno być zaprojektowane tak, by pozwoliło na osiągniecie najwyższego wymaganego poziomu. Jeśli na przykład zapytasz badanych o ich dokładny wiek, możesz ich potem ułożyć w podgrupy nominalne lub porządkowe.

Nie musisz jednak koniecznie mierzyć zmiennych na najwyższym poziomie pomiaru. Jeśli jesteś pewny, że nie są ci potrzebne dane o wieku badanych na poziomie pomiaru wyższym niż porządkowy, możesz ich po prostu poprosić, by wskazali, w którym przedziale wiekowym się mieszczą, na przykład od 20 do 29, od 30 do 39 itd. W badaniach stanu posiadania przedsiębiorstw nie musisz gromadzić szczegółowych informacji – wystarczy, że posłużysz się skalą porządkową Dun & Bradstreet. Jeśli jednak twoje cele badawcze nie są do końca sprecyzowane, staraj się dotrzeć do najwyższego możliwego poziomu pomiaru. Choć można zredukować pomiar ilorazowy do poziomu porządkowego, nie można przekształcić pomiaru porządkowego na ilorazowy. Ogólnie rzecz biorąc, nie można przekształcić zmiennej na niższym poziomie pomiaru w zmienną na poziomie wyższym. Warto dobrze zapamiętać tę jednokierunkową drogę.

Wskaźniki proste i złożone

Mając do dyspozycji tak wiele możliwości operacjonalizacji zmiennych w naukach społecznych, możesz zacząć się poważnie martwić, czy dokonasz właściwego wyboru. Pozwól, że postaram się rozjaśnić twoje myśli, dodając do poprzednich wywodów odrobinę poczucia pewności i stałości.

Wiele zmiennych w naukach społecznych ma w miarę proste, oczywiste wskaźniki. Jakby na to nie patrzeć, płeć okazuje się na ogół kwestią bycia mężczyzną lub kobietą: zmienna na poziomie nominalnym, którą można mierzyć jedną obserwacją – wzrokową (no, może nie zawsze) lub odpowiedzią na pytanie (na ogół). W badaniu dzietności małżeństw mógłbyś co prawda wziąć pod uwagę dzieci adoptowane i przyjęte do rodzin zastępczych oraz dzieci z poprzedniego małżeństwa jednego lub obojga partnerów, zazwyczaj jednak stosunkowo łatwo jest się dowiedzieć, ile dzieci ma dana para. W przypadku większości badań liczba mieszkańców danego kraju jest liczbą mieszkańców tego kraju – sprawdzasz to w roczniku statystycznym i znasz odpowiedź. Tak więc wiele zmiennych ma oczywiste proste wskaźniki. Docierając do pojedynczej informacji, wiesz wszystko, co jest ci potrzebne.

Niekiedy jednak zdarza się, że nie istnieje jeden wskaźnik, który da ci pomiar potrzebnej zmiennej. Jak już pisałem wcześniej w tym rozdziale, wiele pojęć podlega różnorodnym interpretacjom, z których każda ma kilka możliwych wskaźników. W tego typu przypadkach mógłbyś dokonać kilku obserwacji danej zmiennej, następnie zaś połączyć zebrane informacje w celu stworzenia złożonego miernika badanej zmiennej. Rozdział 6 poświęcony jest metodom przeprowadzenia tej procedury, w tym miejscu rozważymy więc jedynie krótki przykład.

Weźmy pojęcie „wyników w nauce". Wszyscy zauważyliśmy, że niektórzy studenci mają dobre, inni zaś słabe wyniki w nauce. Badając te różnice, moglibyśmy zadać sobie pytanie, jakie cechy i doświadczenia są powiązane z osiąganiem dobrych wyników w nauce (wielu badaczy tak właśnie zrobiło). Jak można zmierzyć ogólne wyniki w nauce? Każda ocena z każdego egzaminu jest potencjalnym wskaźnikiem wyników w nauce, może jednak nie być typowa dla ogólnych wyników danego studenta. Rozwiązanie tego problemu ustaliło się już na tyle, że uważane jest za oczywiste: średnia punktów (*grade point average* – GPA). Przypisujemy każdej ocenie określoną liczbę punktów, dodajemy punkty uzyskane przez danego studenta i dzielimy je przez liczbę zdanych egzaminów, otrzymując w ten sposób miernik złożony. (Jeśli egzaminom przypisane są różne

wartości punktowe, obniżamy lub podwyższamy odpowiednią wartość punktową). Tworzenie tego typu mierników złożonych jest często bardzo przydatne w badaniach społecznych.

Kilka przykładów możliwych wyborów w procesie operacjonalizacji

Aby podsumować wszystkie możliwości operacjonalizacji dostępne badaczowi społecznemu, wskazując zarazem potencjał tkwiący w każdej z nich, rozważmy kilka alternatywnych podejść do rozmaitych problemów badawczych. Różne sposoby operacjonalizacji zmiennych w każdym przypadku mają ilustrować, jak wielkie pole do popisu dla naszej wyobraźni i kreatywności stwarzają badania społeczne. Dla uproszczenia nie usiłowałem nawet opisywać wszystkich warunków, które mogłyby przemawiać za jedną lub drugą możliwością, aczkolwiek w danych okolicznościach nie wszystkie one byłyby tak samo godne polecenia.

Oto konkretne problemy badawcze i kilka sposobów ich rozwiązania. Zacznijmy od przykładu omawianego już wcześniej w tym rozdziale – jego zaletą jest fakt, że operacjonalizacja jednej ze zmiennych jest bardzo prosta.

1. Czy kobiety są bardziej skłonne do współczucia niż mężczyźni?
   a. Wybierz grupę badanych, złożoną z takiej samej liczby mężczyzn i kobiet. Przedstaw im hipotetyczne sytuacje, w których ktoś jest w tarapatach. Zapytaj, co by zrobili, gdyby natknęli się na taką sytuację – na przykład małe dziecko płaczące i zagubione w poszukiwaniu rodziców. Każdą odpowiedź, która wskazuje na chęć pocieszenia lub pomocy dziecku, uznaj za wskaźnik współczucia. Sprawdź, czy kobiety, czy też mężczyźni częściej wskazują, że byliby skłonni do współczucia.
   b. Zaprojektuj eksperyment, w którym płacisz małemu dziecku za to, by udawało, że się zgubiło. Postaw dziecko na ruchliwym chodniku i obserwuj, czy częściej oferują mu pomoc kobiety, czy mężczyźni. Upewnij się, że policzyłeś ogólną liczbę przechodniów obu płci, ponieważ jedna z nich mogła być tam liczniej reprezentowana. W takim przypadku po prostu oblicz odsetek kobiet i mężczyzn skłonnych do udzielenia pomocy.
   c. Wybierz próbę ludzi i przeprowadź sondaż, w którym zapytasz ich, do jakich organizacji należą. Przelicz następnie, czy kobiety częściej niż mężczyźni należą do organizacji, które odpowiadają potrzebie współczucia swoich członków. Ażeby uwzględnić fakt, że mężczyźni – albo kobiety – należą w ogóle do większej liczby organizacji, musisz postąpić następująco: dla każdego badanego obliczasz odsetek organizacji związanych z realizacją skłonności do współczucia, do których należy. Sprawdź, czy kobiety mają wyższe niż mężczyźni przeciętne wyniki procentowe.
2. Czy studenci socjologii mają większą wiedzę o świecie współczesnym niż studenci rachunkowości?
   a. Przygotuj krótki quiz na temat zagadnień z wiedzy o świecie współczesnym, następnie zaś przeprowadź go na zajęciach dla studentów równoległych lat kierunków socjologii i rachunkowości. Jeśli chcesz porównać studentów ostatniego roku, upewnij się, z jakiej dziedziny piszą prace magisterskie.
   b. Zapytaj wykładowcę wiedzy o świecie współczesnym o przeciętne oceny studentów socjologii i rachunkowości na jego egzaminach.
   c. Przyjdź na zajęcia dla studentów socjologii i na zajęcia dla studentów rachunkowości z petycją głoszącą „Nowy Jork siedzibą ONZ". Policz tych, którzy podpiszą petycję, i tych, którzy powiedzą ci, że Nowy Jork jest przecież siedzibą ONZ.
3. Czy ludzie uważają za lepsze miejsce do życia Nowy Jork czy Kalifornię?
   a. Korzystając na przykład z *Abstraktu Statystycznego dla Stanów Zjednoczonych* lub innej tego typu publikacji, porównaj stopę imigracji i emigracji dla każdego z tych

stanów. Sprawdź, czy są grupy migrujące bezpośrednio z Nowego Jorku do Kalifornii lub na odwrót.

b. Ogólnokrajowe ośrodki badania opinii publicznej – na przykład Gallup, Harris, Roper i inne – często zadają ludziom pytanie, w jakim stanie żyje się ich zdaniem najlepiej. Sprawdź w bibliotece albo w lokalnej gazecie kilka ostatnich wyników takich sondaży.

c. Porównaj liczbę samobójstw w obu stanach.

4. Czy na twojej uczelni najpopularniejszymi wykładowcami są pracownicy wydziału nauk społecznych, przyrodniczych czy humanistycznych?

a. Jeśli twoja uczelnia prowadzi ocenę prowadzących zajęcia przez studentów, sprawdź kilka ostatnich wyników i oblicz średnią dla każdej z trzech grup.

b. Chodź na zajęcia dla każdego z kierunków w trzech grupach i obserwuj frekwencję.

c. W grudniu wybierz grupę profesorów każdego z trzech kierunków i poproś ich, by notowali, ile świątecznych kartek i prezentów otrzymali od swoich studentów wielbicieli. Sprawdź, kto wygrał.

Powyższe przykłady mają stanowić nie tyle sugestię wartych uwagi projektów badawczych, ile ilustrację wielości możliwych sposobów operacjonalizacji zmiennych.

Operacjonalizacja trwa

Mimo że omawiałem dotąd operacjonalizację i konceptualizację jako czynności poprzedzające zbieranie danych i analizę – musisz na przykład ułożyć pytania kwestionariusza, zanim go rozdasz badanym – oba te procesy trwają również w toku procesu badawczego, nawet jeśli dane zbierane są w ustrukturyzowanym, masowym sondażu. Jak już widzieliśmy, w badaniach o mniej sztywnej strukturze, na przykład w badaniach terenowych, określenie i uszczegółowienie odpowiednich pojęć nie dają się oddzielić od toczącego się procesu obserwacji.

Jako badacz musisz zawsze pozostać otwarty na możliwość przeformułowania twoich pojęć i definicji. Ostatecznym celem badań społecznych jest wyjaśnienie natury życia społecznego. Waga i użyteczność tego, czego się dowiesz na ten temat, nie zależy od momentu, w którym pierwszy raz zrozumiałeś, jak patrzeć na rzeczywistość, podobnie jak nie zależy od tego, czy twój pomysł pochodzi z wyuczonego na pamięć podręcznika, ze snu czy też od szwagra.

Kryteria jakości pomiaru

Zaszliśmy już dość daleko. Ten rozdział zaczął się od ogólnego stwierdzenia, że socjologia może mierzyć wszystko, co istnieje. Potem odkryliśmy, że mierzone przez nas przedmioty w większości nie istnieją. Następnie dowiedzieliśmy się, że mimo to można je jednak mierzyć. Teraz kończymy ten rozdział omówieniem kilku standardów, zgodnie z którymi oceniamy nasz względny sukces lub porażkę w pomiarze przedmiotów – nawet takich, których nie ma.

Precyzja i poprawność

Zacznijmy od tego, że pomiarów można dokonywać z różnym stopniem precyzji. Jak już widzieliśmy przy okazji operacjonalizacji, precyzja dotyczy subtelności rozróżnień odrębnych wartości zmiennej. Opis kobiety jako „43-letniej" jest bardziej precyzyjny niż jako „po czterdziestce". Stwierdzenie, że gang uliczny powstał w lecie 1996 roku, jest bardziej precyzyjne niż stwierdzenie, że było to „w latach 90."

Pomiary precyzyjne mają w zasadzie przewagę nad nieprecyzyjnymi, zgodnie ze wskazaniami zdrowego rozsądku. Nie ma okoliczności, w których pomiar nieprecyzyjny miałby zasadniczą wyższość nad precyzyjnym. Mimo to jednak stuprocentowa precyzja nie zawsze okazuje się konieczna czy pożądana. Jeśli informacja, że kobieta jest po czterdziestce, wystarcza na potrzeby twoich badań, każdy dodatkowy wysiłek zmierzający do ustalenia jej dokładnego wieku należy uznać za zmarnowany. Przy

operacjonalizacji pojęć musisz więc częściowo kierować się wiedzą o wymaganym poziomie precyzji. Jeśli twoje potrzeby nie są jasne, staraj się być raczej bardziej niż mniej precyzyjny.

Nie myl jednak precyzji z poprawnością. Opis osoby jako „urodzonej w Nowej Anglii" jest mniej precyzyjny niż jako „urodzonej w Stowe w Vermont" – co jednak, jeśli osoba ta urodziła się w rzeczywistości w Bostonie? Opis mniej precyzyjny jest w tym przypadku bardziej poprawny – lepiej bowiem odzwierciedla rzeczywistość.

Precyzja i poprawność są oczywiście istotnymi cechami pomiaru badawczego i zapewne nie ma potrzeby ich dłużej wyjaśniać. Badacze społeczni, tworząc i oceniając pomiary, zwracają jednak szczególną uwagę na dwa aspekty techniczne: rzetelność i trafność.

Rzetelność

W ujęciu abstrakcyjnym **rzetelność** polega na tym, że dana technika, stosowana do tego samego przedmiotu, daje za każdym razem ten sam wynik. Powiedzmy, że chciałbyś wiedzieć, ile ważę. (Nie, nie wiem dlaczego). Jedną z technik byłoby zadawanie dwóm różnym osobom pytania o moją szacunkową wagę. Jeśli jedna z nich oceni moją wagę na 70, druga zaś na 140 kilogramów, będziemy musieli uznać, że technika opierająca się na ocenie mojej wagi przez innych ludzi nie jest szczególnie rzetelna.

Załóżmy jednak, że jako technikę pomiarową stosujesz wagę łazienkową. Wchodzę na wagę dwa razy, ty zaś za każdym razem notujesz wyniki. Byłyby one najprawdopodobniej identyczne, co dowodziłoby, że waga łazienkowa jest rzetelniejszą techniką pomiaru wagi danej osoby niż szacunkowa ocena dokonywana przez innych ludzi.

Rzetelność nie zapewnia jednak, podobnie jak precyzja, poprawności wyniku. Załóżmy, że przestawiłem moją wagę tak, żeby odjąć sobie dwa czy trzy kilogramy dla poprawy samopoczucia. Mimo że odnotowałbyś za każdym razem (rzetelnie) taką samą wagę, za każdym razem byłbyś w błędzie. Ten nowy element, zwany *obciążeniem*, omawiam w rozdziale 8. Na razie wystarczy nam ostrzeżenie, że rzetelność nie gwarantuje poprawności.

Załóżmy, że chcemy badać nastroje wśród robotników fabrycznych w dwóch różnych typach fabryk. W jednym zbiorze fabryk robotnicy wykonują prace wyspecjalizowane, co stanowi konsekwencję bardzo zaawansowanego podziału pracy. Każdy z robotników przyczynia się w bardzo niewielkim stopniu do ogólnego procesu produkcji taśmowej. W drugim zbiorze fabryk każdy z robotników ma wiele zadań, przy czym niewielkie zespoły robotników wykonują cały proces produkcji.

Jak powinniśmy mierzyć nastroje? Jedna ze strategii mogłaby polegać na obserwacji robotników w każdej z fabryk, ze szczególnym uwzględnieniem takich elementów, jak żarty, częsty śmiech, uśmiechy itp. Moglibyśmy też pytać, jak się im podoba ich praca, a nawet, czy zamieniliby ją na pracę drugiego badanego przez nas typu. Porównując obserwacje w różnych fabrykach, moglibyśmy dojść do wniosku co do tego, która z form organizacji pracy lepiej wpływa na nastroje wśród robotników.

Zajmijmy się teraz problemem rzetelności tej metody. Po pierwsze, nasze samopoczucie podczas obserwacji zabarwi na pewno to, co widzimy. Możemy dokonać błędnej interpretacji. Możemy zobaczyć robotników żartujących ze sobą i pomyśleć, że się kłócą. Możemy trafić na ich słaby dzień. Obserwując tę samą grupę robotników przez kilka dni z rzędu, moglibyśmy codziennie dochodzić do innego wniosku. Kilku obserwatorów, oceniających te same zachowania, mogłoby również dojść do różnych wniosków co do nastrojów wśród robotników.

A oto inna możliwa strategia oceny nastrojów. Załóżmy, że sprawdzamy dokumenty fabryki, by dowiedzieć się, jak wiele skarg złożyły w danym okresie związki zawodowe. Przypuszczalnie byłby to jakiś wskaźnik nastrojów: im więcej skarg, tym gorsze nastroje. Tego typu strategia pomiaru wydaje się bardziej rzetelna: licząc skargi kilka razy od nowa, wciąż otrzymywalibyśmy ten sam wynik.

Jeśli zastanawiasz się w tej chwili nad tym, czy liczba skarg jest naprawdę odzwierciedleniem nastrojów, martwisz się nie o rzetelność, lecz o trafność, o której za chwilę. Na razie cho-

dzi nam o to, że ta ostatnia metoda przypomina moją wagę łazienkową: daje spójne wyniki.

W badaniach społecznych kwestia rzetelności pojawia się w wielu formach. Rzetelność staje się problemem zawsze, gdy źródłem danych jest pojedynczy obserwator, ponieważ nie ma pewnego zabezpieczenia przed wpływem subiektywnego nastawienia tego obserwatora. Nie jesteśmy w stanie określić z pewnością, na ile otrzymane dane pochodzą z obserwowanych zjawisk, na ile zaś od samego obserwatora.

Subiektywność nie jest jednak jedynie problemem pojedynczego obserwatora. Badacze sondażowi wiedzą już od dawna, że różni ankieterzy otrzymują różne odpowiedzi od badanych, zależnie od własnych postaw i zachowań. Albo gdybyśmy mieli zbadać stanowiska redakcji różnych gazet w jakiejś kwestii publicznej, moglibyśmy zatrudnić zespół koderów, którzy zabraliby się za czytanie setek artykułów redakcyjnych i klasyfikowanie w zależności od stanowiska zajętego w tej kwestii. Niestety jednak różni koderzy będą różnie klasyfikować ten sam artykuł redakcyjny. Moglibyśmy też postanowić, że dokonamy klasyfikacji kilkuset zawodów według jakiegoś standardowego klucza kodowego, na przykład zbioru kategorii stworzonego przez Departament Pracy czy przez Biuro Spisowe. Ty i ja nie umieścilibyśmy jednak wszystkich tych zawodów w takich samych kategoriach.

Każdy z tych przykładów jest ilustracją problemu rzetelności. Podobne trudności pojawiają się, gdy staramy się uzyskać od ludzi informacje o nich samych. Zadajemy czasem pytania, na które ludzie nie znają odpowiedzi: ile razy byłeś w kościele? Czasem pytamy ludzi o coś, co ich zdaniem ich nie dotyczy: czy jesteś zadowolony z aktualnego stanu stosunków dyplomatycznych między Chinami a Albanią? W takich przypadkach ludzie będą za każdym razem odpowiadać różnie, ponieważ wymyślają odpowiedzi na bieżąco. Czasem badany temat bywa tak skomplikowany, że osoba, która miała w tej sprawie jasne zdanie, może, gdy zadamy jej dane pytanie po raz drugi, dokonać zupełnie innej interpretacji problemu.

Jak więc tworzyć rzetelne mierniki? Jeśli projekt badawczy wymaga zadawania ludziom pytań, musisz uważać, by zadawać wyłącznie pytania, na które respondenci prawdopodobnie znają odpowiedź. Pytaj o rzeczy, które ich dotyczą, jasno precyzując, o co właściwie pytasz. Wszystko to nie rozwiąże oczywiście każdego możliwego problemu z rzetelnością. Na szczęście jednak badacze społeczni wypracowali liczne techniki krzyżowego badania rzetelności tworzonych mierników.

Metoda testu powtórnego

Czasem dobrze jest przeprowadzić ten sam pomiar więcej niż raz. Taką technikę nazywamy metodą testu powtórnego. Jeśli nie masz powodów, aby oczekiwać zmiany poszukiwanych informacji, powinieneś otrzymać za każdym razem tę samą odpowiedź. Jeśli odpowiedzi są różne, twoja metoda może być nierzetelna – w takim stopniu, w jakim odpowiedzi są różne. Oto przykład.

W badaniach dotyczących oceny zagrożenia zdrowia (*Health Hazard Appraisal* – HHA), będącej elementem profilaktyki, Jeffrey Sacks, W. Mark Krushat i Jeffrey Newman (1980) postanowili określić ryzyko związane z różnymi czynnikami dotyczącymi otoczenia i stylu życia, co ułatwiłoby lekarzom należyte doradzanie pacjentom. Znając sytuację życiową pacjenta, lekarze mogliby powiedzieć mu, jakie ma szanse dalszego życia, oraz poradzić, jak może je poprawić. Osiągnięcie tego celu zależało oczywiście w znacznym stopniu od porawności informacji zebranych o każdym badanym.

Aby sprawdzić rzetelność swoich informacji, Sacks i jego współpracownicy poprosili wszystkich 207 badanych o wypełnienie podstawowego kwestionariusza dotyczącego ich cech i zachowań. Trzy miesiące później w kolejnym kwestionariuszu zadano badanym te same pytania i porównano wyniki. Okazało się, że ogółem tylko 15% badanych podało w obu badaniach te same informacje.

Sacks i jego koledzy piszą:

> „Prawie 10% badanych w drugim badaniu podało inny wzrost. Wiek rodziców zmienił się u niemal co trzeciego badanego – ojciec jedne-

go z badanych podobno postarzał się w ciągu trzech miesięcy o ponad 20 lat. Jeden na pięciu byłych palaczy lub alkoholików miał wyraźne problemy z dokładnym przypomnieniem sobie swoich dawnych nawyków konsumpcyjnych".

(1980, s. 730)

Niektórzy badani wymazali wszelkie ślady deklarowanych poprzednio szmerów serca, cukrzycy, rozedmy płuc, aresztowań czy myśli samobójczych. Matka jednego z badanych zmarła zgodnie z pierwszym kwestionariuszem, wedle drugiego najwyraźniej żyła i cieszyła się dobrym zdrowiem. Jedna z badanych w pierwszym badaniu nie miała jajnika, w drugim badaniu natomiast miała go. W innym przypadku jajnik w pierwszym badaniu był, w drugim natomiast okazało się, że nie ma go już od dziesięciu lat. Jedna z badanych zadeklarowała w pierwszym badaniu wiek 55 lat, trzy miesiące później zaś już 50. (Należałoby się zastanowić, czy lekarze mogliby mieć na pacjentów tak wielki wpływ, jak ich własna pamięć). Tak więc metodą testu powtórnego stwierdzono, że przyjęta metoda zbierania danych nie była zbyt rzetelna.

Metoda połówkowa

Co do zasady, dobrze jest robić więcej niż jeden pomiar każdego skomplikowanego lub złożonego pojęcia społecznego, takiego jak uprzedzenia, alienacja czy klasa społeczna. Procedura ta daje podstawy dla innego sprawdzianu rzetelności. Załóżmy, że stworzyłeś kwestionariusz złożony z dziesięciu pytań, które w twoim przekonaniu mierzą uprzedzenia względem kobiet. Stosując metodę połówkową, mógłbyś losowo podzielić te dziesięć pytań na dwa zbiory pięcioelementowe. Jak widzieliśmy w rozważaniach na temat „wzajemnej wymienialności wskaźników" Lazarsfelda, każdy zbiór powinien stanowić dobry miernik uprzedzeń wobec kobiet, a oba zbiory powinny prowadzić do takiej samej klasyfikacji respondenta. Gdyby klasyfikacja ludzi wypadła różnie według każdego ze zbiorów pytań, najprawdopodobniej masz problem z rzetelnością pomiaru tej zmiennej.

Stosowanie ustalonych mierników

Innym sposobem zapewnienia rzetelności informacji otrzymywanych od ludzi jest stosowanie mierników, których rzetelność została udowodniona w poprzednich badaniach. Mierząc na przykład anomię, mógłbyś pójść śladami Srole'a.

Jednak powszechne stosowanie miernika nie daje gwarancji jego rzetelności. Na przykład Testy Umiejętności Szkolnych (*Scholastic Aptitude Tests*) oraz Wielofazowa Skala Osobowości Minnesota (*Minnesota Multiphasic Personality Inventory – MMPI*) przyjęły się w ciągu dziesięcioleci jako ustalone standardy w swoich dziedzinach. Jednak w ostatnich latach potrzebne okazały się ich fundamentalne zmiany, które pozwoliły uwzględnić przemiany społeczne oraz wyeliminować przestarzałe tematy i seksistowskie obciążenie sformułowań.

Rzetelność osób zatrudnionych przy realizacji badań

Jak już wyżej wspomniałem, pomiar może się okazać nierzetelny także z powodu osób zatrudnionych przy realizacji badań, takich jak ankieterzy czy koderzy. Sprawdzenia rzetelności można w tym przypadku dokonać na kilka sposobów. Powszechną praktyką zabezpieczania się przed nierzetelnością ankieterów w badaniach sondażowych jest zatrudnianie kontrolera, który przeprowadza rozmowy telefoniczne z respondentami z wybranej podpróby i weryfikuje w ten sposób niektóre informacje.

W innych sytuacjach przydatne może się okazać powtórzenie badań. Jeśli sądzisz, że artykuły redakcyjne lub zawody mogą nie zostać rzetelnie sklasyfikowane, możesz kazać je zakodować niezależnie od siebie różnym koderom. Przypadki zakodowane w sposób niespójny można wówczas ocenić bardziej dokładnie i rozwiązać problemy.

Jasność, precyzja i praktyka mogą wreszcie oszczędzić ci wielu problemów z rzetelnością. Gdybyśmy poświęcili trochę czasu na porozumienie się co do sposobu oceniania stanowiska redakcji w jakiejś sprawie – omawiając rozmaite możliwości i czytając razem niektóre arty-

kuły – udałoby się nam prawdopodobnie zaklasyfikować je niezależnie w ten sam sposób.

Rzetelność pomiaru to kwestia fundamentalna dla każdych badań społecznych. Będziemy do niej wracać niejednokrotnie w następnych rozdziałach. Tymczasem jednak postaraj się zapamiętać, że nawet pełna rzetelność nie daje gwarancji, że nasze pomiary dotyczą właśnie tego, co sądzimy, że jest mierzone. Zajmijmy się teraz kwestią trafności.

Trafność

W zwyczajowym ujęciu **trafność** odnosi się do zakresu, w jakim miernik empiryczny adekwatnie odzwierciedla prawdziwe znaczenie danego pojęcia. Ha! Przecież zgodziliśmy się już co do tego, że pojęcia nie mają prawdziwego znaczenia. Jak więc możemy określić, czy dany miernik adekwatnie odzwierciedla znaczenie pojęcia? Oczywiście, ostatecznie nie możemy. Zarazem jednak, jak już widzieliśmy, całość zjawisk społecznych, włącznie z badaniami społecznymi, bazuje na porozumieniu co do znaczeń używanych terminów i odpowiadających im pojęć. Istnieje kilka kryteriów udanego pomiaru odpowiadającego uzgodnionym znaczeniom pojęć.

Po pierwsze, istnieje tak zwana **trafność fasadowa**. Konkretny miernik empiryczny może się zgadzać albo nie zgadzać z naszymi uzgodnieniami i indywidualnymi obrazami mentalnymi dotyczącymi danego pojęcia. Ty i ja moglibyśmy się na przykład spierać co do adekwatności pomiaru nastrojów wśród robotników na podstawie liczby skarg związków zawodowych. Zgodzilibyśmy się jednak z pewnością, że skargi te mają *coś* wspólnego z nastrojami wśród robotników. Oznacza to, że miernik ten jest trafny „co do fasady", niezależnie od tego, czy jest zadowalający. Gdybym chciał mierzyć nastroje wśród robotników liczbą książek wypożyczanych przez nich z biblioteki w czasie wolnym, twoje obiekcje byłyby niewątpliwie poważniejsze: ten miernik nie miałby wiele wspólnego z trafnością fasadową.

Po drugie, jak już wskazywałem wcześniej, istnieje wiele bardziej formalnie ustalonych konwencji co do znaczenia pojęć. Biuro Spisowe stworzyło na przykład definicje operacyjne pojęć takich jak rodzina, gospodarstwo domowe czy status zatrudnienia, które wydają się praktycznie trafne w większości badań wykorzystujących te pojęcia.

Trzy inne typy trafności określają także konkretne sposoby sprawdzania trafności mierników. Pierwszy z nich, **trafność kryterialna**, zwana niekiedy *trafnością predykcyjną*, opiera się na jakimś kryterium zewnętrznym. Na przykład trafność egzaminów na studia przejawia się w możliwości przewidywania na tej podstawie przyszłych sukcesów studenta. Trafność pisemnych testów na prawo jazdy jest w tym sensie określona przez zależność między wynikami zdających a ich późniejszymi doświadczeniami jako kierowców. W tych dwóch przykładach sukcesy studenta i umiejętności kierowcy są owymi kryteriami.

Aby przekonać się, na ile zrozumiałeś ideę trafności kryterialnej, sprawdź, czy potrafisz wymyślić zachowania, które mogłyby służyć ocenie trafności następujących postaw:

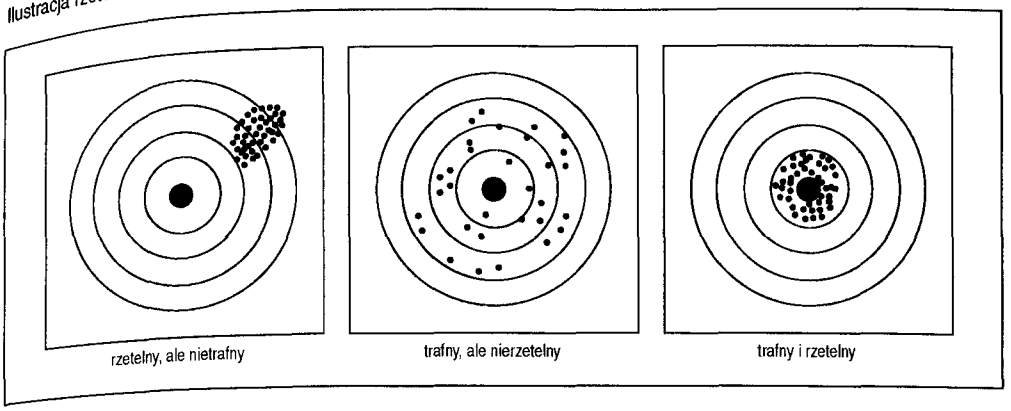
Jest bardzo religijny
Popiera równouprawnienie kobiet i mężczyzn
Popiera skrajnie prawicowe ugrupowania paramilitarne
Troszczy się o środowisko naturalne

Wśród możliwych kryteriów mogłyby się znaleźć na przykład: uczęszcza do kościoła, głosuje na kobiety, należy do NRA, należy do Klubu Sierra. Niekiedy niełatwo znaleźć równie bezpośrednie kryteria behawioralne trafności. W takim przypadku można jednak przybliżyć się do takich kryteriów, stosując inny test. Możemy zastanowić się, jak dana zmienna powinna być, teoretycznie, powiązana z innymi zmiennymi. **Trafność teoretyczna** opiera się na logicznych powiązaniach między zmiennymi.

Załóżmy na przykład, że zamierzasz badać źródła i konsekwencje satysfakcji z pożycia małżeńskiego. W ramach badań opracowujesz miernik owej satysfakcji, po czym sprawdzasz jego trafność.

Oprócz miernika wypracowałeś również pewne przewidywania teoretyczne co do relacji między satysfakcją z pożycia małżeńskiego

Rysunek 5.2
Ilustracja rzetelności i trafności



a innymi zmiennymi. Mógłbyś na przykład dojść do wniosku, że zadowoleni mężowie i żony rzadziej niż niezadowoleni oszukują swoich współmałżonków. Jeśli twój miernik pozostaje w przewidywanej relacji do wierności małżeńskiej, stanowi to dowód jego trafności teoretycznej. Jeśli jednak małżonkowie zadowoleni oszukują równie często jak niezadowoleni, podważałoby to trafność twojego miernika.

Testy trafności teoretycznej mogą więc dostarczyć ważkich oznak tego, że twój miernik dotyczy – bądź nie dotyczy – własności, która ma być mierzona, nie zapewniając jednak ostatecznego na to dowodu. Mimo że twierdziłem, że testy trafności teoretycznej są mniej godne polecenia niż trafności kryterialnej, można by się spierać, jaki rodzaj testu zapewnia w danych okolicznościach określona zmienna porównawcza (doświadczenia kierowców, wierność małżeńska). Ważne jest nie tyle rozróżnienie tych dwóch typów testów trafności, ile zrozumienie logiki walidacji, która stanowi ich cechę wspólną: jeśli udało nam się zmierzyć jakąś zmienną, nasz miernik powinien pozostawać w jakiejś relacji logicznej do innych mierników.

Trafność treściowa wreszcie oznacza zakres, w jakim miernik obejmuje skalę znaczeń zawartych w pojęciu. Na przykład test umiejętności matematycznych nie może się ograniczać tylko do dodawania, ale musi obejmować także odejmowanie, mnożenie, dzielenie itd. Jeśli mierzymy uprzedzenia, to czy nasz miernik uwzględnia wszystkie typy uprzedzeń, w tym uprzedzenia wobec grup rasowych i etnicznych, mniejszości religijnych, kobiet, osób starszych itd.?

Rysunek 5.2 przedstawia ilustrację graficzną różnicy między trafnością a rzetelnością. Jeśli wyobrazisz sobie pomiar jako coś podobnego do wielokrotnego strzelania do tarczy, stwierdzisz, że rzetelność oznacza „mały rozrzut", niezależnie od miejsca, w które dany strzał trafia, ponieważ rzetelność jest funkcją spójności. Trafność z kolei to funkcja rozmieszczenia strzałów wokół celu. Brak rzetelności oznacza losowy rozrzut trafień wokół celu; brak trafności obrazuje rysunek z trafieniami systematycznie poza celem. Zauważ, że zarówno miernik nierzetelny, jak i miernik nietrafny nie będą prawdopodobnie zbyt użyteczne.

Kto decyduje, co jest trafne?

Nasze rozważania na temat trafności rozpocząłem od przypomnienia faktu, że nasze ustalenia co do tego, co jest rzeczywiste, zależą od uzgodnień społecznych. Zapoznaliśmy się też z kilkoma sposobami, jakie stosują przedstawiciele nauk społecznych, by uzgodnić między sobą, że ich pomiary są trafne. Na trafność można jednak spojrzeć w zupełnie inny sposób.

Badacze społeczni krytykują niekiedy siebie samych i innych za przyjmowanie nieświadomego założenia, że badacz jest kimś nadrzędnym w stosunku do badanego. Badacze usiłują na przykład często odkryć motywacje, których sami aktorzy społeczni są nieświadomi. Myślisz, że kupiłeś nowe Burpo-Blasto, bo ma dobre parametry i świetnie wygląda, ale *my* wiemy, że tak naprawdę starasz się podnieść swój status społeczny.

To ukryte poczucie wyższości doskonale pasuje do podejścia skrajnie pozytywistycznego (biolog odczuwa wyższość wobec żaby na stole laboratoryjnym), kłóci się jednak z podejściem bardziej humanistycznym i na ogół jakościowym, reprezentowanym przez wielu przedstawicieli nauk społecznych. Zajmiemy się dokładniej tym tematem w rozdziale 10.

Próbując zrozumieć, w jaki sposób zwykli ludzie interpretują swój świat, etnometodologowie starali się skłonić wszystkich badaczy społecznych, by poświęcili więcej uwagi naturalnym procesom społecznym konceptualizacji i podzielania znaczeń. W końcu zachowanie, które może się wydawać irracjonalne w paradygmacie naukowym, może mieć logiczny sens z punktu widzenia paradygmatu aktora.

Ostatecznie badacze społeczni powinni traktować zarówno swoich kolegów, jak i badanych jako źródła uzgodnień dotyczących najbardziej użytecznych znaczeń i pomiarów badanych przez siebie pojęć. Niekiedy jedno, innym zaś razem drugie z tych źródeł okaże się bardziej przydatne. Żadnego z nich nie należy jednak lekceważyć.

Napięcie między rzetelnością a trafnością

Oczywiste jest, że chcemy, by nasze mierniki były zarazem rzetelne i trafne. Często jednak powstaje napięcie między kryteriami rzetelności i trafności, co wymusza dokonywanie wyboru między nimi.

Przypomnij sobie przykład z mierzeniem nastrojów wśród robotników w różnych fabrykach. Strategia zanurzania się w codzienną rutynę taśmy produkcyjnej, obserwacji zdarzeń i rozmów z robotnikami zapewniłaby na pierwszy rzut oka bardziej trafny miernik owych nastrojów, niż liczenie skarg. Wydaje się oczywiste, że przy użyciu pierwszej metody uzyskalibyśmy jaśniejsze pojęcie o tym, czy nastroje są dobre, czy złe.

Jak już wcześniej wskazywałem, strategia liczenia skarg byłaby jednak rzetelniejsza. Sytuacja ta stanowi przykład bardziej ogólnego problemu pomiarów badawczych. Większość naprawdę interesujących pojęć, które chcielibyśmy badać, ma wiele subtelnych niuansów znaczeniowych i niełatwo jest dokładnie określić, co przez nie rozumiemy. Badacze używają niekiedy określenia „bogactwo znaczeń" w stosunku do tej cechy pojęć. Mimo że na tematy takie jak na przykład anomia napisano już mnóstwo książek i artykułów, ciągle jeszcze nie wyczerpano znaczenia tego pojęcia.

Niezwykle często wydaje się więc, że określenie rzetelnych definicji operacyjnych i pomiarów pozbawia pojęcia całego bogactwa ich znaczeń. Dobre nastroje to coś znacznie więcej niż brak skarg ze strony związków zawodowych, anomia to coś znacznie więcej niż to, co mierzymy za pomocą pięciu pytań Leo Srole'a. Im więcej jednak zachowamy zmienności i bogactwa danego pojęcia, tym więcej będzie okazji do nieporozumień co do jego zastosowania w konkretnej sytuacji, co obniża rzetelność.

Dylemat ten wyjaśnia do pewnego stopnia trwałość dwóch różnych orientacji w badaniach społecznych: technik ilościowych, nomotetycznych, ustrukturyzowanych, takich jak sondaże i eksperymenty oraz metod jakościowych, idiograficznych, takich jak badania terenowe i studia historyczne. Najogólniej można powiedzieć, że metody pierwszego rodzaju są rzetelniejsze, drugiego zaś trafniejsze.

Skoro zostałeś już ostrzeżony, jesteś też wyposażony w narzędzia walki z tym uporczywie powracającym i nieuniknionym dylematem. Jeśli nie istnieje jasne porozumienie co do sposobu mierzenia jakiegoś pojęcia, należy je mierzyć na kilka sposobów. Jeśli pojęcie ma kilka wymiarów, zmierz je wszystkie. Przede wszystkim jednak pamiętaj, że pojęcie to nie ma żadnego innego znaczenia poza tym, które

ty i ja mu przypisujemy. Jedynym uzasadnieniem nadanego pojęciu znaczenia jest użyteczność. Mierz pojęcia w sposób, który ułatwi nam zrozumienie świata wokół nas.

Główne zagadnienia

- Koncepcje to obrazy mentalne, których używamy jako syntetycznych sposobów gromadzenia obserwacji i doświadczeń, które, jak się zdaje, mają ze sobą coś wspólnego. Odwołując się do tych koncepcji używamy pewnych terminów lub etykiet.
- Pojęcia są konstruktami; odzwierciedlają one uzgodnione znaczenie przypisywane terminom. Nasze pojęcia nie istnieją w świecie rzeczywistym, tak więc nie jest możliwy ich bezpośredni pomiar. Możemy jednak mierzyć zjawiska, które one podsumowują.
- Konceptualizacja to proces określania obserwacji i pomiarów, które nadają pojęciom konkretne znaczenie dla celów badań.
- Konceptualizacja wymaga określenia wskaźników pojęcia i opisania jego wymiarów. Definicje operacyjne precyzują, jak będą mierzone zmienne odnoszące się do danego pojęcia.
- Precyzyjne definicje są bardziej istotne w badaniach opisowych niż w wyjaśniających. Stopień niezbędnej precyzji zależy od typu i celu badań.
- Operacjonalizacja to przedłużenie konceptualizacji, kiedy określa się konkretne procedury, które mają być zastosowane do pomiaru wartości zmiennych.
- Operacjonalizacja wymaga dokonania wielu wzajemnie powiązanych wyborów: określenia zakresu zmienności właściwego dla celów badań, oznaczenia stopnia precyzji pomiaru zmiennych, wyjaśnienia istotnych wymiarów zmiennych, jasnego zdefiniowania wartości zmiennych i zależności między nimi oraz decyzji dotyczących właściwego poziomu pomiaru.
- Badacze muszą wybrać jeden z czterech poziomów pomiaru, które zawierają coraz więcej informacji: nominalny, porządkowy, interwałowy lub ilorazowy. Wybór najodpowiedniejszego poziomu pomiaru zależy od celów pomiaru.
- Określoną zmienną można niekiedy mierzyć na różnych poziomach. W razie wątpliwości badacze powinni wybierać najwyższy możliwy poziom pomiaru danej zmiennej, aby uzyskać maksymalną ilość informacji.
- Operacjonalizacja zaczyna się w fazie planowania badań i trwa przez wszystkie fazy projektu badawczego, włącznie z analizą danych.
- Kryteria jakości pomiaru to między innymi precyzja, poprawność, rzetelność i trafność.
- Rzetelność oznacza otrzymywanie spójnych wyników za pomocą tego samego miernika, trafność zaś oznacza uzyskiwanie wyników, które są adekwatnym odbiciem mierzonego pojęcia.
- Badacze mogą testować i poprawiać rzetelność pomiarów dzięki metodzie testu powtórnego, metodzie połówkowej, dzięki użyciu ustalonych mierników oraz kontroli pracy osób zatrudnionych przy przeprowadzaniu badań.
- Do standardów oceny trafności miernika należą: trafność fasadowa, trafność kryterialna, trafność teoretyczna i trafność treściowa.
- Tworzenie precyzyjnych, rzetelnych mierników wpływa często na zubożenie bogactwa znaczeń naszych ogólnych pojęć. Problemu tego nie da się uniknąć. Najlepszym jego rozwiązaniem jest wykorzystanie kilku różnych mierników, ujmujących różne aspekty danego pojęcia.

Kluczowe pojęcia

konceptualizacja
wskaźnik
wymiar
pomiar nominalny
pomiar porządkowy
pomiar interwałowy
pomiar ilorazowy
rzetelność
trafność
trafność fasadowa
trafność kryterialna
trafność teoretyczna
trafność treściowa

Pytania i ćwiczenia

1. Wybierz jedno z pojęć z dziedziny nauk społecznych, takie jak liberalizm lub alienacja, po czym doprecyzuj to pojęcie na tyle, by mogło być wykorzystane w projekcie badawczym. Upewnij się, że określiłeś wskaźniki, które zastosujesz, jak również wymiary, które chciałbyś uwzględnić lub wykluczyć w trakcie konceptualizacji.
2. Znajdź raport z badań w jakiejś książce lub czasopiśmie. Określ, która zmienna jest kluczowa dla badaczy i opisz, jak zmienna ta została zoperacjonalizowana na potrzeby pomiaru.
3. Do którego poziomu pomiaru – nominalnego, porządkowego, interwałowego czy ilorazowego – odnoszą się następujące zmienne:
   a) rasa (biała, afroamerykańska, azjatycka itd.);
   b) kolejność na mecie w wyścigu (pierwszy, drugi, trzeci itd.);
   c) liczba dzieci w rodzinie;
   d) ludność poszczególnych krajów;
   e) stosunek do energii jądrowej (bardzo przychylny, przychylny, nieprzychylny i bardzo nieprzychylny);
   f) region pochodzenia (północny Wschód, środkowy Zachód itd.);
   g) poglądy polityczne (bardzo liberalne, raczej liberalne, raczej konserwatywne, bardzo konserwatywne).
4. Znajdź w jakiejś gazecie lub czasopiśmie przykład nierzetelnego i/lub nietrafnego pomiaru. Uzasadnij swój wybór.
5. Wejdź na stronę *Badań na temat Holokaustu: uprzedzenia* (http://catalog.socialstudies.com/c/@HG432G1hE1Kc2/Pages/holo.html) i przejrzyj opisane tam materiały. Sporządź listę różnych wymiarów uprzedzeń, które uda ci się w nich znaleźć.

Dodatkowe lektury

Bohrnstedt George W., *Measurement*, w: *Handbook of Survey Research*, Peter H. Rossi, James D. Wright, Andy B. Anderson (red.), Academic Press, New York 1983, s. 70–121. Esej ten prezentuje logiczne i statystyczne podstawy rzetelności i trafności pomiaru.

Grimes Michael D., *Class in Twentieth-Century American Sociology: An Analysis of Theories and Measurement Strategies*, Praeger, New York 1991. Książka zawiera świetne, szerokie omówienie konceptualizacji wraz z prezentacją wielu teoretycznych wizji klasy społecznej i właściwych tym teoriom technik pomiaru.

Miller Delbert, *Handbook of Research Design and Social Measurement*, Sage, Newbury Park 1991. Świetna pozycja bibliograficzna. Książka ta, a zwłaszcza jej część 6, zawiera przykłady i opisy wielu mierników operacyjnych stosowanych we wcześniejszych stadiach rozwoju badań społecznych. W kilku przypadkach podane są zastosowane wówczas formaty kwestionariuszy. Mimo nierównego poziomu przykładów dają one dobre wyobrażenie o wielości możliwych odmian.

Silverman David, *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*, rozdz. 7, Sage, Newbury Park 1993. Rozdział ten traktuje o trafności i rzetelności w kontekście badań jakościowych.

The Language of Social Research, Paul F. Lazarsfeld, Morris Rosenberg (red.), Free Press of Glencoe, New York 1955. Wspaniały i urozmaicony klasyczny zbiór opisów konkretnych pomiarów, dokonanych w przeszłości w badaniach społecznych. W 14 artykułach przedstawiono praktyczne i czytelne opisy faktycznych operacji pomiaru, wykonanych przez badaczy społecznych, a także omówienie ogólnych zagadnień pomiaru o charakterze bardziej pojęciowym.

U.S. Department of Health and Human Services, *Survey Measurement of Drug Use*, Government Printing Office, Washington 1992. Obszerny przegląd technik wykorzystywanych przy pomiarach różnych rodzajów używania narkotyków.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodat-

kowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 6

Indeksy, skale i typologie

Przegląd holograficzny

Badacze muszą często stosować liczne wskaźniki, by zmierzyć daną zmienną w sposób trafny i adekwatny. Indeksy, skale i typologie to praktyczne mierniki złożone z kilku wskaźników zmiennych.

Wprowadzenie

Indeksy a skale

Budowa indeksu
Dobór pytań
Badanie zależności empirycznych
Punktacja indeksu
Braki danych
Walidacja indeksu
Status kobiety: przykład procesu tworzenia indeksu

Budowa skali
Skala dystansu społecznego Bogardusa
Skala Thurstone'a
Skala Likerta
Dyferencjał semantyczny
Skala Guttmana

Typologie

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Jak się przekonaliśmy w rozdziale 5, wiele pojęć w naukach społecznych ma złożone i zróżnicowane znaczenie. Przeprowadzenie pomiaru dotyczącego takich pojęć może być ryzykowne. Przypomnij sobie rozważania na temat trafności treściowej, dotyczące oceny tego, czy właściwie uchwyciliśmy wszystkie wymiary pojęcia w ich zróżnicowaniu.

Aby uzyskać dane uwzględniające różne wymiary pojęcia, musimy zwykle poczynić rozmaite obserwacje zjawisk z nim związanych. Bruce Berg (1989, s. 21) na przykład radzi przeprowadzić wywiady pogłębione w celu przygotowania zasadniczych pytań, „obliczonych na wydobycie konkretnych, pożądanych informacji". Badacz powinien prócz tego przygotować pytania dodatkowe: „pytania z grubsza odpowiadające pewnym pytaniom zasadniczym, lecz nieco inaczej sformułowane".

W badaniach ilościowych również stosuje się wielorakie wskaźniki. Załóżmy, że planujesz sondaż. Mógłbyś co prawda niekiedy stworzyć pojedyncze pytanie, które wyczerpuje całość informacji o danej zmiennej – „Płeć: ☐ mężczyzna ☐ kobieta" to najprostszy przykład. Inne zmienne nie są jednak proste i ich pomiar może wymagać zastosowania kilku pytań kwestionariusza.

Analitycy danych ilościowych wypracowali specjalne techniki służące połączeniu wskaźników w pojedynczy miernik. Ten rozdział omawia budowę dwóch typów złożonych mierników zmiennych: indeksów i skal. Mierniki te najczęściej spotykane są w badaniach sondażowych i w innych metodach ilościowych, mogą być jednak stosowane w każdym typie badań społecznych. Krótki fragment pod koniec tego rozdziału dotyczy typologii, które pojawiają się zarówno w badaniach ilościowych, jak i jakościowych.

Mierniki złożone wykorzystywane są często w badaniach ilościowych, a to z kilku powodów. Po pierwsze, nauki społeczne często badają zmienne, dla których nie ma jasnych i jednoznacznych pojedynczych wskaźników. Dla niektórych zmiennych, takich jak wiek, wystarcza pojedynczy wskaźnik. Możemy określić wiek respondenta, zadając mu proste pytanie: „Ile masz lat?". Możemy też określić nakład gazety patrząc po prostu na liczbę w odpowiedniej rubryce. Jednak w przypadku pojęć złożonych rzadko udaje się badaczom wypracować pojedyncze wskaźniki przed przeprowadzeniem badań. Dotyczy to zwłaszcza postaw i poglądów. Badacz prowadzący sondaż zwykle nie jest w stanie zaproponować pojedynczych pytań, które mierzyłyby należycie poziom uprzedzeń respondenta, jego religijność, poglądy polityczne, alienację itp. Jest znacznie bardziej prawdopodobne, że badacz użyje kilku pytań, z których każde stanowi jakiś wskaźnik zmiennej. Każde z tych pytań z osobna mogłoby okazać się nierzetelne lub nietrafne w odniesieniu do wielu respondentów. Miernik złożony może jednak przezwyciężyć ten problem.

Po drugie, badacze mogą się zdecydować na użycie dość precyzyjnego miernika porządkowego danej zmiennej – na przykład alienacji – grupując przypadki w kilku kategoriach porządkowych, na przykład od bardzo wysokiej do bardzo niskiej. Pojedyncze pytanie może nie mieć dostatecznie wielu kategorii, by zapewniać pożądany zakres zmienności. Indeks lub skala złożona z kilku pytań może jednak obejmować konieczny zakres.

Indeksy i skale są wreszcie wydajnymi narzędziami analizy danych. Jeśli analiza pojedynczej informacji stanowi tylko bardzo ogólny wskaźnik danej zmiennej, to uwzględnienie kilku informacji może być źródłem bardziej całościowych i dokładniejszych wskazań. Na przykład pojedynczy artykuł redakcyjny może być pewnym wskaźnikiem orientacji politycznej danej gazety. Zbadanie kilku artykułów redakcyjnych umożliwiłoby nam prawdopodobnie lepszą ocenę, ale jednoczesne posługiwanie się wieloma danymi mogłoby się okazać bardzo skomplikowane. Indeksy i skale (a zwłaszcza skale) to wydajne narzędzia redukcji danych: pozwalają one podsumować wiele wskaźników w pojedynczym wyniku liczbowym, zachowując niekiedy niemal wszystkie szczegóły każdego ze wskaźników.

Indeksy a skale

Terminy *indeks* i *skala* są na ogół w literaturze z zakresu badań społecznych stosowane nie-

precyzyjnie i zamiennie. Te dwa typy mierników mają pewne cechy wspólne, w tej książce będziemy je jednak rozróżniać.

Po pierwsze, zastanówmy się, co je łączy. Zarówno skale, jak i indeksy są porządkowymi miernikami zmiennych. Jedne i drugie porządkują jednostki analizy w kategoriach konkretnych zmiennych, takich jak religijność, alienacja, status społeczno-ekonomiczny, uprzedzenia lub poziom intelektualny. Wynik danej osoby na przykład na skali lub indeksie religijności stanowi wskazanie co do względnej religijności tej osoby w porównaniu z innymi badanymi.

Co więcej, zarówno skale, jak i indeksy są złożonymi miernikami zmiennych – oznacza to, że pomiar opiera się na więcej niż tylko jednej jednostce danych. Tak więc wynik respondenta na indeksie lub skali religijności będzie uzależniony od jego odpowiedzi na kilka pytań kwestionariusza, z których każde stanowi pewien wskaźnik religijności. Podobnie wynik danej osoby na skali IQ jest pochodną odpowiedzi na wiele pytań testowych. Orientacja polityczna gazety może znaleźć odbicie w wyniku, uzyskanym za pomocą indeksu lub skali odzwierciedlającej politykę redakcji gazety wobec różnych kwestii politycznych.

Pomimo tych cech wspólnych zasadne jest rozróżnianie indeksów i skal. W tej książce rozróżnimy je na podstawie sposobu ustalania wyniku punktowego przy każdej z metod. **Indeks** tworzymy przez proste zsumowanie wyników przypisanych poszczególnym wartościom. Moglibyśmy na przykład mierzyć uprzedzenia dodając liczbę uznawanych przez danego respondenta twierdzeń świadczących o uprzedzeniach. **Skalę** natomiast tworzymy przypisując określony wynik punktowy pewnym układom odpowiedzi, przy założeniu, że niektóre pytania świadczą o relatywnie niższym, inne zaś o wyższym poziomie danej zmiennej. Na przykład akceptacja twierdzenia „kobiety różnią się od mężczyzn" jest w najlepszym razie bardzo słabym dowodem seksizmu w porównaniu z akceptacją twierdzenia „kobiety nie powinny mieć praw wyborczych". Skala wykorzystuje różnice w intensywności wartości tej samej zmiennej, by odnaleźć różne wzorce odpowiedzi.

Na rysunku 6.1 graficznie przedstawiona jest różnica między skalami a indeksami. Załóżmy, że chcemy stworzyć miernik aktywności politycznej, dzieląc ludzi na bardzo aktywnych w życiu politycznym, nie biorących w nim udziału oraz wszystkich innych pomiędzy tymi skrajnościami.

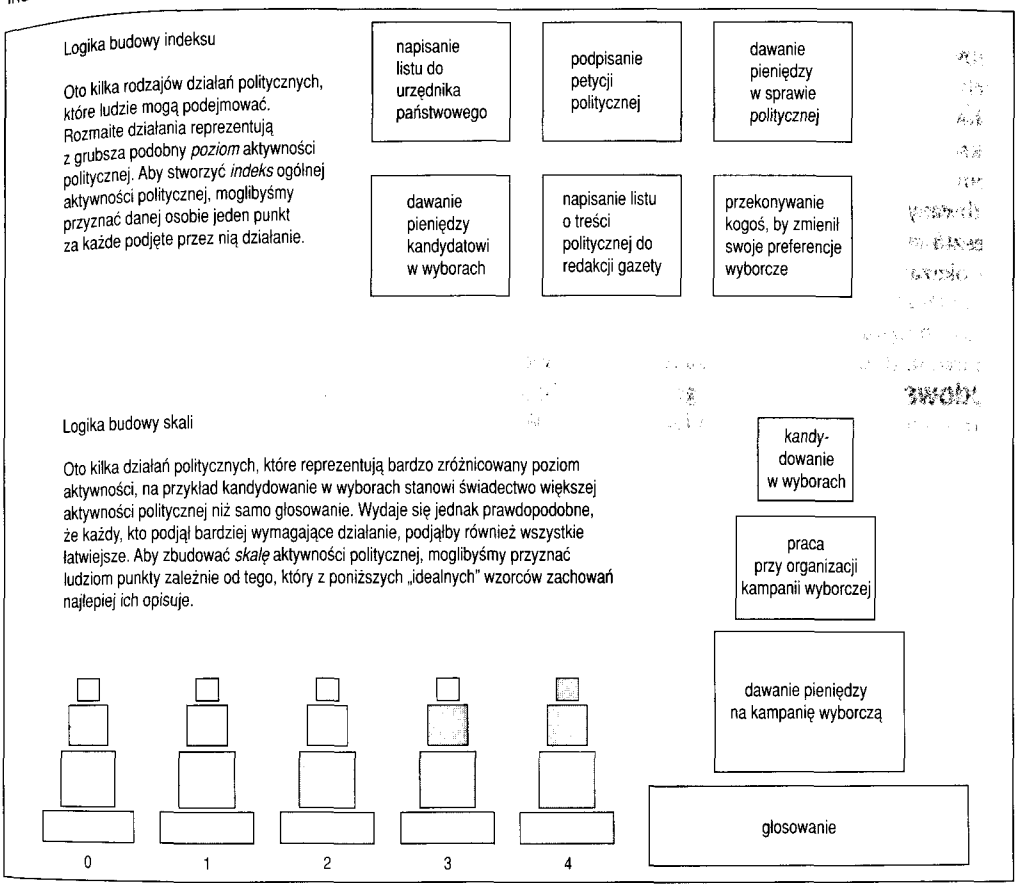
Część górna rysunku 6.1 ilustruje logikę indeksu. Rysunek przedstawia sześć różnych działań politycznych. Moglibyśmy oczywiście spierać się co do pewnych szczegółów, myślę jednak, że zgodzimy się, iż owe sześć form działania stanowi odbicie mniej więcej tego samego stopnia aktywności politycznej.

Za pomocą takich sześciu pytań moglibyśmy zbudować indeks aktywności politycznej, przyznając każdej osobie jeden punkt za każde z podjętych działań. Gdybyś napisał do urzędnika państwowego i podpisał petycję, dostałbyś w sumie dwa punkty. Gdybym dał pieniądze kandydatowi i przekonywał kogoś, by zmienił swoje preferencje wyborcze, uzyskałbym taki sam wynik jak ty. Przy takim podejściu uznalibyśmy, że ty i ja reprezentujemy ten sam poziom aktywności politycznej, mimo że podjęliśmy różne działania.

Część dolna rysunku 6.1 opisuje logikę tworzenia skali. W tym przypadku działania reprezentują wyraźnie różny poziom aktywności politycznej, począwszy od zwykłego głosowania aż po kandydowanie w wyborach. Co więcej, założenie istnienia pewnych prawidłowości działań wydaje się tutaj zupełnie bezpieczne. Na przykład wszyscy, którzy dali pieniądze, najprawdopodobniej również głosowali. Ci, którzy pracowali przy kampanii wyborczej, prawdopodobnie dali jakieś pieniądze i głosowali. Oznacza to, że większość ludzi przejawia jeden z pięciu wyidealizowanych wzorców zachowań przedstawionych w dolnej części rysunku. Rozważania na temat skal w dalszej części tego rozdziału pozwolą nam opisać sposoby przypisania danej osoby do typu, który ją najlepiej charakteryzuje.

Jak się już zapewne zorientowałeś, skale mają generalnie tę wyższość nad indeksami, że uwzględniają intensywność obrazowania danej zmiennej przez różne pytania. Jak widać na rysunku 6.1, wyniki na skali obejmują więcej informacji niż wyniki na indeksie. Musisz jednak

Rysunek 6.1
Indeksy i skale



być świadom, że termin *skala* jest często mylnie stosowany na oznaczenie mierników, które są w istocie tylko indeksami. Fakt, że nazywamy jakiś miernik skalą zamiast indeksem, nie wpływa bynajmniej na jakość tego miernika.

Powinieneś zapoznać się także z dwoma innymi błędnymi wyobrażeniami na temat użycia skal. Po pierwsze, to, czy kombinacja kilku pytań okaże się skalą, niemal zawsze zależy od konkretnej próby objętych badaniem obserwacji. Pewne informacje mogą wchodzić do skali w *jednej* próbie, w *innej* zaś nie. Z tego powodu nie powinieneś zakładać, że dany zestaw informacji stanowi skalę, ponieważ taki był wynik w poprzednich badaniach.

Po drugie, użycie konkretnych technik skalowania – na przykład skali Guttmana, o której powiemy później – nie daje gwarancji stworzenia skali. Techniki te pozwalają raczej ocenić, czy dany zestaw jednostek informacji stanowi skalę.

Przegląd faktycznych raportów z badań z zakresu nauk społecznych pozwala stwierdzić, że badacze znacznie częściej używają indeksów niż skal. Jak na ironię jednak w literaturze metodologicznej nie znajdujemy właściwie żadnych tekstów na temat tworzenia indeksów, choć jest wiele tekstów o tworzeniu skal. Zróżnicowanie to wynika, jak się wydaje, z dwóch przyczyn. Po pierwsze, indeksy są częściej stosowane, ponieważ z danych, które są do dyspozycji, często nie da się stworzyć skali – albo jest to trudne. Po drugie, metody tworzenia indeksów wydają się tak proste i oczywiste, że niewiele się o nich mówi.

Budowa indeksów nie jest jednak przedsięwzięciem prostym. Niepowodzenie wysiłków mających na celu opracowanie technik budowy indeksów przyniosło skutek w postaci wielu kiepskich indeksów w badaniach socjologicznych. Mając to na uwadze, poświęciłem ponad połowę tego rozdziału metodom budowy indeksów. Dysponując gruntownym zrozumieniem logiki tej czynności, będziesz lepiej przygotowany do podjęcia próby utworzenia skali. Zresztą starannie skonstruowany indeks może też okazać się skalą.

Budowa indeksu

Przyjrzyjmy się teraz czterem zasadniczym etapom budowy indeksu: wyborowi możliwych pytań, zbadaniu zależności empirycznych między nimi, ustaleniu punktacji indeksu i jego walidacji. Zakończymy tę część rozważań analizą procesu tworzenia indeksu, który dostarczył ciekawych informacji na temat statusu kobiet w różnych krajach.

Dobór pytań

Pierwszym krokiem przy tworzeniu indeksu jest dobór pytań składających się na złożony indeks, który tworzymy w celu pomiaru jakiejś zmiennej.

Trafność fasadowa

Pierwszym kryterium wyboru pytań do indeku jest trafność fasadowa (albo trafność logiczna). Jeśli chcesz mierzyć na przykład konserwatyzm polityczny, każde z twoich pytań powinno wskazywać na konserwatyzm (albo jego przeciwieństwo, czyli liberalizm). Sympatie partyjne byłyby jednym z tego typu pytań. Innym byłoby pytanie, czy badani zgadzają się, czy nie zgadzają z poglądami jakiegoś znanego polityka konserwatywnego. Tworząc indeks religijności, mógłbyś wziąć pod uwagę takie pytania, jak pytanie o chodzenie do kościoła, o akceptację pewnych wierzeń religijnych, o częstotliwość modlitw. Wydaje się, że każde z tych pytań w jakiś sposób wskazuje na religijność.

Jednowymiarowość

W literaturze metodologicznej na temat konceptualizacji i pomiaru podkreśla się potrzebę *jednowymiarowości* w tworzeniu skal i indeksów. Oznacza to, że złożony miernik powinien odzwierciedlać tylko jeden wymiar danego pojęcia. Nie powinniśmy zatem umieszczać pytań o fundamentalizm religijny w mierniku konserwatyzmu politycznego, mimo możliwej empirycznej zależności między tymi zmiennymi.

Ogólnie czy szczegółowo

Mierniki powinny co prawda trafiać w ten sam wymiar mierzonego pojęcia, może on jednak mieć wiele niuansów. Na przykładzie religijności widać, że wymienione wcześniej wskaźniki – partycypacja, wierzenia itp. – odpowiadają różnym typom religijności. Gdybyś chciał się skoncentrować na uczestnictwie w obrzędach religijnych, wybrałbyś pytania wskazujące na ten właśnie typ religijności: chodzenie do kościoła, komunia, spowiedź itd. Gdybyś chciał mierzyć religijność w znaczeniu bardziej ogólnym, wybrałbyś zrównoważony zbiór pytań, reprezentujących każdy z różnych typów religijności. Ostatecznie to charakter dobranych przez ciebie pytań zadecyduje o tym, jak ogólnie czy jak szczegółowo dana zmienna będzie mierzona.

Zmienność

Wybierając pytania do indeksu, musisz również zwrócić uwagę na obejmowany przez nie zakres zmienności. Jeśli jakieś pytanie ma wskazywać na konserwatyzm polityczny, powinieneś sprawdzić, jaki odsetek respondentów zostanie przez nie zaklasyfikowany jako konserwatyści. Jeśli dane pytanie nie klasyfikuje nikogo jako konserwatysty albo klasyfikuje w ten sposób wszystkich – na przykład jeśli żaden respondent nie zadeklarował poparcia dla jakiegoś skrajnie prawicowego polityka – pytanie to nie na wiele ci się przyda przy tworzeniu indeksu.

Aby zapewnić odpowiedni poziom zróżnicowania, możesz wybrać jedną z dwóch opcji. Po pierwsze, możesz dobrać kilka pytań, według odpowiedzi na które można podzielić respondentów mniej więcej na połowy w kategoriach danej zmiennej, na przykład na konserwatystów i liberałów. Mimo że żadna z odpowiedzi samodzielnie nie daje podstaw, by uznać daną osobę za bardzo konserwatywną, możemy ją jednak w ten sposób zaklasyfikować, jeśli odpowiedziała w sposób konserwatywny na wszystkie pytania.

Druga możliwość polega na wyborze pytań o różnym stopniu różnicowania. Jedno pytanie może klasyfikować jako konserwatystów około połowy badanych, inne zaś zaledwie kilku respondentów. Zauważ, że ta druga możliwość staje się koniecznością w przypadku skalowania, ale jest godna polecenia także w przypadku indeksu.

Badanie zależności empirycznych

Kolejnym krokiem w tworzeniu indeksu jest badanie zależności empirycznych między pytaniami wchodzącymi potencjalnie w jego skład. Wyprzedzamy w tym miejscu rozważania rozdziałów 15 i 16. Zależność empiryczna polega na tym, że odpowiedź respondenta na jedno pytanie – na przykład w kwestionariuszu – pozwala nam przewidzieć jego odpowiedź na inne pytania. Jeśli dwa pytania są ze sobą empirycznie powiązane, możemy zakładać, że wskazują one na tę samą zmienną i możemy włączyć je obie do jednego indeksu. Istnieją dwa typy możliwych zależności między zmiennymi: dwuzmiennowe i wielozmiennowe.

Zależności dwuzmiennowe

Zależność dwuzmiennowa to po prostu zależność między dwiema zmiennymi. Załóżmy, że chcemy mierzyć poparcie respondentów dla członkostwa Stanów Zjednoczonych w ONZ. Jednym ze wskaźników różnych poziomów tego poparcia mogłoby być pytanie: „Czy sądzi Pan(i), że nakłady finansowe USA na działalność ONZ są ☐ zbyt wysokie ☐ właściwe ☐ zbyt niskie?".

Innym wskaźnikiem poparcia dla ONZ mogłoby być pytanie: „Czy Stany Zjednoczone powinny udostępniać swoje siły zbrojne dla celów operacji pokojowych ONZ? ☐ zdecydowanie tak ☐ raczej tak ☐ raczej nie ☐ zdecydowanie nie".

W warstwie fasadowej treść obu tych pytań, jak się wydaje, odzwierciedla różne stopnie poparcia dla ONZ. Niektórzy ludzie mogą być jednak zdania, że USA powinny dawać więcej pieniędzy, ale nie dostarczać wojska. Inni mogą uważać, że należy wysyłać oddziały wojskowe, ale obniżyć nakłady finansowe.

Jeśli jednak oba te pytania odzwierciedlają różne natężenie tej samej postawy, możemy oczekiwać, że udzielone na nie odpowiedzi będą sobie w ogólnych zarysach odpowiadać. W szczególności osoby, które popierają wsparcie militarne dla ONZ, powinny częściej niż wrogowie udostępnienia amerykańskich sił zbrojnych wypowiadać się także za jej wspomaganiem finansowym. I na odwrót: zwolennicy pomocy finansowej powinni częściej niż osoby niechętne wsparciu finansowemu Organizacji popierać użycie amerykańskich wojsk w operacjach ONZ. Jeśli przewidywania te się sprawdzą, możemy powiedzieć, że między obiema zmiennymi istnieje zależność dwuzmiennowa.

Oto kolejny przykład. Załóżmy, że chcemy określić stopień poparcia respondenta dla prawa kobiet do aborcji. Moglibyśmy zapytać: (1) „Czy sądzi Pan(i), że kobieta powinna mieć prawo do aborcji, jeśli ciąża jest wynikiem gwałtu?" i (2) „Czy uważa Pan(i), że kobieta powinna mieć prawo do aborcji, jeśli dalsza ciąża zagrażałaby poważnie jej zdrowiu?".

Z całą pewnością niektórzy respondenci zgodziliby się z pytaniem (1) i nie zgodziliby się z pytaniem (2), inni zaś na odwrót. Jeśli jednak oba pytania dotyczą pewnej ogólnej opinii ludzi na temat aborcji, to odpowiedzi na oba pytania powinny być wzajemnie powiązane. Osoby popierające prawo do aborcji w przypadku gwałtu powinny częściej popierać je też w razie zagrożenia życia kobiety, niż przeciwnicy aborcji w przypadku ciąży będącej wynikiem gwałtu. Mielibyśmy zatem kolejny przykład zależności dwuzmiennowej.

Powinieneś zbadać wszystkie możliwe zależności dwuzmiennowe między pytaniami

mającymi ewentualnie wchodzić w skład indeksu. Pozwala to określić względną siłę związków między poszczególnymi parami pytań. Można w tym celu wykorzystać tabele procentowe lub współczynniki korelacji (patrz rozdział 17) – albo jedne i drugie. Ocena siły zależności może się jednak okazać sprawą dość trudną. Niektóre z tych trudności omówiono w ramce „Wskaźniki «przyczynowe» i «skutkowe»".

Uważaj na pytania, które nie są ze sobą powiązane empirycznie: jest mało prawdopodobne, by mierzyły tę samą zmienną. Powinieneś prawdopodobnie usunąć każde pytanie, które nie jest powiązane z kilkoma innymi pytaniami.

Bardzo silna zależność między dwoma pytaniami może stanowić inny problem. Jeśli dwa pytania są doskonale związane, wystarczy, jeśli tylko jedno z nich zostanie włączone do indeksu. Skoro jedno pytanie w pełni odtwarza odpowiedzi na drugie, włączenie ich obu do indeksu nie wnosi nic nowego. (Wyjaśnię ten problem w następnym podrozdziale).

Oto przykład ilustrujący testowanie zależności dwuzmiennowych w tworzeniu indeksu. Przeprowadziłem kiedyś sondaż wśród profesorów uczelni medycznych, aby sprawdzić wpływ „orientacji naukowej" na jakość opieki sprawowanej przez lekarzy nad pacjentami.

Wskaźniki „przyczynowe" i „skutkowe"

Kenneth Bollen

Wydział Socjologii, Uniwersytet North Carolina, Chapel Hill

Zazwyczaj rozsądnie jest oczekiwać, że wskaźniki tej samej zmiennej będą ze sobą dodatnio powiązane, jak to zostało omówione w tekście. Zdarzają się jednak również inne przypadki.

Wskaźniki powinny być ze sobą powiązane, jeśli są zasadniczo „skutkami" danej zmiennej. Mierząc na przykład samoocenę, moglibyśmy zadać badanym pytanie, czy zgadzają się, czy też nie zgadzają z następującymi stwierdzeniami: (1) „Jestem dobrym człowiekiem" i (2) „Jestem z siebie zadowolony(a)". Osoba o wysokiej samoocenie powinna zgodzić się z oboma stwierdzeniami, podczas gdy ktoś z niską samooceną prawdopodobnie nie zgodziłby się z żadnym z nich. Skoro oba wskaźniki zależą od poziomu samooceny – lub też go „odzwierciedlają", oczekujemy, że będą ze sobą dodatnio skorelowane. Ogólniej rzecz biorąc, wskaźniki zależne od tej samej zmiennej, jeśli są trafne, powinny być wzajemnie powiązane.

Prawidłowość taka nie zachodzi jednak w przypadku wskaźników będących raczej „przyczyną" niż „skutkiem" zmiennej. W takiej sytuacji wskaźniki mogą być skorelowane dodatnio, ujemnie lub mogą w ogóle nie być skorelowane. Na przykład moglibyśmy uznać płeć i rasę za wskaźniki zmiennej *zagrożenie dyskryminacją*. Bycie nie-białym lub kobietą zwiększa prawdopodobieństwo doświadczenia dyskryminacji, tak więc są to dobre wskaźniki naszej zmiennej. Nie przewidywalibyśmy jednak zapewne, że płeć i rasa pojedynczych ludzi są ze sobą silnie powiązane.

Moglibyśmy też mierzyć na przykład *interakcje społeczne* za pomocą trzech wskaźników: czasu spędzanego z przyjaciółmi, czasu spędzanego z rodziną i czasu spędzanego ze współpracownikami. Mimo że każdy ze wskaźników jest trafny, nie muszą one być pozytywnie skorelowane. Czas spędzany z przyjaciółmi może być na przykład ujemnie związany z czasem spędzanym z rodziną. W tym przypadku te trzy wskaźniki są „przyczyną" poziomu interakcji społecznych.

Ostatni przykład: *zagrożenie stresem* można mierzyć tym, czy dana osoba doświadczyła ostatnio rozwodu, śmierci współmałżonka lub utraty pracy. Mimo że każde z tych zdarzeń może wskazywać na stres, nie muszą one być wzajemnie powiązane.

Krótko mówiąc, możemy oczekiwać związku między wskaźnikami, które są zależne od jakiejś zmiennej lub ją „odzwierciedlają", czyli takimi, które są jej „skutkami". Jeśli jednak to dana zmienna jest zależna od wskaźników – jeśli wskaźniki są jej „przyczynami" – wskaźniki te mogą być skorelowane dodatnio lub ujemnie bądź też mogą nawet nie być ze sobą powiązane. Powinniśmy więc określić, czy wskaźniki są przyczynami, czy też skutkami zmiennej, zanim postanowimy ocenić ich trafność za pomocą ich wzajemnych korelacji.

Miałem pierwotnie zamiar stwierdzić, czy lekarze z inklinacjami naukowymi traktują pacjentów bardziej bezosobowo niż inni.

Kwestionariusz sondażu zawierał kilka możliwych wskaźników orientacji naukowej respondenta. Trzy pytania wydawały się szczególnie jasnymi wskaźnikami tego, że dany lekarz jest zorientowany naukowo:

1. Jako profesor uczelni medycznej, w jakiej roli może Pan(i), Pana(i) zdaniem, być lepszym *nauczycielem*: praktykującego lekarza czy badacza w dziedzinie medycyny?
2. Czy sądzi Pan(i), że w miarę pogłębiania Pana(i) wiedzy medycznej, Pana(i) zainteresowania medyczne kierują się raczej w stronę kompleksowego leczenia pacjenta czy też w stronę znajomości podstawowych mechanizmów? [Celem tego pytania było oddzielenie osób zainteresowanych przede wszystkim całościowym leczeniem pacjenta od tych, które interesowały się raczej procesami biologicznymi].
3. W zakresie badań terapeutycznych, czy interesują Pana(ią) bardziej artykuły dotyczące oceny efektywności różnych terapii, czy też te, które analizują uzasadnienie poszczególnych metod leczenia? [W tym przypadku było podobnie – starałem się odróżnić osoby zainteresowane artykułami na temat opieki nad pacjentem od tych, które zajmowały się raczej procesami biologicznymi].

(Babbie 1970, s. 27–31)

W przypadku każdego z powyższych pytań moglibyśmy dojść do wniosku, że respondenci, którzy wybrali drugą odpowiedź, są zorientowani bardziej naukowo niż ci, którzy wybrali odpowiedź pierwszą. Mimo że ten wniosek porównawczy wydaje się logiczny, nie powinniśmy dać się zwieść przekonaniu, że respondenci, którzy wybrali drugą odpowiedź na każde pytanie, są naukowcami w jakimś sensie absolutnym. Są po prostu zorientowani bardziej teoretycznie niż osoby wybierające odpowiedź pierwszą.

By uzmysłowić to sobie należycie, przyjrzyjmy się rozkładowi odpowiedzi na każde pytanie. W kategoriach pierwszego pytania – skuteczności nauczania – tylko jedna trzecia respondentów okazała się zorientowana naukowo. Oznacza to, że w przybliżeniu co trzeci z nich stwierdził, że lepiej naucza jako badacz w dziedzinie medycyny. Odpowiadając na drugie pytanie – kierunek zainteresowań medycznych – około dwóch trzecich badanych wybrało odpowiedź naukową, twierdząc, że bardziej niż całościowe leczenie pacjenta interesuje ich znajomość podstawowych mechanizmów. Przy pytaniu trzecim – o preferencje czytelnicze – około 80% wybrało odpowiedź naukową.

Te trzy pytania kwestionariusza nie pozwalają nam powiedzieć, ilu mamy „naukowców" w próbie, ponieważ żadne z nich nie jest powiązane ze zbiorem kryteriów bycia naukowcem w sensie bezwzględnym. Wykorzystanie tych pytań w takim celu stawia nas przed problemem trzech zupełnie różnych oszacowań liczby naukowców w próbie.

Pytania te stanowią jednak trzy niezależne wskaźniki względnych skłonności respondentów do zajmowania się medycyną jako nauką. Każde z pytań dzieli respondentów na bardziej i mniej naukowo zorientowanych. Jednak każde grupowanie respondentów jako mniej lub bardziej naukowo zorientowanych będzie nieco inne niż pozostałe. Respondenci zorientowani naukowo w świetle odpowiedzi na jedno pytanie nie okażą się tacy w przypadku innego pytania. Niemniej jednak w takim zakresie, w jakim każde z pytań mierzy ten sam ogólny wymiar, powinniśmy odnaleźć jakąś odpowiedniość między tymi trzema grupowaniami. Respondenci, których uznaliśmy za naukowo zorientowanych na podstawie odpowiedzi na jedno pytanie, powinni częściej okazywać się zorientowani naukowo przy odpowiedzi na inne pytanie niż ci, którzy na pierwsze pytanie nie udzielili odpowiedzi naukowej.

Rysunek 6.2 przedstawia powiązania między odpowiedziami na te trzy pytania. Zaprezentowałem tam trzy tabele dwuzmiennowe, ukazujące rozkład odpowiedzi dla każdej z możliwych par pytań. Analiza trzech relacji dwuzmiennowych przedstawionych na rysunku pozwala podtrzymać przypuszczenie, że te trzy pytania mierzą tę samą zmienną: *orientację naukową*. Aby przekonać się, dlaczego, przyjrzyjmy się pierwszej zależności dwuzmiennowej

w tabeli. Tabela wykazuje, że profesorowie, którzy zadeklarowali, że to w roli „badacza" są najbardziej skuteczni jako nauczyciele, częściej też twierdzili, że ich zainteresowania medyczne dotyczą „podstawowych mechanizmów" (87%) niż respondenci, którzy wybrali wariant „lekarz" (51%). Fakt, że „lekarze" podzielili się niemal równo na pół w kwestii zainteresowań medycznych, jest dla naszych celów nieistotny. Ważne jest tylko, że ich zainteresowania medyczne są mniej naukowe niż „badaczy". Siłę tej zależności można podsumować jako różnicę 36 punktów procentowych.

Podobny wniosek ogólny stosuje się do innych relacji dwuzmiennowych. Siłę zależności między preferencjami czytelniczymi a kierunkiem zainteresowań medycznych można określić w skrócie jako różnicę 38 punktów procentowych, natomiast między preferencjami czytelniczymi a rolą w nauczaniu – 21 punktów. Każde pytanie daje zatem w efekcie podział respondentów na zorientowanych „naukowo" i „nienaukowo". Odpowiedzi udzielone na każde z pytań odpowiadają jednak w mniejszym lub większym stopniu odpowiedziom na pozostałe pytania.

Trzy omawiane pytania zostały początkowo wybrane na podstawie ich trafności fasadowej – każde z nich zdawało się w jakiś sposób wskazywać orientację naukową profesora. Badając zależności dwuzmiennowe między parami pytań, znaleźliśmy argumenty na poparcie naszego oczekiwania, że wszystkie pytania mierzą zasadniczo to samo zjawisko. Argumenty te nie wystarczą jednak dla uzasadnienia decyzji o włączeniu tych pytań do złożonego indeksu.

Rysunek 6.2

Zależności dwuzmiennowe między pytaniami dotyczącymi orientacji naukowej

A.

Kierunek zainteresowań medycznych	Skuteczność nauczania: lekarz	Skuteczność nauczania: badacz
całościowe leczenie pacjenta	49%	13%
podstawowe mechanizmy	51%	87%
	100% (268)	100% (159)

B.

Kierunek zainteresowań medycznych	Preferencje czytelnicze: efektywność	Preferencje czytelnicze: uzasadnienie
całościowe leczenie pacjenta	68%	30%
podstawowe mechanizmy	32%	70%
	100% (78)	100% (349)

C.

Skuteczność nauczania	Preferencje czytelnicze: efektywność	Preferencje czytelnicze: uzasadnienie
lekarz	85%	64%
badacz	15%	36%
	100% (78)	100% (349)

Przed ich połączeniem w jeden indeks musimy zbadać zależności wielozmiennowe.

Zależności wielozmiennowe między pytaniami

Zależność dwuzmiennowa dotyczy jednocześnie dwóch zmiennych, zależność wielozmiennowa zachodzi natomiast między więcej niż dwiema zmiennymi. Przypomnij sobie, że zasadniczym celem tworzenia indeksu jest wypracowanie metody klasyfikacji badanych w kategoriach pewnej zmiennej, takiej jak *konserwatyzm polityczny*, *religijność* czy też *orientacja naukowa*. Indeks konserwatyzmu politycznego powinien rozróżniać osoby bardzo konserwatywne, umiarkowanie konserwatywne, niezbyt konserwatywne i zupełnie niekonserwatywne (względnie odpowiednio umiarkowanie liberalne i bardzo liberalne w miejsce dwóch ostatnich kategorii). Gradację zmiennej zapewnia kombinacja odpowiedzi udzielonych na kilka pytań zawartych w indeksie. Tak więc respondent, który okazałby się konserwatywny na podstawie wszystkich pytań, zostałby uznany za ogólnie bardzo konserwatywnego.

Aby indeks mógł zapewnić sensowną gradację w powyższym rozumieniu, każde pytanie musi dodawać coś nowego do oceny respondenta. Pamiętasz zapewne z poprzedniej części rozważań, że dwa doskonale ze sobą związane pytania nie powinny się znaleźć w tym samym indeksie. Skoro już włączymy jedno z nich, drugie nie wnosi nic do naszej oceny respondenta. Badanie zależności wielozmiennowych między pytaniami to inny sposób na wyeliminowanie bezwartościowych danych. Pozwala ono przy tym określić całkowitą moc pomiaru poszukiwanej zmiennej przez dany zbiór pytań.

Zilustrujmy cele analizy wielozmiennowej, wracając do przykładu pomiaru orientacji naukowej wśród profesorów uczelni medycznych. Rysunek 6.3 przedstawia trójzmiennowe zależności między trzema pytaniami.

Respondenci podzieleni są na cztery kategorie według ich (1) skuteczności nauczania i (2) preferencji czytelniczych. Liczby w nawiasach oznaczają liczbę respondentów w każdej z grup. Zatem 66 profesorów, którzy stwierdzili, że uczą najlepiej jako lekarze, zadeklarowało też większe zainteresowanie artykułami na temat efektywności różnych terapii. W każdej z czterech grup liczba oznacza odsetek tych respondentów, którzy twierdzili, że są bardziej zainteresowani znajomością podstawowych mechanizmów. Tak więc na przykład we wzmiankowanej grupie 66 profesorów 27% było zainteresowanych przede wszystkim podstawowymi mechanizmami.

Porządek czterech grup opiera się na uprzednio wyciągniętych wnioskach co do orientacji naukowej. Grupa w lewym górnym rogu tabeli jest przypuszczalnie najmniej naukowo zorientowana, jeśli przyjmiemy jako kryterium skuteczność nauczania i preferencje czytelnicze. Grupa w prawym dolnym rogu jest przypuszczalnie najbardziej naukowo zorientowana – w kategoriach tych pytań.

Przypomnij sobie, że wyrażenie zainteresowania podstawowymi mechanizmami uznaliśmy również za wskaźnik orientacji naukowej. Zgodnie z naszymi oczekiwaniami osoby w prawym dolnym rogu najczęściej udzielały

Rysunek 6.3

Zależności trójzmiennowe między pytaniami dotyczącymi orientacji naukowej

Odsetek zainteresowanych podstawowymi mechanizmami

Preferencje czytelnicze	Skuteczność nauczania: lekarz	Skuteczność nauczania: badacz
efektywność	27% (66)	58% (12)
uzasadnienie	58% (219)	89% (130)

tej odpowiedzi (89%), osoby zaś w lewym górnym rogu – najrzadziej (27%). Respondenci, którzy udzielili mieszanych odpowiedzi na pytania o skuteczność nauczania i preferencje czytelnicze, wykazują pośredni stopień zainteresowania podstawowymi mechanizmami (58% w obu przypadkach).

Omówiona tabela niesie wiele istotnych dla nas informacji. Po pierwsze, możemy zauważyć, że obecność trzeciego pytania nie wpływa istotnie na pierwotne zależności między parami pytań. Przypomnij sobie na przykład, że zależność między skutecznością nauczania a kierunkiem zainteresowań medycznych opisaliśmy jako różnicę 36 punktów procentowych. Na rysunku 6.3 widzimy, że tylko wśród respondentów zainteresowanych artykułami na temat efektywności leczenia zależność między skutecznością nauczania a kierunkiem zainteresowań medycznych sięga 31 punktów procentowych różnicy (58% minus 27%: pierwszy rząd). Podobnie wygląda sytuacja wśród respondentów zainteresowanych artykułami na temat uzasadnienia poszczególnych metod leczenia (89% minus 58%: drugi rząd). Pierwotna zależność między skutecznością nauczania a kierunkiem zainteresowań medycznych jest właściwie taka sama, jak na rysunku 6.2, nawet wśród respondentów ocenionych jako zorientowani naukowo lub nienaukowo w kategoriach ich preferencji czytelniczych.

Ten sam wniosek możemy wyciągnąć z analizy danych zawartych w tabeli na rysunku 6.3. Przypomnij sobie, że zależność początkowa między preferencjami czytelniczymi a kierunkiem zainteresowań medycznych została opisana jako różnica 38 punktów procentowych. Patrząc tylko na „lekarzy" na rysunku 6.3, widzimy, że zależność między pozostałymi dwoma pytaniami wynosi teraz 31 punktów procentowych różnicy. Taką samą zależność odnajdujemy w grupie „badaczy" w drugiej kolumnie.

Waga tych obserwacji staje się jeszcze bardziej oczywista, gdy zastanowimy się, co mogło się stać. Na rysunku 6.4 przedstawiono dane hipotetyczne, których wymowa jest zupełnie inna niż faktycznych danych z rysunku 6.3. Jak widać na rysunku 6.4, początkowa zależność między skutecznością nauczania a kierunkiem zainteresowań medycznych utrzymuje się także po uzupełnieniu obrazu o preferencje czytelnicze. W każdym rzędzie tabeli „badacze" częściej niż „lekarze" wyrażają zainteresowanie znajomością podstawowych mechanizmów. Patrząc jednak na kolumny, widzimy, że nie ma zależności między preferencjami czytelniczymi a kierunkiem zainteresowań medycznych. Jeśli wiemy, czy dany respondent uważa, że jest skuteczniejszym nauczycielem jako badacz, czy jako lekarz, wiedza o jego preferencjach czytelniczych nic nie wnosi do naszej oceny jego orientacji naukowej. Gdyby prawdziwe dane dały wynik podobny do tego na rysunku 6.4, doszlibyśmy do wniosku, że nie powinniśmy włączać preferencji czytelniczych do tego samego indeksu, co skuteczność nauczania, ponieważ nie wnoszą one nic do całego indeksu.

Powyższy przykład dotyczył tylko trzech pytań kwestionariusza. Gdybyśmy wzięli pod uwagę więcej pytań, należałoby stworzyć bardziej skomplikowane tabele wielozmiennowe, uwzględniające cztery, pięć i więcej zmiennych.

Rysunek 6.4

Zależności trójzmiennowe między pytaniami dotyczącymi orientacji naukowej

Odsetek zainteresowanych podstawowymi mechanizmami

Preferencje czytelnicze	Skuteczność nauczania: lekarz	Skuteczność nauczania: badacz
efektywność	51% (66)	87% (12)
uzasadnienie	51% (219)	87% (130)

Celem tego etapu tworzenia indeksu jest wykrycie jednoczesnych wzajemnych oddziaływań między pytaniami, aby określić, które z nich powinny się znaleźć w jednym indeksie.

Punktacja indeksu

Jeśli wybrałeś już najlepsze pytania do twojego indeksu, powinieneś teraz przypisać poszczególnym odpowiedziom wartości punktowe, tworząc w ten sposób jeden złożony miernik z kilku pytań. Na tym etapie musisz podjąć dwie zasadnicze decyzje.

Przede wszystkim musisz określić pożądany zakres punktacji indeksu. Główną zaletą indeksu w porównaniu z pojedynczym pytaniem jest większa gradacja stopni pomiaru zmiennej. Zwracałem już wcześniej uwagę, że konserwatyzm polityczny można mierzyć od poziomu „bardzo konserwatywnego" aż po „zupełnie niekonserwatywny" czy „bardzo liberalny". Gdzie zatem powinny się znaleźć punkty skrajne indeksu?

Podejmując tę decyzję, raz jeszcze natykasz się na kwestię zmienności. W miarę rozszerzania zakresu indeksu niemal zawsze na jego krańcach mieści się coraz mniej przypadków. Badacz, który chciałby mierzyć konserwatyzm polityczny włącznie z jego najbardziej skrajnymi postaciami (parafrazując słynną wypowiedź pewnego polityka, gdzieś na prawo od ściany), mógłby stwierdzić, że w tej kategorii nie ma prawie nikogo. W pewnym punkcie zwiększanie gradacji nie zwiększa znaczenia wyników.

Pierwszego wyboru musisz więc dokonać między sprzecznymi dążeniami do (1) szerokiego zakresu pomiaru indeksu a (2) odpowiednią liczbą przypadków dla każdego punktu indeksu. Będziesz zmuszony wypracować jakiś kompromis między tymi sprzecznymi dążeniami.

Druga decyzja dotyczy faktycznego przypisywania wartości punktowych poszczególnym odpowiedziom. Przede wszystkim musisz zdecydować, czy przyznajesz poszczególnym pytaniom indeksu taką samą czy różną wagę. Nie ma w tym względzie sztywnych reguł, sugeruję jednak – a praktyka zdaje się potwierdzać moje słowa – aby przyznawać poszczególnym pytaniom taką samą wagę, chyba że istnieją naprawdę ważne powody, by postąpić inaczej. Uzasadnienia powinny zatem wymagać różne wagi, wagi równe zaś powinny być normą.

Decyzja w tej sprawie musi być oczywiście powiązana z wcześniejszymi kwestiami dotyczącymi równowagi wybranych pytań. Jeśli indeks ma być kompozycją nieco zróżnicowanych aspektów danej zmiennej, powinieneś nadać każdemu aspektowi tę samą wagę. Niekiedy może się jednak okazać, że dwa pytania odzwierciedlają właściwie ten sam aspekt, trzecie zaś zupełnie inny. Gdybyś chciał, by oba aspekty były w tym samym stopniu uwzględniane przez indeks, mógłbyś przyznać trzeciemu pytaniu wagę równą kombinacji wag dwóch pierwszych pytań, na przykład przypisując trzeciemu maksymalny wynik w wysokości dwóch punktów, dwóm pierwszym zaś maksymalnie po jednym punkcie.

Mimo że uzasadnienie punktacji odpowiedzi powinno uwzględniać powyższe i podobne rozważania, badacze na ogół eksperymentują z różnymi metodami punktowania, badając względne wagi przyznawane różnym aspektom, ale nie zapominając zarazem o zasięgu i rozkładzie przypadków. Wybrana w końcu metoda będzie wyrazem kompromisu między kilkoma dążeniami. Oczywiście tak jak w przypadku większości czynności badawczych, decyzja taka może zostać zweryfikowana na podstawie dalszych badań. Ocena trafności indeksu, o której wkrótce, może zmusić badacza do powtórnego wysiłku w celu stworzenia zupełnie innego indeksu.

W przykładzie zaczerpniętym z sondażu wśród profesorów uczelni medycznych postanowiłem przyznać pytaniom taką samą wagę, ponieważ wybrałem je między innymi dlatego, że odzwierciedlają nieco odmienne aspekty ogólnej zmiennej *orientacja naukowa*. W każdym pytaniu respondentom przyznawano 1 punkt za odpowiedź „naukową" i 0 za „nienaukową". Każdy respondent mógł zatem otrzymać łącznie 0, 1, 2 lub 3 punkty. Ta metoda punktacji zapewniła zakres zmienności, który uznałem za użyteczny – cztery kategorie indeksowe – a zarazem zapewniała dostatecznie wiele przypadków do analizy w każdej kategorii.

Jaka jest najlepsza uczelnia w USA

Czasopismo „U.S. News and World Report" zamieszcza co roku specjalny raport z rankingiem wyższych uczelni z całego kraju. Ranking ten powstaje na podstawie indeksu, uwzględniającego kilka pozycji: wydatki na kształcenie na głowę studenta, odsetek magistrów wśród studentów, selekcję wstępną (odsetek przyjętych kandydatów), przeciętne wyniki naukowe studentów pierwszego roku uzyskane w teście umiejętności szkolnych i inne podobne wskaźniki jakości.

Harvard zajmuje na ogół pierwsze miejsce w rankingu, za nim plasują się Yale i Princeton. Jednak wyniki rankingu „Najlepsza Uczelnia Ameryki" z 1999 roku zaskoczyły nauczycieli, przyszłych studentów i ich rodziców. California Institute of Technology skoczył z dziewiątego miejsca w 1998 roku na pierwsze rok później. Harvard, Yale i Princeton, mimo że dalej radziły sobie świetnie, przesunęły się w dół. Co się zdarzyło w instytucie w Kalifornii, co spowodowało tak znaczącą poprawę jakości?

Odpowiedzi dostarcza „U.S. News and World Report", nie zaś kalifornijska uczelnia. Czasopismo zmieniło w 1999 roku strukturę indeksu, co zaowocowało znacznymi zmianami w wynikach szkół.

Bruce Gottlieb (1999) podaje to jako przykład wpływu zmiany punktacji na wynik badań:

> „Jak więc Instytut Kalifornijski znalazł się na szczycie? Cóż, jedną ze zmiennych w rankingu szkół były od dawna wydatki na wykształcenie studenta, w której to dziedzinie uczelnia ta tradycyjnie zajmowała czołową pozycję. Jednak aż do tego roku «U.S. News and World Report» uwzględniał tylko kolejność szkół w świetle tego kryterium – pierwsze miejsce, drugie itd. – nie zaś poziom wydatków na tle innych szkół. Nie miało więc znaczenia, czy Instytut Kalifornijski pobił Harvard o jednego dolara, czy o sto tysięcy dolarów. Dwie inne szkoły, które uzyskały wyraźnie wyższe wyniki w tym roku, to Massachussets Institute of Technology (z czwartego

Oto podobny przykład punktacji indeksu, z badań dotyczących zadowolenia z pracy. Jedną z kluczowych zmiennych była depresja związana z pracą, mierzona za pomocą indeksu złożonego z następujących czterech stwierdzeń, dotyczących odczuć robotników, gdy myślą o sobie i swojej pracy:

- Czuję się przygnębiony i smutny.
- Czuję się zmęczony bez powodu.
- Jestem niespokojny i nie mogę wytrzymać bez ruchu.
- Jestem bardziej rozdrażniony niż zwykle.

Badacze, Amy Wharton i James Baron, piszą: „Każdemu z tych pytań przypisano kody: 4 = często, 3 = czasami, 2 = rzadko, 1 = nigdy" (Wharton, Baron 1987, s. 578). Następnie wyjaśniają, jak mierzyli inną zmienną, samoocenę związaną z pracą:

> „Samoocena związana z pracą opierała się na czterech pytaniach, które dotyczyły tego, jak respondenci postrzegają siebie w pracy: zadowolony/smutny, odnoszę sukcesy/nie odnoszę sukcesów, ważny/nieważny, staram się najlepiej jak mogę/nie staram się najlepiej jak mogę. Każde pytanie otrzymało punktację od 1 do 7, gdzie 1 oznacza postrzeganie samego siebie jako osoby niezadowolonej, nieodnoszącej sukcesów, nieważnej, niestarającej się pracować najlepiej jak może".
>
> (1987, s. 578)

Przeglądając literaturę z zakresu badań społecznych, znajdziesz wiele podobnych przykładów zastosowań tego rodzaju zbiorczych indeksów do pomiaru zmiennych. Niekiedy procedura indeksacji bywa kontrowersyjna, jak tego dowodzi tekst w ramce „Jaka jest najlepsza uczelnia w USA?".

Braki danych

Niezależnie od metody zbierania danych będziesz często spotykał się z problemem braków danych. W analizie treści artykułów prasowych pod kątem orientacji politycznej redakcji możesz na przykład odkryć, że dana gazeta nigdy nie zaprezentowała

miejsca na trzecie) i Uniwersytet Johna Hopkinsa (z czternastego na siódme). Wszystkie trzy mają wysokie wydatki na głowę studenta i wszystkie specjalizują się w naukach ścisłych. Uczelnie takie mogą obliczać swoje budżety na badania w kategoriach wydatków na głowę studenta, mimo że studenci nie odnoszą żadnych bezpośrednich korzyści z faktu, że ich profesorowie prowadzą kosztowne badania poza zajęciami.

W rankingu wyższych uczelni sprzed dwóch lat «U.S. News and World Report» zwrócił uwagę na to właśnie zagadnienie, twierdząc, że brał pod uwagę tylko kolejność uczelni według wydatków na studenta, nie zaś faktyczne kwoty, ponieważ wydatki takie na uczelniach z szerokim programem badań oraz w szkołach medycznych są znacznie wyższe niż w innych szkołach danej kategorii. Inaczej mówiąc, dwa lata temu redakcja czasopisma uważała, że nie ma powodu, by przyznawać punkty dla California Institute of Technology, Massachussets Institute of Technology oraz Uniwersytetu Johna Hopkinsa ze względu na ich liczne, bogato wyposażone laboratoria, które nie wpływają faktycznie na poprawę kształcenia studentów".

Gottlieb zbadał każdą ze zmian indeksu i zadał następnie pytanie, jak wypadłby w rankingu California Institute of Technology, zajmujący w 1998 roku dziewiąte miejsce, gdyby zmiana formuły indeksu została przeprowadzona rok wcześniej. Jego wniosek jest następujący: w 1998 roku uczelnia ta również byłaby pierwsza. Innymi słowy, pozorna poprawa była jedynie funkcją punktacji indeksu.

Mierniki złożone, takie jak skale i indeksy, są cennymi narzędziami, mogącymi służyć zrozumieniu społeczeństwa. Ważne jednak, byśmy pamiętali, jak tworzy się te mierniki i co wynika z ich budowy.

Która zatem uczelnia jest *naprawdę* najlepsza w Stanach Zjednoczonych? Zależy to od definicji słowa „najlepszy". Nie ma „naprawdę najlepszej", są tylko rozmaite konstrukcje społeczne, które możemy tworzyć.

Źródło: „U.S. News and World Report", „America's Best Colleges", 30 sierpnia 1999; Bruce Gottlieb, *Cooking the School Books: How U.S. News Cheats in Picking Its Best American College*, artykuł zamieszczony 31 sierpnia 1999 przez Slate (online), dostępny pod adresem http://www.slate.com/crapshoot/99-08-31/crapshoot.asp

stanowiska redakcji w jednej z badanych kwestii. W schemacie eksperymentalnym, wymagającym kilkakrotnego powtórzenia testu, niektórzy badani mogą nie być w stanie uczestniczyć w niektórych sesjach. W niemal każdych badaniach sondażowych niektórzy respondenci nie odpowiadają na pewne pytania (albo wybierają odpowiedź „nie wiem"). Choć braki danych są problemem na każdym etapie analizy, są one szczególnie uciążliwe przy tworzeniu indeksu. Istnieje jednak kilka metod radzenia sobie z tym problemem.

Po pierwsze, jeśli braki występują w dość niewielu przypadkach, można wyłączyć je z budowy indeksu i z analizy (tak zrobiłem w przykładzie z profesorami uczelni medycznych). Najważniejsze w takim przypadku jest ustalenie, czy liczba przypadków dostępnych do analizy będzie dostateczna i czy wyłączenie nie będzie powodować niereprezentatywności próby przy stosowaniu tego indeksu w analizie. Tę drugą możliwość można zbadać w drodze porównań – pod kątem innych relewantnych zmiennych – przypadków włączonych do indeksu i wyłączonych z niego.

Po drugie, można niekiedy potraktować braki danych jako jedną z dostępnych kategorii odpowiedzi. Na przykład jeśli w kwestionariuszu proszono respondentów o wskazanie ich uczestnictwa w różnych działaniach przez wybranie przy każdym z nich odpowiedzi „tak" lub „nie", wielu respondentów mogło wybrać „tak" przy niektórych działaniach, a pozostałe zostawić bez odpowiedzi. W takim przypadku możesz uznać, że brak odpowiedzi znaczy „nie", i przyznać punkty w taki sposób, jak gdyby respondent wybrał odpowiedź „nie".

Po trzecie, uważna analiza braków danych może pozwolić na interpretację ich znaczenia. Tworząc miernik konserwatyzmu politycznego, możesz na przykład stwierdzić, że respondenci, którzy nie udzielili odpowiedzi na dane pytanie, byli ogólnie równie konserwatywni w swoich odpowiedziach na pozostałe pytania jak ci, którzy wybrali w danym przypadku odpowiedź konserwatywną. Inny przykład: w jednym z nowszych badań dotyczących wierzeń religijnych odkryto, że ludzie, którzy odpowiadali „nie wiem" na pytanie o dane wierzenie, udzielali na pytania do-

tyczące innych wierzeń niemal identycznych odpowiedzi jak „niewierzący". (*Uwaga:* we własnych badaniach powinieneś traktować te przykłady nie jako wskazówki empiryczne, lecz jako ogólne sugestie co do sposobu analizy twoich danych). Jeśli analiza braków danych pozwala na taką interpretację, możesz zdecydować o odpowiedniej punktacji tego typu przypadków.

Istnieje wiele innych sposobów radzenia sobie z problemem braków danych. Jeśli dane pytanie ma wiele możliwych wartości, odpowiadających poszczególnym kategoriom odpowiedzi, można przypisać przypadkom z brakami danych wartość średnią: na przykład jeśli możliwe wartości to 0, 1, 2, 3 i 4, mógłbyś przypisać brakowi danych wartość 2. Dla zmiennych ciągłych, takich jak wiek, mógłbyś także przypisać przypadkom z brakami danych średnią (więcej o tym w rozdziale 15). Brakujące dane można też uzupełnić, przypisując różne wartości losowo. Wszystkie te rozwiązania należy uznać za zachowawcze, ponieważ osłabiają „czystość" indeksu i redukują prawdopodobieństwo istnienia hipotetycznie przez ciebie przewidywanej zależności między indeksem a innymi zmiennymi.

Jeśli tworzysz indeks z dużej liczby pytań, możesz niekiedy traktować brak danych proporcjonalnie do uzyskanych obserwacji. Załóżmy, że twój indeks składa się z sześciu wskaźników, a ty dysponujesz tylko czterema obserwacjami na temat danego badanego. Jeśli badany otrzymał 4 punkty na 4 możliwe, mógłbyś mu przyznać 6 punktów, jeśli dostał 2 punkty (czyli połowę możliwego maksymalnego wyniku przy czterech pytaniach), mógłbyś mu przyznać 3 punkty (połowa możliwego maksymalnego wyniku przy sześciu obserwacjach).

Wybór konkretnej metody jest tak bardzo uzależniony od sytuacji badawczej, że nie jestem w stanie rozsądnie doradzić jednej „najlepszej" metody czy też uszeregować pod tym kątem metod powyżej opisanych. Wyłączenie wszystkich przypadków z brakami danych może wpłynąć na reprezentatywność wyników, jednak ich włączenie przez przypisanie brakom danych wyników punktowych może wpłynąć na treść wyników. Najbezpieczniejsza i najlepsza metoda polega na stosowaniu więcej niż jednej z tych metod, co pozwala sprawdzić, czy za każdym razem używając indeksu dochodzimy do tych samych wniosków. Ostatecznym celem analizy jest i tak zawsze wyjaśnienie danych.

Walidacja indeksu

Jak dotąd, omówiliśmy wszystkie etapy procesu wyboru i ustalania punktacji pytań, które składają się na indeks przeznaczony do pomiaru danej zmiennej. Jeśli każdy z tych etapów został przeprowadzony z należytą starannością, wzrasta prawdopodobieństwo tego, że indeks istotnie mierzy daną zmienną. Aby dowieść, że osiągnęliśmy sukces, musimy jednak ocenić trafność indeksu, czyli dokonać tzw. walidacji. Zgodnie z podstawową logiką oceny trafności zakładamy, że indeks stanowi miernik jakiejś zmiennej, czyli przypadki można uporządkować według kolejności w kategoriach tej zmiennej za pomocą wyników uzyskanych na tym indeksie. Indeks konserwatyzmu politycznego porządkuje ludzi w kategoriach ich względnego konserwatyzmu. Jeśli indeks spełnia to zadanie, ludzie ocenieni jako względnie konserwatywni za pomocą tego indeksu okazaliby się także względnie konserwatywni według innych wskaźników orientacji politycznej, na przykład według odpowiedzi na inne pytania kwestionariusza. Istnieje kilka metod walidacji indeksu.

Analiza pytań

Pierwszym krokiem w ocenie trafności indeksu jest walidacja wewnętrzna, czyli **analiza pytań**. W trakcie analizy pytań badamy, w jakim zakresie indeks jest powiązany z poszczególnymi pytaniami w nim zawartymi (lub w jakim zakresie pozwala przewidywać odpowiedzi na nie udzielone). Oto przykład przebiegu tego etapu.

W indeksie orientacji naukowej profesorów uczelni medycznych wyniki uzyskane za pomocą indeksu wynoszą od 0 (zainteresowany przede wszystkim opieką nad pacjentem) do 3 (zainteresowany przede wszystkim badaniami naukowymi). Zajmijmy się teraz jednym z pytań indeksu: czy respondent chciałby poszerzyć swoją wiedzę raczej z zakresu całościowego leczenia pacjenta, czy też raczej z zakresu pod-

stawowych mechanizmów rządzących skutecznością terapii. To ostatnie traktujemy jako wyraz orientacji bardziej naukowej. Poniższa pusta tabela ilustruje metodę badania zależności między tym indeksem a pojedynczym pytaniem.

	Indeks orientacji naukowej			
	0	1	2	3
Odsetek respondentów zainteresowanych podstawowymi mechanizmami	??	??	??	??

Gdybyś poświęcił kilka chwil na przyjrzenie się tabeli, zauważyłbyś może, iż znamy już liczby, które znajdą się w dwóch polach. Aby otrzymać 3 punkty w indeksie, respondent musiał odpowiedzieć na to pytanie: „podstawowe mechanizmy", musiał też udzielić „naukowych" odpowiedzi na pozostałe dwa pytania. Zatem 100% trójek uzyskanych za pomocą tego indeksu odpowiedziało „podstawowe mechanizmy". Na mocy tego samego rozumowania wszyscy z wynikiem 0 musieli odpowiedzieć na to pytanie „całościowe leczenie pacjenta". Zatem 0% tej grupy respondentów odpowiedziało „podstawowe mechanizmy". Oto jak wygląda tabela ze znanymi nam już informacjami.

	Indeks orientacji naukowej			
	0	1	2	3
Odsetek respondentów zainteresowanych podstawowymi mechanizmami	0	??	??	100

Jeśli pojedyncze pytanie ma stanowić dobre odbicie całości indeksu, powinniśmy oczekiwać, że jedynki i dwójki wypełnią w stopniowalny sposób przestrzeń między 0% a 100%. Więcej dwójek niż jedynek powinno wybrać „podstawowe mechanizmy". Rezultat ten nie wynika jednak z samej budowy indeksu – jest to kwestia empiryczna, którą rozwiązujemy w toku analizy pytań. Oto wynik analizy tego konkretnego pytania.

	Indeks orientacji naukowej			
	0	1	2	3
Odsetek respondentów zainteresowanych podstawowymi mechanizmami	0	16	91	100

Jak widać, zgodnie z naszymi przypuszczeniami, że dwójki są bardziej naukowo zorientowane niż jedynki, odkrywamy więcej dwójek (91%) niż jedynek (16%), które udzieliły odpowiedzi „podstawowe mechanizmy".

Analiza dwóch pozostałych pytań, wchodzących w skład indeksu, daje podobne wyniki, przedstawione poniżej.

	Indeks orientacji naukowej			
	0	1	2	3
Odsetek respondentów twierdzących, że są lepszymi nauczycielami jako naukowcy	0	4	14	100
Odsetek respondentów preferujących artykuły na temat uzasadnień terapii	0	80	97	100

Każde z tych pytań wydaje się zatem właściwym składnikiem indeksu. Wydaje się, że każde z nich odbija tę samą cechę, którą mierzy indeks jako taki.

W indeksie złożonym z wielu pytań etap analizy stanowi wygodny test niezależnego wkładu każdego z pytań do całości indeksu. Jeśli dane pytanie okazuje się słabo związane z indeksem, można przypuszczać, że inne pytania wchodzące w skład tego indeksu znoszą wkład tego pytania. Jeśli jakieś pytanie nic nie wnosi do ogólnej mocy indeksu, należy je wyłączyć.

Mimo że analiza pytań jest istotnym pierwszym sprawdzianem trafności indeksu, nie jest to sprawdzian wystarczający. Jeśli indeks adekwatnie mierzy daną zmienną, powinien skutecznie przewidywać inne wskaźniki tej zmiennej. By to sprawdzić, musimy przejść do pytań nie wchodzących w skład indeksu.

Walidacja zewnętrzna

Osoby ocenione na podstawie indeksu jako konserwatywne politycznie powinny okazać się konserwatywne również wedle innych mierników, na przykład odpowiedzi na inne pytania kwestionariusza. Mówimy oczywiście o względnym konserwatyzmie, ponieważ nie możemy zdefiniować *konserwatyzmu* w sposób absolutny. Jednak respondenci ocenieni jako najbardziej konserwatywni na podstawie indeksu powinni

Tabela 6.1

Walidacja indeksu orientacji naukowej

	Indeks orientacji naukowej			
	niski poziom		wysoki poziom	
	0	1	2	3
Odsetek zainteresowanych uczęszczaniem na wykłady naukowe w szkołach medycznych	34	42	46	65
Odsetek twierdzących, że profesorowie powinni mieć doświadczenie jako badacze	43	60	65	89
Odsetek tych, którzy woleliby, żeby ich obowiązki obejmowały tylko działalność badawczą	0	8	32	66
Odsetek zaangażowanych w badania w ubiegłym roku akademickim	61	76	94	99

również okazać się najbardziej konserwatywni na podstawie odpowiedzi na inne pytania. Ci, których oceniono jako najmniej konserwatywnych na podstawie indeksu, powinni okazać się najmniej konserwatywni w odpowiedziach na inne pytania. Właściwe uporządkowanie grup respondentów według indeksu powinno umożliwić przewidywanie porządku tych grup według odpowiedzi na inne pytania dotyczące poglądów politycznych.

W naszym przykładzie indeksu orientacji naukowej w kwestionariuszu znalazło się kilka pytań, które umożliwiły taką **walidację zewnętrzną**. Tabela 6.1 przedstawia kilka spośród tych pytań, które dostarczają nam informacji o trafności indeksu. Po pierwsze, zauważamy, że indeks pozwala w dużym stopniu przewidywać odpowiedzi na pytania użyte w sprawdzianie w tym znaczeniu, że porządek odpowiedzi naukowych we wszystkich czterech grupach jest ten sam, co porządek ustalony na podstawie samego indeksu. Oznacza to, że odsetki odzwierciedlają coraz wyższy poziom orientacji naukowej w miarę przesuwania się wzdłuż rzędów tabeli. Jednocześnie każde pytanie daje inny opis ogólnej orientacji naukowej. Na przykład ostatnie pytanie wskazuje, że znakomita większość profesorów była w ciągu ubiegłego roku zaangażowana w projekty badawcze. Gdyby był to jedyny wskaźnik orientacji naukowej, musielibyśmy uznać, że niemal wszyscy profesorowie są zorientowani naukowo. Niemniej jednak osoby, które uzyskały wyższy wynik na podstawie indeksu, były częściej zaangażowane w projekty badawcze niż osoby, które na podstawie indeksu ocenione zostały jako stosunkowo mniej naukowo zorientowane. Trzecie pytanie rysuje obraz nieco odmienny: jedynie mniejszość ogółu profesorów twierdzi, że woleliby, gdyby ich obowiązki ograniczały się tylko do badań naukowych. Jednak względne odsetki badanych podających tę odpowiedź są zgodne z wynikami uzyskanymi na podstawie indeksu.

Zły indeks kontra złe wskaźniki trafności

Niemal każdy twórca indeksu musi czasem stawić czoło wyraźnej porażce zewnętrznego sprawdzianu trafności indeksu. Jeśli wewnętrzna analiza pytań wykazuje niespójne relacje między pytaniami wchodzącymi w skład indeksu a samym indeksem, coś jest nie tak z tym indeksem. Jeśli jednak indeks nie umożliwia dobrego przewidywania odpowiedzi na pytania używane do walidacji zewnętrznej, wnioski są mniej jednoznaczne. Musimy w takiej sytuacji wybrać jedną z dwóch możliwości: (1) indeks nie mierzy adekwatnie badanej zmiennej lub (2) pytania użyte do walidacji nie mierzą adekwatnie owej zmiennej i nie stanowią tym samym wystarczającego testu dla indeksu.

Biorąc pod uwagę, jak długo i sumiennie pracowałeś nad stworzeniem danego indeksu, zapewne uznasz drugie wyjaśnienie za bardziej przekonujące. Na ogół będziesz sądził, że skoro włączyłeś do indeksu najlepsze wskaźniki badanej zmiennej, to pytania użyte do walidacji są wskaźnikami drugorzędnymi. Powinieneś jednak pamiętać, że indeks jest pomyślany jako bardzo skuteczny miernik zmiennej. Powinien więc być jakoś powiązany z każdym pytaniem,

które choćby w niewielkim stopniu dotyczy mierzonej przezeń zmiennej.

W razie niepowodzenia zewnętrznej walidacji indeksu powinieneś raz jeszcze zbadać indeks, zanim zdecydujesz, że pytania użyte do walidacji są niewystarczające. Jednym ze sposobów może być badanie relacji między pytaniami użytymi do walidacji a pojedynczymi pytaniami wchodzącymi w skład indeksu. Jeśli stwierdzisz, że niektóre pytania indeksu są powiązane z pytaniami użytymi do walidacji, inne zaś nie, pozwoli ci to lepiej zrozumieć pierwotną budowę indeksu.

Nie ma prostego rozwiązania tego problemu: tę mękę musi przejść każdy poważny badacz. To, na ile rozsądna okaże się twoja decyzja przyjęcia danego indeksu, zależy ostatecznie od użyteczności tego indeksu w dalszej analizie. Być może w pierwszej chwili uznasz, że indeks jest dobry, a pytania użyte do walidacji wadliwe, potem jednak stwierdzisz, że badana zmienna (mierzona przez indeks) nie pozostaje w oczekiwanych zależnościach z innymi zmiennymi. Być może będziesz wtedy zmuszony stworzyć nowy indeks.

Status kobiety: przykład procesu tworzenia indeksu

Nasze rozważania na temat tworzenia indeksu koncentrowały się jak dotąd na konkretnym kontekście badań sondażowych, jednak inne typy badań również coraz częściej posługują się tym rodzajem miernika złożonego. Na przykład kiedy ONZ postanowiła zbadać status kobiety w świecie (1995), zdecydowano się stworzyć dwa indeksy, mierzące dwa różne wymiary.

Indeks Rozwoju według Płci (*Gender-related Development Index – GDI*) porównywał mężczyzn i kobiety pod kątem trzech wskaźników: przewidywanej długości życia, wykształcenia i dochodu. Wskaźników tych często używa się w badaniach statusu kobiety w świecie. Najwyżej na podstawie tego indeksu uplasowały się kraje skandynawskie: Norwegia, Szwecja, Finlandia i Dania.

Drugi indeks, Miara Dostępu do Władzy według Płci (*Gender Empowerment Measure – GEM*), dotyczył raczej kwestii władzy i zawierał trzy różne wskaźniki:

- udział kobiet wśród parlamentarzystów,
- udział kobiet wśród osób na stanowiskach administracyjnych, kierowniczych, wysoko kwalifikowanych i technicznych,
- miernik dostępu do pracy i wynagrodzeń.

Raz jeszcze kraje skandynawskie uplasowały się bardzo wysoko, ale dołączyły do nich Kanada, Nowa Zelandia, Holandia, Stany Zjednoczone i Austria. Wykorzystanie dwóch mierników równości płci zamiast jednego pozwoliło badaczom na wprowadzenie bardziej złożonych różnicowań. Na przykład w kilku krajach, zwłaszcza w Grecji, Francji czy Japonii, status kobiety został oceniony stosunkowo dobrze według GDI, ale wypadł kiepsko według GEM. Tak więc kobiety, radzące sobie nieźle w zakresie dochodu, wykształcenia i długości życia, nadal nie miały dostępu do władzy. Wyniki GDI okazywały się wyższe w krajach zamożniejszych, natomiast wyniki GEM dowodziły, że dostęp kobiet do władzy w mniejszym stopniu zależy od bogactwa kraju, jako że wiele biednych krajów rozwijających się wyprzedziło pod tym względem niektóre bogate kraje przemysłowe.

Badanie kilku różnych wymiarów analizowanych zmiennych pozwoliło badaczom ONZ odkryć zwykle niezauważany aspekt zarobków kobiet. Międzynarodowe Komunikaty o Stanie Populacji (*Population Communications International*) podsumowały to odkrycie następująco:

> „Co roku kobiety wnoszą w gospodarkę światową niewidzialny wkład w wysokości 11 bilionów dolarów, twierdzą autorzy raportu UNDP (*United Nations Development Programme*), po obliczeniu wartości zarówno bezpłatnej, jak i niedostatecznie opłacanej pracy kobiet według cen rynkowych. «Niedoocena» pracy kobiet nie tylko wpływa negatywnie na ich siłę nabywczą, zgodnie z HDR (*Human Development Report*) z 1995 roku, ale także obniża i tak już niski status społeczny kobiet i wpływa na możliwość posiadania przez nie własności i wykorzystywania kredytu. Mahbub al Haq, główny autor raportu, twierdzi: «gdyby praca kobiet została należycie uwzględniona w krajowych statystykach, wstrząsnęłoby

to mitem o mężczyznach jako głównych żywicielach świata». Raport UNDP ujawnia, że kobiety niemal we wszystkich krajach pracują dłużej niż mężczyźni, biorąc pod uwagę pracę płatną i bezpłatną. W krajach rozwijających się kobiety wykonują około 53% pracy ogółem i spędzają dwie trzecie czasu pracy na zajęciach bezpłatnych. W krajach uprzemysłowionych kobiety wykonują przeciętnie 51% pracy ogółem i – podobnie jak ich odpowiedniczki w krajach rozwijających się – wykonują około dwóch trzecich swojej pracy bezpłatnie. Mężczyźni w krajach uprzemysłowionych są wynagradzani za dwie trzecie swojej pracy".

(1996, s. 1)

Jak widać, można tworzyć indeksy dla rozmaitych rodzajów danych i w różnych celach. Teraz przejdziemy od tworzenia indeksów do rozważań na temat technik skalowania.

Budowa skali

Dobry indeks umożliwia uporządkowanie przypadków według wartości określonej zmiennej. Wszystkie indeksy opierają się na tej przesłance: senator, który głosował za siedmioma konserwatywnymi projektami ustaw, jest uważany za bardziej konserwatywnego niż ten, który głosował tylko za czterema. Indeks może jednak nie uwzględniać faktu, że nie wszystkie wskaźniki danej zmiennej są równie ważne lub równie silne. Pierwszy senator mógł głosować za siedmioma umiarkowanie konserwatywnymi projektami ustaw, natomiast drugi mógł głosować za czterema projektami skrajnie konserwatywnymi. (Drugi senator mógł uznać pozostałe siedem projektów ustaw za zbyt liberalne i głosować przeciw nim).

Skale zapewniają większy stopień pewności uszeregowania dzięki uwzględnieniu intensywności wskaźników. Kilka pytań wchodzących w skład miernika złożonego może mierzyć różne natężenie zmiennej. Istnieje wiele metod skalowania. Omówimy cztery spośród nich, aby zilustrować różnorodność dostępnych technik, włącznie z techniką zwaną dyferencjałem semantycznym. Mimo że w przykładach tych odwołuję się przede wszystkim do kwestionariuszy, logika skalowania, podobnie jak logika indeksowania, stosuje się również do innych metod badawczych.

Skala dystansu społecznego Bogardusa

Załóżmy, że interesuje cię, do jakiego stopnia obywatele Stanów Zjednoczonych skłonni są integrować się, powiedzmy, z Albańczykami. Mógłbyś zadać następujące pytania:

1. Czy zgodziłbyś się, by Albańczycy mieszkali w twoim kraju?
2. Czy zgodziłbyś się, by Albańczycy mieszkali w twojej miejscowości?
3. Czy zgodziłbyś się, by Albańczycy mieszkali w twojej dzielnicy?
4. Czy zgodziłbyś się, by Albańczyk był twoim sąsiadem?
5. Czy zgodziłbyś się, by twoje dziecko poślubiło Albankę/Albańczyka?

Pytania te stopniują bliskość kontaktu respondenta z Albańczykami. Wychodząc od początkowego zainteresowania pomiarem skłonności do integracji z Albańczykami, opracowałeś oto kilka pytań wskaźnikujących różny stopień intensywności tej zmiennej. Zaprezentowane typy pytań tworzą **skalę dystansu społecznego Bogardusa** (stworzoną przez Emory Bogardusa). Skala ta służy do pomiaru skłonności ludzi do uczestniczenia w – zróżnicowanych co do stopnia bliskości – stosunkach społecznych z innymi grupami ludzi.

Wyraźne różnice intensywności sugerują istnienie struktury pytań. Osoba, która akceptuje pewien stopień integracji, będzie też przypuszczalnie skłonna zaakceptować wszystkie możliwości poprzedzające go na liście – czyli o mniejszej intensywności. Na przykład osoba zgadzająca się, by Albańczycy mieszkali z nią w jednej dzielnicy, zgodzi się z pewnością, by mieszkali oni w tej samej miejscowości lub kraju, może jednak nie zgadzać się, by byli jej sąsiadami czy krewnymi. Tak zatem wygląda struktura logiczna zawarta w wymienionych pytaniach.

Można by oczekiwać, że w badaniach empirycznych najwięcej osób zaakceptuje Albańczyków jako współobywateli, najmniej zaś jako członków rodziny. W tym sensie można mówić o „łatwych" (jak zamieszkiwanie Albańczyków w USA) i „trudnych" (jak małżeństwo dziecka z Albańczykiem) pytaniach kwestionariusza. Na pytania łatwe przystanie więcej badanych, na trudne – mniej. Logika wymaga, by – z uwzględnieniem kilku nieuniknionych wyjątków – osoba, która odrzuciła daną relację przedstawioną na skali, odrzuciła również wszystkie następujące po niej trudniejsze relacje.

Skala dystansu społecznego Bogardusa ukazuje znaczącą oszczędność skalowania jako narzędzia redukcji danych. Wiedząc, ile rodzajów stosunków z Albańczykami zaakceptuje dany badany, wiemy też, które z nich zaakceptuje. Tak więc jedna liczba pozwala poprawnie podsumować pięć czy sześć pytań bez żadnych strat informacji.

Skala Thurstone'a

Wewnętrzna struktura skali Bogardusa często nie odpowiada rodzajowi mierzonych zmiennych. Tego typu logiczna struktura kilku wskaźników zdarza się tak naprawdę dość rzadko. **Skala Thurstone'a** (stworzona przez Louisa Thurstone'a) stanowi próbę opracowania formatu grupowania wskaźników zmiennej, które powiązane są przynajmniej w strukturę empiryczną. Grupa sędziów otrzymuje na przykład setki pytań, które uznawane są za wskaźniki danej zmiennej. Każdy sędzia ma następnie ocenić, jak mocnym wskaźnikiem zmiennej jest każde z pytań, przypisując im wartość punktową, powiedzmy, od 1 do 13. Na przykład gdyby tą zmienną były *uprzedzenia*, sędziowie przypisywaliby wartość 1 najsłabszym wskaźnikom uprzedzeń, 13 – najsilniejszym wskaźnikom, wartości zaś pośrednie wskaźnikom leżącym gdzieś pośrodku.

Po wykonaniu tego zadania przez sędziów badacz analizuje oceny przyznane poszczególnym pytaniom przez wszystkich sędziów, by ustalić, co do których pytań byli oni najbardziej zgodni. Pytania, co do których zdecydowanie nie ma zgodności, zostałyby odrzucone jako niejednoznaczne. Spośród pytań, których punktacja okazała się zgodna, badacz wybrałby jedno lub więcej, by odpowiadały każdemu punktowi od 1 do 13 na skali.

Wybrane w ten sposób pytania mogłyby zostać włączone do kwestionariusza. Oczekiwalibyśmy wówczas, że respondenci, którzy okazali się uprzedzeni w kategoriach pytań o wartości punktowej 5, byliby uprzedzeni również w kategoriach pytań o niższych wartościach punktowych. Można by się również spodziewać, że jeśli nie okazali się uprzedzeni na podstawie pytań o mocy 6 punktów, nie będą też uprzedzeni na podstawie pytań o większej mocy.

Starannie opracowana i wypunktowana skala Thurstone'a odpowiada pod względem oszczędności i efektywności, związanej z redukcją danych, skali dystansu społecznego Bogardusa. Każdemu respondentowi można przypisać pojedynczy wynik (moc najtrudniejszego stwierdzenia, z którym zgodził się respondent), który w sposób adekwatny odzwierciedlałby odpowiedzi na kilka pytań kwestionariusza. Podobnie jak w przypadku skali Bogardusa, respondenta z wynikiem 6 możemy uważać za bardziej uprzedzonego niż respondenta z wynikiem 5 lub niższym.

Skali Thurstone'a nie używa się obecnie zbyt często, przede wszystkim ze względu na ogromny nakład energii i czasu, którego wymaga zaangażowanie od 10 do 15 sędziów do oceny pytań. Możliwe, że trzeba by było skorzystać z usług profesjonalnych badaczy, ponieważ jakość sądu zależy od doświadczeń sędziów związanych z określoną zmienną. Co więcej, znaczenie pytań wskaźnikujących daną zmienną zmienia się zwykle z czasem, tak więc może się okazać, że pytanie o określonej wadze zyskuje później zupełnie inną wagę. Zachowanie użyteczności skali Thurstone'a wymaga jej okresowej aktualizacji.

Skala Likerta

Być może zdarzyło ci się już zetknąć z terminem skala Likerta, używanym na określenie pytania kwestionariuszowego zawierającego kategorie

odpowiedzi w rodzaju „zdecydowanie się nie zgadzam", „nie zgadzam się", „zgadzam się" i „zdecydowanie się zgadzam". Technicznie rzecz biorąc jest to określenie niewłaściwe, aczkolwiek Rensis Likert (wymawia się tak, jak się pisze: LIKert) istotnie jest twórcą tego powszechnie stosowanego formatu pytania.

Szczególna przydatność tego formatu zasadza się na jednoznacznym uporządkowaniu kategorii odpowiedzi. Gdyby respondentom pozostawiono swobodę lub dano wybór pomiędzy kategoriami takimi, jak: „w zasadzie się zgadzam", „właściwie się zgadzam", „naprawdę się zgadzam" itp., badacz nie byłby w stanie ocenić względnej intensywności akceptacji danego stwierdzenia przez poszczególnych respondentów. Format Likerta rozwiązuje ten problem.

Likertowi chodziło jednak o coś więcej. Stworzył metodę, za pomocą której ten format pytania umożliwia określenie względnej intensywności różnych pytań. Weźmy najprostszy przykład: chcemy zmierzyć uprzedzenia wobec kobiet. Tworzymy w tym celu zestaw 20 stwierdzeń, z których każde jest wyrazem tego typu uprzedzeń. Jedno z nich mogłoby brzmieć: „kobiety są gorszymi kierowcami niż mężczyźni", inne zaś na przykład: „kobiety nie powinny mieć praw wyborczych". Technika skalowania Likerta ukazałaby różnice intensywności między tymi pytaniami, a zarazem ustaliła intensywność pozostałych 18 stwierdzeń.

Załóżmy, że prosimy próbę ludzi, by powiedzieli, czy się zgadzają, czy nie zgadzają z każdym z tych 20 twierdzeń. Przyznanie każdemu ze wskaźników uprzedzeń względem kobiet jednego punktu dałoby nam indeks z zakresem wyników od 0 do 20. **Skala Likerta** idzie o krok dalej i oblicza przeciętne wyniki uzyskane za pomocą indeksu dla osób zgadzających się z poszczególnymi zdaniami. Załóżmy, że wszyscy, którzy zgodzili się, że kobiety prowadzą gorzej niż mężczyźni, uzyskali na podstawie indeksu przeciętny wynik 1,5 punktu (na 20 możliwych). Osoby, które zgodziły się, że kobiety nie powinny mieć praw wyborczych, mogłyby uzyskać przeciętnie 19,5 punktu – co oznacza, że odpowiedź ta wskazuje na wyższy poziom uprzedzeń.

W wyniku analizy pytań można ponownie przypisać respondentom wyniki punktowe składające się na skalę: 1,5 dla osób uważających kobiety za gorszych niż mężczyźni kierowców, 19,5 dla tych, którzy twierdzą, że kobiety nie powinny głosować, i odpowiednie ilości punktów dla innych odpowiedzi, odzwierciedlające związek poszczególnych pytań z początkowym prostym indeksem. Jeśli osoby nie zgadzające się ze zdaniem: „mógłbym głosować na kobietę w wyborach prezydenckich" uzyskiwały na podstawie indeksu przeciętny wynik 15, skala przyznawałaby 15 punktów respondentom nie zgadzającym się z tym stwierdzeniem.

W praktyce skalę Likerta stosuje się obecnie dość rzadko. Nie wiem dlaczego – być może wydaje się zbyt skomplikowana. Format pytań zaproponowany przez Likerta jest natomiast jednym z najpowszechniej używanych we współczesnych kwestionariuszach. Jest on na ogół wykorzystywany obecnie do tworzenia prostych indeksów. Na przykład przy pięciu kategoriach odpowiedzi można ustalić punktację od 0 do 4 lub od 1 do 5, uwzględniając kierunek pytań (na przykład przypisując wynik 5 odpowiedzi „zdecydowanie się zgadzam" przy pytaniu pozytywnym i „zdecydowanie się nie zgadzam" przy negatywnym). Każdemu z respondentów zostałby wówczas przypisany wynik ogólny, stanowiący podsumowanie wyników uzyskanych przez niego za odpowiedzi na poszczególne pytania.

Dyferencjał semantyczny

Podobnie jak w formacie Likerta, w przypadku **dyferencjału semantycznego** prosimy respondentów w kwestionariuszu, by wybrali między dwiema przeciwnymi postawami, stosując przy tym wyrażenia określające stopień intensywności postawy między przeciwieństwami. Oto przykład.

Załóżmy, że oceniasz skuteczność wpływu wykładu na temat odbioru nowej muzyki na odbiór muzyki przez badanych. W ramach badań zamierzasz odtwarzać badanym kilka wybranych utworów muzycznych i prosić ich, by opisali swoje wrażenia. Dobrym narzędziem

dla uchwycenia tych wrażeń będzie format dyferencjału semantycznego.

Na początku musisz określić, jakie aspekty utworów będą oceniane przez badanych. Następnie musisz znaleźć dwa przeciwstawne terminy na określenie dwóch biegunów każdego aspektu. Załóżmy, że jeden z interesujących cię aspektów to pytanie, czy badanemu podoba się dany kawałek, czy nie. W tym przypadku dwa przeciwstawne określenia brzmiałyby: „podoba mi się" i „nie podoba mi się". Mógłbyś również starać się dowiedzieć, czy badany uznał dany fragment za „prosty", czy „skomplikowany", „harmonijny", czy „nieharmonijny" itp.

Skoro już masz odpowiednie aspekty i określenia ich skrajności, mógłbyś przygotować arkusz ocen, który badany wypełniałby po wysłuchaniu każdego fragmentu muzyki. Rysunek 6.5 przedstawia, jak mogłoby to wyglądać.

W każdym wierszu arkusza ocen badany wskazywałby, jakie były jego wrażenia: czy muzyka mu się podobała, czy nie podobała, i czy „trochę", czy „bardzo". By uniknąć obciążeń we wzorcach odpowiedzi na tego typu pytania, nie należy umieszczać po tej samej stronie terminów, co do których możemy przypuszczać, że są ze sobą powiązane. Zauważ na przykład, że „nieharmonijny" i „tradycyjny" są po lewej stronie arkusza ocen, natomiast „harmonijny" i „nowoczesny" są po prawej. Fragmenty ocenione jako „nieharmonijne" będą prawdopodobnie zarazem odbierane jako „nowoczesne", w przeciwieństwie do „tradycyjnych".

Zarówno skala Likerta, jak i dyferencjał semantyczny charakteryzują się większą ścisłością i ustrukturyzowaniem niż inne formaty pytań kwestionariuszowych. Jak już wcześniej wspominałem, formaty te przydatne są przy pozyskiwaniu danych zarówno dla celów skalowania, jak i indeksowania.

Skala Guttmana

Badacze stosują dziś często skalę stworzoną przez Louisa Guttmana. Podobnie jak Bogardus, Thurstone i Likert, Guttman wyszedł z założenia, że niektóre spośród analizowanych pytań mogą okazać się bardziej intensywnymi wskaźnikami zmiennej. Jeden przykład powinien zapewnić dostateczną ilustrację tej prawidłowości.

Przypominasz sobie zapewne, że w użytym poprzednio przykładzie pomiaru orientacji naukowej wśród profesorów uczelni medycznych stworzyliśmy prosty indeks. Trzy pytania, wchodzące w skład tego indeksu, zasadniczo tworzą jednak – niekiedy tak się zdarza – skalę Guttmana.

Tworzenie **skali Guttmana** zaczyna się od tych samych kroków, co tworzenie indeksu. Wychodzimy od badania trafności fasadowej pytań dostępnych analizie. Następnie badamy zależności dwuzmiennowe i – niekiedy – wielozmiennowe między pytaniami. Jednak konstruując skalę, poszukujemy również dość „trudnych" i „łatwych" wskaźników badanej zmiennej.

Nieco wcześniej, gdy mówiliśmy o postawach wobec prawa kobiety do aborcji, omówiliśmy kilka czynników, które mogą wpływać na opinie badanych: to, czy kobieta jest zamężna, czy jej zdrowie jest zagrożone itp. Te różne warunki stanowią doskonały przykład skali Guttmana.

Rysunek 6.5

Dyferencjał semantyczny: wrażenia dotyczące utworów muzycznych

	bardzo	raczej	obojętnie	raczej	bardzo	
podoba mi się	☐	☐	☐	☐	☐	nie podoba mi się
prosty	☐	☐	☐	☐	☐	skomplikowany
nieharmonijny	☐	☐	☐	☐	☐	harmonijny
tradycyjny	☐	☐	☐	☐	☐	nowoczesny

itd.

Oto odsetki ludzi, którzy w badaniach GSS w 1996 roku popierali prawo kobiety do aborcji w trzech różnych sytuacjach:

Poważne zagrożenie zdrowia kobiety	92%
Ciąża w wyniku gwałtu	86%
Kobieta nie jest zamężna	48%

Różnice w odsetkach osób popierających aborcję w trzech wymienionych sytuacjach zdają się prowadzić do wniosku, że każde pytanie wskaźnikuje inny poziom poparcia. Jeśli na przykład ktoś popiera aborcję w sytuacji poważnego zagrożenia życia matki, nie jest to zbyt silny wskaźnik ogólnego poparcia dla aborcji, ponieważ niemal wszyscy się z tym zgadzają. Poparcie prawa kobiet niezamężnych do aborcji jest wskaźnikiem znacznie większego poparcia dla aborcji w ogóle – postawę tę przyjęła mniej niż połowa próby.

Skala Guttmana opiera się na założeniu, że każdy, kto wykazuje wysoki poziom danej zmiennej, wykaże też poziom niższy. W tym przypadku moglibyśmy założyć, że każdy, kto popiera prawo niezamężnych kobiet do aborcji, popierałby także aborcję w przypadku ciąży powstałej w wyniku gwałtu lub zagrażającej życiu matki. Tabela 6.2 podaje wyniki sprawdzianu tego założenia, prezentując odsetki respondentów wpisujących się w każdy z możliwych wzorców odpowiedzi.

Pierwsze cztery wzorce odpowiedzi w tabeli można by nazwać *typami skalarnymi*: są to te wzorce, które tworzą strukturę skali. Przechodząc od respondentów, którzy popierali aborcję w każdej z trzech sytuacji (wiersz 1), do tych, którzy podali tylko dwie odpowiedzi pozytywne (wiersz 2), widzimy, że ci ostatni wybrali dwie łatwiejsze opcje. Osoby, które udzieliły tylko jednej pozytywnej odpowiedzi, wybierały najłatwiejszy wariant (zagrożenie zdrowia matki). Są wreszcie i tacy respondenci, którzy sprzeciwiają się przyznaniu kobiecie prawa do aborcji w każdej z trzech sytuacji (wiersz 4).

W drugiej części tabeli przedstawione zostały te wzorce odpowiedzi, które wyłamują się ze skalarnej struktury pytań. Najbardziej radykalne odstępstwo reprezentują tu dwa ostatnie wzorce: osoby akceptujące tylko najtrudniejszą opcję i osoby odrzucające tylko najłatwiejszą.

Ostatnia kolumna tabeli wskazuje liczbę respondentów w sondażu, którzy dali odpowiedzi pasujące do każdego z wzorców. Znakomita większość (1231 osób, czyli 98%) pasuje do jednego z typów skalarnych. Obecność typów mieszanych wskazuje jednak, że pytania te nie stanowią doskonałej skali Guttmana.

Przypomnijmy w tym miejscu, że jedną z głównych funkcji skali jest efektywna redukcja danych. Skale są techniką prezentacji da-

Tabela 6.2

Skalowanie poparcia dla prawa do aborcji

	Zdrowie kobiety	Gwałt	Kobieta niezamężna	Liczba przypadków
Typy skalarne	+	+	+	612
	+	+	–	448
	+	–	–	92
	–	–	–	79
				ogółem: 1231
Typy mieszane	–	+	–	15
	+	–	+	5
	–	–	+	2
	–	+	+	5
				ogółem: 27

+ = popiera prawo kobiety do wyboru aborcji
– = nie popiera prawa kobiety do wyboru aborcji

nych w formie syntetycznej przy zachowaniu możliwie dużej części pierwotnych informacji. Kiedy w poprzedniej części rozważań ujęliśmy pytania o orientację naukową w formę indeksu, przyznawaliśmy respondentom po jednym punkcie za każdą naukową odpowiedź. Gdyby te same trzy pytania zostały wypunktowane jako skala Guttmana, niektórzy respondenci uzyskaliby na skali wynik umożliwiający najtrafniejsze odtworzenie ich pierwotnych odpowiedzi na wszystkie trzy pytania.

W obecnym przykładzie dotyczącym postaw wobec aborcji respondenci wpisujący się w typy skalarne otrzymaliby te same wyniki, co przy tworzeniu indeksu. Osoby wybierające trzy odpowiedzi „za" (+ + +) tak samo miałyby 3 punkty; te, które wybrały dwie łatwiejsze opcje, a nie zaakceptowały najtrudniejszego wariantu (+ + –) miałyby 2 punkty itd. Dla każdego z czterech typów skalarnych moglibyśmy na podstawie ustalonych wyników trafnie przewidzieć wszystkie faktyczne odpowiedzi udzielone przez wszystkich respondentów.

Typy mieszane w tabeli stanowią jednak problem. Pierwszy typ mieszany (+ – +) na podstawie indeksu otrzymał tylko 1 punkt, co wskazywało na jedną odpowiedź pozytywną. Jeśli jednak 1 punkt zostałby przyznany na podstawie skali, przewidywalibyśmy, że 15 respondentów w tej grupie wybrało tylko najłatwiejsze stwierdzenie (akceptację aborcji w sytuacji zagrożenia zdrowia matki), przez co popełnilibyśmy dwa błędy w przypadku każdego takiego respondenta. Punkty na skali są więc przyznawane w taki sposób, by zminimalizować ewentualne błędy w rekonstrukcji pierwotnych odpowiedzi.

Tabela 6.3 przedstawia wyniki, które zostałyby przypisane każdemu z wzorców odpowiedzi w naszym przykładzie na podstawie indeksu i na podstawie skali. Zauważ, że w przypadku każdego respondenta z typu mieszanego popełniamy co najmniej jeden błąd. Jest to minimum, do którego możemy dążyć w przypadku typów mieszanych. Na przykład w pierwszym typie mieszanym przewidywalibyśmy mylnie, że każdy z 15 respondentów w tej grupie odpowiedział pozytywnie na pytanie o najłatwiejszy wariant, co dałoby w sumie 15 błędów.

To, w jakim stopniu zbiór danych empirycznych układa się w skalę Guttmana, zależy od stopnia poprawności, z jaką pierwotne odpowiedzi dają się odtworzyć na podstawie wyników uzyskanych na skali. Dla każdego z 1258 respondentów w tych badaniach przewidywalibyśmy trzy odpowiedzi, co daje w sumie 3774

Tabela 6.3

Punktacja indeksu i skali

	Wzorzec odpowiedzi	Liczba przypadków	Wynik na podstawie indeksu	Wynik na podstawie skali*	Błędy skali ogółem
Typy skalarne	+ + +	612	3	3	0
	+ + –	448	2	2	0
	+ – –	92	1	1	0
	– – –	79	0	0	0
Typy mieszane	– + –	15	1	2	15
	+ – +	5	2	3	5
	– – +	2	1	0	2
	– + +	5	2	3	5
				błędy skali ogółem:	27

$$\text{Współczynnik odtwarzalności} = 1 - \frac{\text{liczba błędów}}{\text{liczba prób przewidywania}} = 1 - \frac{27}{1258 \times 3} = \frac{27}{3774} = 0{,}993 = 99{,}3\%$$

* Tabela ta przedstawia jedną z popularnych metod ustalania punktacji typów mieszanych, powinieneś jednak pamiętać, że w użyciu są również inne metody.

przewidywania. Tabela 6.3 wykazuje, że używając wyników uzyskanych na skali popełnimy 27 błędów. Odsetek poprawnych przewidywań nosi nazwę *współczynnika odtwarzalności*: odsetek pierwotnych odpowiedzi, które możemy odtworzyć, znając wyniki uzyskane na skali, stanowiące ich podsumowanie. W omawianym przykładzie współczynnik odtwarzalności wynosi 3747/3774 lub 99,3%.

Z wyjątkiem przypadków pełnej (100%) odtwarzalności nie możemy nigdy stwierdzić, że dany zbiór pytań stanowi skalę Guttmana w jakimkolwiek sensie bezwzględnym. Niemal wszystkie zbiory tego typu pytań zbliżają się do skali. Generalnie należy jednak kierować się powszechnie przyjętą zasadą, że standard stanowią współczynniki od 90% do 95%. Jeśli zaobserwowana odtwarzalność przekracza ustalony poziom, zechcesz prawdopodobnie wypunktować i wykorzystać dany zbiór pytań jako skalę.

Decyzja co do kryteriów jest w tym względzie oczywiście arbitralna. Co więcej, wysoki stopień odtwarzalności nie gwarantuje, że skala mierzy naprawdę badane pojęcie. Jedyna korzyść z wysokiego współczynnika odtwarzalności to większa pewność, że wszystkie pozycje składowe mierzą *to samo*. Należy też zdawać sobie sprawę, że prawdopodobieństwo uzyskania wysokiego współczynnika odtwarzalności jest większe przy mniejszej liczbie pytań.

Na koniec zauważmy, że skalowanie Guttmana opiera się na obserwacji struktury faktycznych danych. Fakt ten, bardzo istotny, często jest mylnie rozumiany. Nie ma sensu mówić, że zbiór pytań kwestionariusza (być może opracowany i wykorzystany wcześniej przez innego badacza) stanowi skalę Guttmana. Powinniśmy raczej mówić, że stanowią one skalę w obrębie konkretnego zestawu analizowanych danych. Skalowalność jest więc cechą danych zależną od próby, empiryczną. Zestaw pytań może stanowić skalę Guttmana na przykład dla jednej próby respondentów w sondażu, nie daje to jednak gwarancji, że zestaw ten okaże się skalą dla innej próby. W tym sensie zestaw pytań kwestionariusza sam w sobie nigdy nie tworzy skali, natomiast zbiór obserwacji empirycznych niekiedy owszem.

Zamykamy w ten sposób rozważania na temat indeksowania i skalowania. Podobnie jak indeksy, skale są złożonymi miernikami zmiennych, umożliwiającymi na ogół rozszerzenie znaczenia zmiennej poza zakres pojedynczego wskaźnika. Zarówno skale, jak i indeksy służą pomiarowi zmiennej na poziomie porządkowym. Skale jednak, inaczej niż indeksy, wykorzystują każdą strukturę intensywności, jaką można wykryć wśród pojedynczych wskaźników. W takim zakresie, w jakim ową strukturę intensywności można wykryć, a jej logika daje się pogodzić z danymi uzyskiwanymi od ludzi czy innych jednostek analizy, możemy być pewni, że stworzyliśmy miernik porządkowy.

Typologie

Zamkniemy ten rozdział krótkim omówieniem tworzenia i analizy typologii. Przypomnij sobie, że indeksy i skale tworzymy, by uzyskać miernik porządkowy danej zmiennej. Przypisujemy poszczególnym przypadkom punkty na skali lub indeksie w taki sposób, by wskaźnikowały wzrastający poziom uprzedzeń, religijności, konserwatyzmu itd. W takich przypadkach mamy do czynienia z pojedynczymi wymiarami.

Często jednak badacz postanawia podsumować skrzyżowanie dwóch lub więcej zmiennych, tworząc w ten sposób zbiór kategorii lub typów – zmienną nominalną – zwany **typologią**. Można by na przykład badać orientację polityczną gazet oddzielnie w kategoriach spraw wewnętrznych i polityki zagranicznej. Czteropolowa tabela 6.4 przedstawia taką właśnie typologię.

Gazety w polu A są konserwatywne zarówno w polityce zagranicznej, jak i wewnętrznej; gazety w polu D są w obu tych dziedzinach polityki liberalne. W polach B i C znajdują się gazety konserwatywne w jednej z tych dziedzin i liberalne w drugiej.

Będziesz się często natykał na typologie, próbując stworzyć indeks lub skalę. Nagle okazuje się, że pytania, które twoim zdaniem miały odzwierciedlać pojedynczą zmienną, tak naprawdę odzwierciedlają dwie. Być może próbowaliśmy stworzyć pojedynczy indeks orientacji politycznej gazet, lecz odkryliśmy empirycznie, że należy rozdzielić politykę wewnętrzną i zagraniczną.

Tabela 6.4

Typologia polityczna gazet

		Polityka zagraniczna	
		konserwatywna	liberalna
Sprawy wewnętrzne	konserwatywna	A	B
	liberalna	C	D

Powinieneś być zawsze świadom trudności analizy typologicznej. Jeśli używasz typologii jako zmiennej niezależnej, prawdopodobnie nie będzie problemu. W powyższym przykładzie mógłbyś obliczyć w każdym polu odsetki gazet na ogół popierających kandydatów demokratycznych, następnie z łatwością mógłbyś sprawdzić wpływ orientacji politycznej w zakresie polityki wewnętrznej i zagranicznej na poparcie polityczne.

Niezwykle trudno jest jednak analizować typologię jako zmienną zależną. Gdybyś chciał odkryć, *dlaczego* gazety znajdują się w różnych polach typologii, miałbyś problem. Staje się on oczywisty, kiedy zastanowimy się, w jaki sposób mógłbyś tworzyć i odczytywać tabele. Załóżmy na przykład, że chciałbyś zbadać wpływ wielkości miejscowości na poglądy polityczne gazety. Dysponując jednym wymiarem, z łatwością mógłbyś przeliczyć odsetki gazet wiejskich i miejskich ocenionych na podstawie twojej skali lub indeksu jako konserwatywne lub liberalne.

W przypadku typologii musiałbyś jednak przedstawić rozkład gazet wiejskich w twojej próbie między typami A, B, C i D. Następnie musiałbyś powtórzyć tę procedurę dla gazet miejskich i porównać oba rozkłady. Załóżmy, że 80% gazet wiejskich i 30% miejskich zostałoby zaszeregowanych do typu A (konserwatywny w obu wymiarach). Załóżmy też, że tylko 5% gazet wiejskich – a 40% miejskich – należy do typu B (konserwatywny tylko w sprawach wewnętrznych). Niepoprawne byłoby wnioskowanie na podstawie typu B, że gazety miejskie są bardziej niż wiejskie konserwatywne w sprawach wewnętrznych, ponieważ cechę tę wykazuje 85% gazet wiejskich, a tylko 70% miejskich. Względna niedoreprezentacja gazet wiejskich w typie B spowodowana jest ich koncentracją w typie A. Jasne jest, że interpretacja tego typu danych w celach innych niż opisowe byłaby niezwykle trudna.

W rzeczywistości zbadałbyś prawdopodobnie każdy wymiar oddzielnie, zwłaszcza gdyby zmienna zależna miała więcej kategorii odpowiedzi niż w podanym przykładzie.

Nie myśl, że w badaniach społecznych należy zawsze unikać typologii – często stanowią one najbardziej użyteczne narzędzie wyjaśniania danych. Aby zbadać w sposób pogłębiony poglądy określone jako „pro-life" (dosłownie: „za życiem" – przyp. red. nauk.), mógłbyś stworzyć typologię dotyczącą zarówno aborcji, jak i kary śmierci. Libertarianizm można badać zarówno w kategoriach permisywizmu społecznego, jak i ekonomicznego. Powinieneś jednak być świadom szczególnych trudności, pojawiających się przy wykorzystaniu typologii jako zmiennych zależnych.

Główne zagadnienia

- Pojedynczy wskaźnik zmiennej rzadko uwzględnia wszystkie wymiary pojęcia i rzadko jest dostatecznie trafny, by uzasadnić jego użycie. Rzadko też zapewnia dostateczny zakres zmienności, by umożliwić wprowadzenie uszeregowania porządkowego. Mierniki złożone, takie jak skale i indeksy, rozwiązują te problemy, gdyż łączą kilka wskaźników zmiennej w jeden syntetyczny miernik.
- Mimo że zarówno indeksy, jak i skale są w zamierzeniu porządkowymi miernikami zmiennych, skale na ogół spełniają to zadanie lepiej niż indeksy.

- Indeksy opierają się na prostej kumulacji wskaźników zmiennej, skale natomiast wykorzystują każdą logiczną lub empiryczną strukturę intensywności wśród wskaźników danej zmiennej.
- Główne etapy tworzenia indeksu to wybór możliwych pytań, zbadanie zależności empirycznych między nimi, ustalenie punktacji indeksu i sprawdzenie jego trafności.
- Do kryteriów wyboru pytań zaliczamy trafność fasadową, jednowymiarowość, stopień szczegółowości planowanego pomiaru danego aspektu i zakres zmienności zapewniany przez dany zbiór pytań.
- Jeśli różne pytania są naprawdę wskaźnikami tej samej zmiennej, powinny one być ze sobą empirycznie powiązane. Tworząc indeks badacz musi sprawdzić zależności dwu- i wielozmiennowe między pytaniami.
- Punktacja indeksu wymaga podjęcia decyzji co do pożądanego zakresu punktów i ustalenia, czy poszczególne pytania mają mieć równe czy też różne wagi.
- Istnieją różne techniki umożliwiające wykorzystanie pytań w indeksie pomimo braków danych.
- Analiza pytań jest rodzajem walidacji wewnętrznej. Opiera się ona na zależności między pojedynczymi pytaniami wchodzącymi w skład złożonego miernika a samym tym miernikiem. Walidacja zewnętrzna odwołuje się do zależności między miernikiem złożonym a innymi wskaźnikami zmiennej, nie uwzględnionymi w mierniku.
- Przykład czterech technik skalowania mogą stanowić odpowiednio: skala dystansu społecznego Bogardusa, narzędzie pomiaru różnego stopnia skłonności danej osoby do integracji z określoną klasą ludzi; skala Thurstone'a, technika wykorzystująca opinie sędziów do określenia stopnia intensywności poszczególnych wskaźników; skala Likerta, technika pomiarowa oparta na wykorzystaniu standaryzowanych kategorii odpowiedzi i skala Guttmana, metoda służąca do odkrywania i wykorzystywania empirycznej struktury intensywności między kilkoma wskaźnikami danej zmiennej. Skala Guttmana jest chyba najbardziej popularną techniką skalowania we współczesnych badaniach społecznych.
- Dyferencjał semantyczny to format pytania, w którym prosimy respondentów, by dokonali oceny pomiędzy dwiema skrajnościami, takimi jak „bardzo pozytywne" i „bardzo negatywne".
- Typologia to nominalny miernik złożony, stosowany często w badaniach społecznych. Typologie mogą być skutecznie wykorzystywane jako zmienne niezależne, lecz wykorzystanie ich jako zmiennych zależnych powoduje trudności interpretacyjne.

Kluczowe pojęcia

indeks
skala
analiza pytań
walidacja zewnętrzna
skala dystansu społecznego Bogardusa
skala Guttmana
skala Thurstone'a
skala Likerta
dyferencjał semantyczny
typologia

Pytania i ćwiczenia

1. Opisz własnymi słowami różnicę między indeksem a skalą.
2. Załóżmy, że chciałbyś stworzyć indeks oceny jakości kształcenia na wyższych uczelniach. Wymień trzy pytania, które mogłyby się znaleźć w takim indeksie.
3. Wymyśl trzy pytania kwestionariusza mierzące postawy wobec energii jądrowej, które składałyby się prawdopodobnie na skalę Guttmana.
4. Stwórz typologię postaw „pro-life", omówioną w tym rozdziale.
5. Ekonomiści używają często indeksów do pomiaru zmiennych ekonomicznych, takich jak koszty utrzymania. Wejdź na stronę Biura Statystyk Pracy (http://stats.bls.gov) i znajdź sondaż dotyczący Indeksu Cen Konsumenta (*Consumer Price Index*). Jakie na przykład wymiary kosztów utrzymania uwzględnia ten miernik?

Dodatkowe lektury

Anderson Andy B., Basilevsky Alexander, Hum Derek P. J., *Measurement: Theory and Techniques*, w: *Handbook of Social Research*, Peter H. Rossi, James D. Wright, Andy B. Anderson (red.), Academic Press, New York 1983, s. 231–287. Analiza logiki pomiaru w kontekście mierników złożonych.

Bobo Lawrence, Licari Frederick C., *Education and Political Tolerance: Testing the Effects of Cognitive Sophistication and Target Group Effect*, „Public Opinion Quarterly" 1989, 53, s. 285–308. Autorzy wykorzystują wiele rozmaitych technik, by określić, jak można najlepiej mierzyć tolerancję wobec różnych grup społecznych.

Indrayan A., Wysocki M. J., Chawla A., Kumar R., Singh N., *Three-Decade Trend in Human Development Index in India and Its Major States*, „Social Indicators Research" 1999, 1, s. 91–120. Autorzy wykorzystują kilka indeksów rozwoju ludności, by porównać sytuację w różnych stanach indyjskich.

Lazarsfeld Paul, Passanella Ann, Rosenberg Morris, *Continuities in the Language of Social Research*, Free Press, New York 1972. Zapoznaj się przede wszystkim z częścią 1. Świetny zbiór dyskusji pojęciowych i konkretnych przykładów. Tworzenie mierników złożonych zostało tu przedstawione w ogólniejszym kontekście konceptualizacji i pomiaru.

Mclver John P., Carmines Edward G., *Unidimensional Scaling*, Sage, Newbury Park 1981. Świetny sposób, by lepiej poznać skalowanie Thurstone'a, Likerta i Guttmana.

Miller Delbert, *Handbook of Research Design and Social Measurement*, Sage, Newbury Park 1991. Świetny zbiór często stosowanych i na wpół standaryzowanych skal. Liczne przykłady z części 4 książki Millera dają się wprost zastosować w badaniach, w każdym zaś razie mogą służyć jako sugestia co do modyfikacji mierników. Co więcej, przestudiowanie kilku różnych przykładów może ułatwić ci ogólne zrozumienie logiki mierników złożonych.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 7

Logika doboru próby

Przegląd holograficzny

Jak to się dzieje, że badacze społeczni wybierają do rozmowy niewielką liczbę ludzi, a dokonują odkryć odnoszących się do setek milionów, których nie objęli obserwacją? Logika doboru próby dostarcza odpowiedzi na to pytanie.

Wprowadzenie

Krótka historia doboru próby
Prezydent Alf Landon
Prezydent Thomas E. Dewey
Dwa typy doboru próby

Próby nieprobabilistyczne
Dobór oparty na dostępności badanych
Dobór celowy lub arbitralny
Metoda kuli śnieżnej
Dobór kwotowy
Wybór informatorów

Teoria i logika doboru losowego
Świadome i nieświadome obciążenie próby
Reprezentatywność i prawdopodobieństwo wyboru
Dobór losowy
Rachunek prawdopodobieństwa, rozkład z prób i szacowanie błędu z próby

Populacje i operaty losowania
Przegląd populacji i operatów losowania

Typy doboru próby
Prosty dobór losowy
Dobór systematyczny
Dobór warstwowy
Ukryty podział na warstwy w doborze systematycznym
Przykład: dobór próby studentów uniwersytetu

Dobór próby wielostopniowy grupowy
Schematy doboru wielostopniowego a błąd z próby
Warstwowanie w wielostopniowej próbie grupowej
Dobór proporcjonalny
Dobór nieproporcjonalny i ważenie
Przykład: dobór próby kobiet chodzących do kościoła

Podsumowanie probabilistycznego doboru próby

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

W listopadzie 1996 roku demokrata Bill Clinton został ponownie wybrany w powszechnym głosowaniu na prezydenta USA. Clinton zdobył 49% głosów, republikanin Robert Dole zaś uzyskał 41%. Trzeci kandydat, Ross Perot zdobył 9%. Przed wyborami wiele sondaży przewidziało to zwycięstwo Clintona. W tabeli 7.1 przytoczono wyniki kilku ogólnokrajowych sondaży, przeprowadzonych na krótko przed wyborami, pod koniec października i na początku listopada. Dla zapewnienia porównywalności przydzieliłem „niezdecydowanych" wyborców poszczególnym kandydatom proporcjonalnie do liczby głosów oddanych na nich przez tych, którzy byli zdecydowani. Mimo że sondaże z reguły przeceniały liczbę wyborców Clintona w 1996 roku, to jednak większość z nich nie odbiegała o więcej niż 2–3% od faktycznego wyniku.

sowań. Zatem decyzja o tym, co obserwować, a czego nie obserwować, stanowi krytyczną część badań w naukach społecznych. Jeśli na przykład zamierzasz badać wyborców, to których wyborców weźmiesz pod uwagę?

Proces wyboru obserwacji nazywa się *doborem próby*. I chociaż dobór próby może oznaczać każdą procedurę wyboru jednostek obserwacji – na przykład ankietowanie co dziesiątego przechodnia na ruchliwej ulicy – kluczem do uogólniania wniosków z próby na szerszą populację jest *probabilistyczny dobór próby*[1], powiązany z ważną ideą *doboru losowego*.

Dużo miejsca w tym rozdziale zajmuje logika i warsztat doboru probabilistycznego. Temat ten podlega ściślejszym rygorom i wymaga większej precyzji niż niektóre inne tematy omówione w tej książce. Badania społeczne jako całość należą zarówno do nauk humanistycznych, jak i ścisłych, dobór próby natomiast skła-

Tabela 7.1

Sondaże przedwyborcze prognozujące wyniki wyborów prezydenckich

		Procent głosów oddanych na			
Data	Agencja	Clintona	Dole'a	Perota	Innych
28–31.10	Hotline/Battleground	49	40	9	2
30.10–2.11	CBS/New York Times	54	35	9	2
31.10–3.11	Pew Research Center	52	38	8	2
1–3.11	Reuter/Zogby	49	41	8	2
1–3.11	Harris	51	39	9	1
2–3.11	ABC	52	39	7	2
2–3.11	NBC/Wall St. Journal	51	38	9	2
3–4.11	Gallup/CNN/USAToday	51	38	9	2
	Wyniki wyborów	**49**	**41**	**9**	**2**

Źródło: Robert M. Worcester, *Reporting the Polls: You Can Do Better*, „Public Perspective" grudzień–styczeń 1997, s. 54.

Jak sądzisz, ile każdy z ośrodków badawczych przeprowadził wywiadów, aby zmieścić się w tych kilku procentach błędu przy przewidywaniu zachowania około 90 mln wyborców? Mniej niż dwa tysiące! Niniejszy rozdział wyjaśnia, w jaki sposób badacze dokonują takich sztuczek.

Badania zachowań politycznych, a także inne formy badań społecznych opierają się na obserwacji. Ale ani badacze zachowań politycznych, ani inni badacze społeczni nie mogą obserwować wszystkiego, co dotyczy ich zainte-

[1] W polskim tłumaczeniu określenie „próba losowa" stosowane jest zarówno dla angielskiego określenia „probability sample", jak i dla (nielicznych) przypadków używania przez autora określenia „random sample". W niektórych przypadkach, kiedy chodziło o szerszą klasę metod losowego doboru próby, a zwłaszcza kiedy podkreślana była odrębność tych metod w stosunku do metod nieprobabilistycznych, zastosowano tłumaczenie „probability sampling" jako „probabilistyczny dobór próby" (przyp. tłum.).

nia się bardziej w stronę nauk ścisłych. Mimo że temat ten jest nieco techniczny, podstawową logikę doboru próby nietrudno zrozumieć. Tak naprawdę logiczna zgrabność tego zagadnienia sprawia, że można je łatwiej zrozumieć niż, powiedzmy, konceptualizację.

Mimo że probabilistyczny dobór próby zajmuje centralne miejsce w dzisiejszych badaniach społecznych, poświęcimy nieco czasu także nieprobabilistycznym metodom doboru próby. Metody te mają własną logikę i mogą być całkiem przydatne dla osiągnięcia niektórych celów badawczych.

Zanim omówimy te dwa główne rodzaje doboru próby, wprowadźmy kilka podstawowych idei poprzez krótką historię doboru próby. Jak się przekonamy, badacze, którzy poprawnie przewidzieli zwycięstwo Billa Clintona w głosowaniu powszechnym w 1996 roku, częściowo zawdzięczają to temu, że nauczyli się unikać pewnych pułapek, w które wpadali ich poprzednicy.

Krótka historia doboru próby

Dobór próby w naukach społecznych rozwijał się równolegle z badaniami zachowań i preferencji politycznych. Działo się tak z całą pewnością dlatego, że badania zachowań politycznych stanowią jedną z niewielu okazji do przekonania się o trafności szacunków. W dniu wyborów badacze społeczni przekonują się, na ile ich badania były trafne lub nie.

Prezydent Alf Landon

Prezydent Alf Landon? Kto to? Czyżbyś przespał całą jedną prezydenturę na lekcji historii Stanów Zjednoczonych? Nie – ale Alf Landon *byłby* prezydentem, gdyby słynny sondaż przeprowadzony przez „Literary Digest" okazał się trafny. „Literary Digest" był popularnym magazynem informacyjnym, wydawanym w latach 1890–1938. W 1920 roku wydawcy „Literary Digest" rozesłali kartki pocztowe do mieszkańców sześciu stanów, zapytując, na kogo zamierzają głosować w wyborach prezydenckich – czy na Warrena Hardinga, czy na Jamesa Coxa. Nazwiska respondentów zostały wybrane z książek telefonicznych i z list rejestracyjnych samochodów. Na podstawie kart pocztowych otrzymanych z powrotem „Literary Digest" trafnie przewidział, że wybory wygra Harding. Przed następnymi wyborami „Literary Digest" zwiększył zasięg sondażu i sformułował trafne prognozy w latach 1924, 1928 i 1932.

W 1936 roku „Literary Digest" przeprowadził swój najambitniejszy sondaż: do ludzi wybranych z książek telefonicznych i z list rejestracyjnych samochodów wysłano dziesięć milionów kart. Odpowiedziały ponad dwa miliony ludzi, dając w skali całego kraju znaczącą przewagę 57% głosów kandydatowi republikanów Alfowi Landonowi nad jego kontrkandydatem, urzędującym prezydentem Franklinem Rooseveltem, który otrzymał 43%. Wydawcy zastrzegali się skromnie:

> „Nie rościmy sobie pretensji do nieomylności. Nie my ukuliśmy zwrot «niesamowita dokładność», którym tak swobodnie szafowano w stosunku do naszych sondaży. Aż nadto dobrze znamy ograniczenia każdego próbnego głosowania, bez względu na wielkość próby i naukowość metody. Zakrawałoby na cud, gdyby w dniu wyborów każdy z 48 stanów zachował się tak, jak przewiduje nasz sondaż".
>
> (*Landon* 1936a, s. 6)

Dwa tygodnie później wydawcy „Literary Digest" uświadomili sobie jeszcze lepiej ograniczenia próbnych głosowań: wyborcy wybrali Roosevelta na drugą kadencję największą przewagą głosów w historii – otrzymał 61% głosów wyborców. Landon otrzymał jedynie 8 głosów elektorskich w stosunku do 523 głosów na Roosevelta.

Wydawcy byli zdumieni tym nieszczęsnym odwróceniem losu. Jak to się stało, że sondaż był tak błędny? Częściowo problem związany był z pewnością z 22-procentowym poziomem zwrotów uzyskanym w sondażu. Wydawcy zapytywali:

> „Dlaczego w Chicago zaledwie co piąty wyborca, do którego «Literary Digest» zwrócił się

z ankietą, zadał sobie trud odpowiedzi? I dlaczego republikanie mieli przewagę wśród tej jednej piątej wyborców, którzy odpowiedzieli? Doświadczyliśmy lepszej współpracy w zakresie tego, co wydawało nam się zawsze współpracą dla dobra ogólnego ze strony wyborców republikańskich niż ze strony wyborców demokratycznych. Czy republikanie mieszkają bliżej skrzynek pocztowych? Czy demokraci są ogólnie nieprzychylni próbnym wyborom?"

(*What Went Wrong* 1936b, s. 7)

Tymczasem istniało lepsze wytłumaczenie – w postaci tego, co technicznie nazywa się *operatem doboru próby* użytym przez „Literary Digest". W tym przypadku operat doboru próby składał się z abonentów telefonów i właścicieli samochodów. W warunkach 1936 roku taka konstrukcja próby daje nadreprezentację zamożnych wyborców, zwłaszcza pod sam koniec najgorszej zapaści gospodarczej w historii kraju. Próba ta w rezultacie pomijała ludzi biednych, ci zaś w większości głosowali na program reform Nowy Ład (*New Deal*) Roosevelta. Sondaż „Literary Digest" mógł – poprawnie lub nie – przewidywać zamiary wyborcze abonentów telefonów i właścicieli samochodów. Na nieszczęście wydawców zdecydowanie nie odzwierciedlał on zamiarów wyborczych całej populacji.

Prezydent Thomas E. Dewey

Podczas owych wyborów w 1936 roku na scenie pojawił się młody badacz, którego nazwisko stało się później synonimem opinii publicznej. W przeciwieństwie do „Literary Digest", George Gallup trafnie przewidział, że Roosevelt wygra z Landonem. U podstaw sukcesu Gallupa w 1936 roku leżało użycie przez niego czegoś, co nazywa się *próbą kwotową*[2], której przyjrzymy się dokładniej w dalszej części tego rozdziału. Na razie wystarczy wiedzieć, że kwotowy dobór próby opiera się na znajomości określonych cech populacji, z której dobiera się próbę: jaki jest w niej udział mężczyzn i kobiet, jaki jest udział różnych grup dochodu, wieku itd. Ludzie dobierani są do próby kwotowej tak, by zachować proporcje tych cech występujące w całej populacji: odpowiednia liczba biednych białych mężczyzn mieszkających na wsi; odpowiednia liczba bogatych Afroamerykanek zamieszkujących miasta itd. Kwoty oparte są na cechach, które mają największy związek z badaniami. W przypadku sondażu Gallupa dobór próby opierał się na poziomie dochodu; procedura selekcji zapewniała właściwy udział respondentów o różnym poziomie dochodu.

Gallup i jego Amerykański Instytut Opinii Publicznej stosowali z powodzeniem próby kwotowe w latach 1936, 1940 i 1944 – za każdym razem trafnie wskazując zwycięzcę wyborów prezydenckich. I wówczas, w 1948 roku, Gallup i większość badaczy opinii publicznej przeżyli wstyd przewidując, że gubernator Nowego Jorku Thomas Dewey wygra w wyborach z urzędującym prezydentem Harrym Trumanem. Ów żenujący błąd badaczy trwał aż do wieczoru wyborczego. Słynna fotografia pokazuje świętującego zwycięstwo Trumana – którego zwolennicy mieli okrzyk bojowy „dołóż im, Harry!" – trzymającego uniesioną gazetę z ogromnym nagłówkiem „Dewey zwycięża Trumana".

Do porażki sondaży w 1948 roku przyczyniło się kilka czynników. Przede wszystkim większość badaczy zakończyła sondaże na początku października, pomimo stałej tendencji do zwiększania się liczby zwolenników Trumana podczas kampanii. Poza tym wielu wyborców było niezdecydowanych przez całą kampanię, a potem oddali nieproporcjonalnie więcej głosów na Trumana, gdy przyszli do lokali wyborczych.

Co ważniejsze, porażka Gallupa polegała na niereprezentatywności jego prób. Próba kwotowa – która była skuteczna w poprzednich latach – zgubiła go w 1948 roku. Technika ta wymaga, by badacz wiedział cokolwiek o całej populacji – w tym przypadku o populacji wyborców. Dla ogólnokrajowych sondaży informacje takie czerpano przede wszystkim z danych ze spisu powszechnego. Jednak przed 1948 rokiem II wojna

[2] Próba kwotowa występuje niekiedy w polskiej literaturze badawczej pod nazwą „próba udziałowa" (przyp. tłum.).

światowa spowodowała masowe migracje ze wsi do miast, co zmieniło radykalnie charakter populacji Stanów Zjednoczonych w stosunku do tego, co pokazywały dane ze spisu powszechnego z 1940 roku. Na danych z tego właśnie spisu opierał się Gallup. Co więcej, mieszkańcy miast skłaniali się ku demokratom, zatem nadreprezentacja wiejskich wyborców w sondażu Gallupa przyniosła w efekcie niedoszacowanie liczby głosów oddanych na demokratów.

Dwa typy doboru próby

Już przed 1948 rokiem niektórzy badacze akademiccy eksperymentowali z pewną formą doboru próby, opartą na rachunku prawdopodobieństwa. U podstaw tej techniki leży dobranie „próby losowej" z listy zawierającej nazwiska wszystkich osób z badanej populacji. Generalnie, probabilistyczne metody doboru próby użyte do badań w 1948 roku okazały się znacznie trafniejsze niż techniki kwotowego doboru próby.

W dzisiejszych czasach probabilistyczny dobór próby jest nadal podstawową metodą dobierania dużych reprezentatywnych prób do badań społecznych, w tym do ogólnokrajowych sondaży wyborczych. Jednocześnie jednak może to być metoda niemożliwa lub niewłaściwa do zastosowania w wielu sytuacjach badawczych. Zatem zanim zwrócimy się ku logice i technikom probabilistycznego doboru próby, przyjrzymy się technikom doboru nieprobabilistycznego i ich zastosowaniu w naukach społecznych.

Próby nieprobabilistyczne

Badania społeczne prowadzone są często w sytuacjach, w których niemożliwe jest dobranie prób probabilistycznych tego rodzaju, jakie stosuje się w dużych sondażach. Wyobraźmy sobie, że ktoś chce badać bezdomnych: nie istnieje lista osób bezdomnych i jest mało prawdopodobne, że zdołamy ją stworzyć. Co więcej, jak się przekonamy, istnieją sytuacje, w których dobór probabilistyczny nie byłby właściwy, nawet gdyby był możliwy. Wiele z takich sytuacji wymaga **doboru nieprobabilistycznego**.

W tym rozdziale zajmiemy się czterema typami doboru nieprobabilistycznego: doborem opartym na dostępności badanych, doborem celowym lub arbitralnym, metodą kuli śnieżnej oraz doborem kwotowym. Zakończymy krótkim omówieniem technik zdobywania informacji o grupach społecznych poprzez informatorów.

Dobór oparty na dostępności badanych

Zdawanie się na dostępność przedmiotów badania dobieranych do próby, czyli na przykład dobór próby poprzez zatrzymywanie ludzi na ulicy albo w innym miejscu, to bardzo ryzykowna metoda; mimo to w praktyce jest ona stosowana aż nazbyt często. Metoda ta w sposób oczywisty nie dopuszcza żadnej kontroli nad reprezentatywnością próby. Może być ona uzasadniona tylko wtedy, gdy badacz chce zbadać cechy osób przechodzących przez dane miejsce o określonych porach lub gdy nie ma możliwości zastosowania mniej ryzykownych metod doboru próby. Nawet jeśli przyczyną zastosowania tej metody jest niemożność zastosowania innych metod, trzeba zachować wielką ostrożność w uogólnianiu wniosków z tak uzyskanych danych. Należy również uczulić odbiorców wyników na ryzyko związane z tą metodą.

Badacze akademiccy często prowadzą sondaże wśród studentów zapisanych na wykłady w dużych grupach. Poręczność i niskie koszty takiej metody wyjaśniają jej popularność, ale rzadko kiedy dostarcza ona danych o jakiejś ogólniejszej wartości. Może być przydatna do pilotażu kwestionariusza, ale nie powinna być używana do badań rzekomo opisujących studentów jako całość.

Rozważmy poniższy fragment raportu, dotyczący doboru próby w badaniach na temat stanu wiedzy i opinii studentów medycyny oraz lekarzy rodzinnych o odżywianiu w chorobach nowotworowych:

„Populacja[3] studentów w tych badaniach składała się ze studentów IV roku medycyny Wydziału Medycznego (Medical School) Uniwersytetu Minnesota w Minneapolis. Populację lekarzy stanowili wszyscy lekarze uczęszczający na kurs «Praktyka lekarza rodzinnego. Przegląd i aktualności», sponsorowany przez Wydział Ustawicznego Kształcenia Medycznego Uniwersytetu Minnesota".

(Cooper-Stephenson, Theologides 1981, s. 472)

Co będą reprezentować wyniki tych badań, tak jak możemy je sobie wyobrazić na podstawie powyższego opisu? Nie dostarczają one sensownego porównania między studentami medycyny a lekarzami rodzinnymi w USA ani nawet w Minnesocie. Kim byli lekarze uczęszczający na kurs? Możemy zgadywać, że zapewne byli oni bardziej zainteresowani swoim dalszym kształceniem niż inni lekarze, ale nie wiemy tego na pewno. Choć tego typu badania mogą być źródłem użytecznych informacji, musimy uważać, by nadmiernie nie uogólniać wyciąganych z nich wniosków.

Dobór celowy lub arbitralny

Niekiedy właściwe jest dobranie próby na podstawie własnej wiedzy o badanej populacji oraz celach badań. Ten typ doboru jest nazywany **próbą celową** (*purposive sampling*) lub **arbitralną** (*judgemental sampling*). Na przykład w początkowej fazie tworzenia kwestionariusza można chcieć dobrać jak najróżniejszych respondentów, by sprawdzić szerokie zastosowanie pytań. Chociaż wyniki takich badań nie reprezentowałyby żadnej sensownej populacji, przebieg testu mógłby z powodzeniem ujawnić jakieś specyficzne wady kwestionariusza. Sytuację taką należałoby jednak uznać raczej za pilotaż niż za badanie końcowe.

W niektórych przypadkach można chcieć zbadać mały podzbiór większej populacji, w której wielu członków tego podzbioru da się łatwo rozpoznać, ale sporządzenie wykazu ich wszystkich byłoby najprawdopodobniej niemożliwe. Weźmy dla przykładu badania przywódców jakiegoś studenckiego ruchu protestu: wielu liderów jest łatwo rozpoznawalnych, ale zidentyfikowanie ich wszystkich i dobranie z nich próby byłoby zadaniem nie do wykonania. Badając wszystkich najbardziej rzucających się w oczy przywódców albo ich próbę, można zgromadzić dane wystarczające dla celów badań.

Innym przykładem mogłyby być porównawcze badania studentów lewicowych i prawicowych. Ponieważ sporządzenie wykazu wszystkich takich studentów i dobranie z nich próby może być niemożliwe, można zdecydować się na dobór próby członków grup lewicowych i prawicowych, takich jak Partia Zielonych i ugrupowanie Młodzi Amerykanie dla Wolności. Chociaż tak skonstruowana próba nie pozwoli na dobry opis ani lewicowych, ani prawicowych studentów jako całości, może wystarczyć dla ogólnych celów porównawczych.

Badacze empiryczni są często szczególnie zainteresowani badaniem przypadków odbiegających od normy – takich, które nie pasują do ściśle określonych wzorców postaw czy zachowań – aby lepiej zrozumieć te bardziej regularne wzorce. Można na przykład uzyskać istotne informacje na temat charakteru szkoły, ujawnianego podczas zjazdu integracyjnego, poprzez przeprowadzenie wywiadów z tymi, którzy, jak się wydaje, nie ulegają powszechnie podzielanym emocjom, albo też z uczniami, którzy w ogóle nie brali udziału w tym zjeździe. Dobór przypadków odbiegających od normy też jest przykładem próby celowej.

Metoda kuli śnieżnej

Druga nieprobabilistyczna technika doboru próby, uważana przez niektórych za formę doboru przypadkowego, jest nazywana **metodą kuli śnieżnej** (*snowball sampling*). Ta procedura jest właściwa, gdy trudno jest odszukać członków jakiejś specyficznej populacji, takiej jak

[3] Wątpliwości budzi w omawianym fragmencie również użycie terminu „populacja" na określenie faktycznie przebadanej grupy osób, a nie szerszej zbiorowości, do której będą odniesione wnioski z badań, a której faktycznie przebadana grupa ma być tylko reprezentacją (przyp. tłum.).

bezdomni, robotnicy napływowi czy nielegalni imigranci. W metodzie kuli śnieżnej badacz zbiera dane o kilku członkach badanej populacji, których da się odszukać, a następnie prosi te osoby o dostarczenie informacji potrzebnych do odszukania innych członków tej populacji, których akurat znają. Określenie „kula śnieżna" odnosi się do procesu akumulacji, gdyż każda odszukana osoba podaje inne osoby. Jako że reprezentatywność prób wyłonionych za pomocą tej procedury można podać w wątpliwość, jest ona używana przede wszystkim do celów eksploracyjnych.

Powiedzmy, że chcemy poznać wzorce rekrutacji do jakiejś organizacji lokalnej w danym okresie. Można by zacząć od przeprowadzenia wywiadów z ostatnio zrekrutowanymi członkami, pytając ich o to, kto je wprowadził do grupy. Wtedy można by przeprowadzić wywiady z wymienionymi osobami, pytając *je* o to, kto je wprowadził do tej grupy. Następnie można by przeprowadzić wywiady z tymi wymienionymi osobami, pytając z kolei o to, kto *je* wprowadził. Albo też, badając luźno ustrukturyzowaną grupę polityczną, można by zapytać jedną z należących do niej osób o to, kogo uważa za najbardziej wpływowych członków grupy. Można by przeprowadzić wywiady ze wskazanymi osobami i zapytać podczas wywiadu, kogo one uważają za najbardziej wpływowego. W każdym z tych przykładów będzie to próba dobrana metodą kuli śnieżnej, gdyż każdy z badanych wskazuje innych, których można badać.

Dobór kwotowy

Kwotowy dobór próby jest metodą, która pomogła George'owi Gallupowi uniknąć katastrofy w 1936 roku – i spowodowała ową katastrofę w 1948 roku. Próba kwotowa, podobnie jak losowa, uwzględnia kwestię reprezentatywności, chociaż każda z tych dwóch metod podchodzi do niej zupełnie odmiennie.

W **próbie kwotowej** (*quota sampling*) wychodzi się od macierzy lub tabeli, opisującej istotne dla badania cechy populacji. W zależności od celów badania badaczowi może być potrzebna wiedza, jaki jest na przykład udział mężczyzn i kobiet w badanej populacji oraz jaki jest udział każdej płci w poszczególnych kategoriach wieku, poziomu wykształcenia, grupach etnicznych itd. Dla dobrania ogólnokrajowej próby kwotowej badaczowi może być potrzebna wiedza o tym, jaki procent ludności kraju stanowią na przykład ludzie o następujących cechach: mieszkańcy miast, zamieszkujący wschodni region kraju, płci męskiej, poniżej 25 lat, biali, zawodowo czynni itp., a także wszystkie możliwe kombinacje tych cech.

Gdy macierz jest już stworzona i dla każdej komórki tabeli określone zostały odpowiednie udziały procentowe, należy następnie zebrać dane od osób mających wszystkie cechy określone w danej komórce tabeli. Potem wszystkim osobom w danej komórce przypisuje się wagę odpowiednią do ich udziału w całej populacji. Gdy zostaną w ten sposób przeważone wszystkie elementy próby, całość danych powinna dać sensowną reprezentację całej populacji.

Mimo że próba kwotowa przypomina dobór losowy, to jednak wiąże się ona nieodłącznie z paroma problemami. Po pierwsze, operat kwotowania (udziały procentowe reprezentowane przez poszczególne komórki tabeli) musi być dokładny, podczas gdy często trudno jest znaleźć dla tego celu aktualne informacje. Porażka Gallupa, który nie przewidział zwycięstwa Trumana w wyborach prezydenckich w 1948 roku, częściowo wynikała z tego problemu. Po drugie, odchylenia od reprezentatywności mogą wynikać z doboru elementów próby w ramach danej komórki tabeli – nawet jeśli jej udział procentowy w całej populacji został poprawnie oszacowany. Mimo że ankieter został poinstruowany, że ma przeprowadzić wywiady z pięcioma osobami, charakteryzującymi się określonym złożonym zestawem cech, może się zdarzyć, że będzie on unikał robienia wywiadów z osobami mieszkającymi na ostatnich piętrach siedmiopiętrowych bloków bez wind, z mieszkańcami szczególnie zdewastowanych domów czy też z właścicielami złych psów.

W ostatnich latach czynione były próby łączenia metod probabilistycznych i kwotowych, ale na poznanie ich skuteczności pozostaje jeszcze poczekać. Na razie zalecałbym

ostrożne traktowanie próby kwotowej, gdy celem jest opis statystyczny.

Jednocześnie ze względu na swą logikę dobór kwotowy może być niekiedy z powodzeniem stosowany do badań terenowych. Na przykład w badaniach jakiejś grupy formalnej można przeprowadzić wywiady zarówno z jej przywódcami, jak i członkami nie będącymi przywódcami. Badając studencką organizację polityczną, można przeprowadzić wywiady ze skrajnymi, umiarkowanymi i konserwatywnymi członkami tej grupy. Można w tych przypadkach osiągnąć wystarczający stopień reprezentatywności, stosując dobór kwotowy, który zapewni objęcie badaniem zarówno mężczyzn, jak i kobiety, zarówno ludzi młodych, jak i starych itd.

Wybór informatorów

W sytuacjach, w których badania terenowe wymagają od badacza podjęcia próby zrozumienia jakiegoś układu społecznego – na przykład gangu młodzieżowego albo lokalnego sąsiedztwa – znaczną część takiego zrozumienia uzyskuje się poprzez współpracę z niektórymi członkami badanej grupy. W badaniach społecznych pojęcie respondentów odnosi się do ludzi, którzy dostarczają informacji o sobie, umożliwiając w ten sposób badaczowi stworzenie złożonego obrazu grupy, którą owi respondenci reprezentują, natomiast **informator** jest członkiem grupy, który może wypowiadać się bezpośrednio o grupie.

Informatorzy mają szczególne znaczenie dla antropologów, ale mogą być również wielce pomocni innym badaczom społecznym. Gdybyśmy na przykład chcieli zbadać nieformalne układy społeczne w jakimś budynku komunalnym, byłoby dobrze odszukać jednostki, które mogą zrozumieć, o co nam chodzi i pomóc znaleźć istotne informacje.

Kiedy Jeffrey Johnson (1990) podjął badania społeczności rybackiej specjalizującej się w połowach łososia, zastosował on kilka kryteriów do oceny potencjalnych informatorów. Dla przykładu, czy ich pozycja pozwala im wchodzić regularnie w interakcje z innymi, czy też są izolowani? (Stwierdził on, że stolarz miał szerszy zakres interakcji niż kapitan łodzi). Czy posiadane przez nich informacje o obozie ograniczają się w znacznym stopniu do ich specyficznych zajęć, czy też obejmują wiele aspektów całego przedsięwzięcia? Te i inne kryteria wpłynęły na ocenę przydatności potencjalnych informatorów.

Zazwyczaj badacz chce dobrać sobie informatorów, którzy byliby w jakiś sposób typowi dla badanych grup. W przeciwnym przypadku ich obserwacje i opinie mogłyby wprowadzać w błąd. Na przykład przeprowadzenie wywiadów wyłącznie z lekarzami nie dałoby całościowego oglądu pracy przychodni gminnej. Na tej samej zasadzie antropolog, który robi wywiady wyłącznie z mężczyznami w społeczności, w której kobiety chroni się przed obcymi, otrzyma zniekształcony obraz rzeczywistości. Albo też, mimo że informatorzy biegle władający angielskim są przydatni dla anglojęzycznych badaczy ze Stanów Zjednoczonych, nie są oni typowymi członkami wielu społeczeństw, a nawet wielu podgrup w krajach anglojęzycznych.

Z tego prostego powodu, że informatorzy chcą pracować z badaczami z zewnątrz, niemal zawsze będą oni nieco „marginalni" albo nietypowi w ramach swojej grupy. Niekiedy jest to oczywiste. W innych sytuacjach jednak można uświadomić sobie ich marginalną pozycję dopiero w trakcie badań.

W badaniach Johnsona lokalny urzędnik wskazał mu pewnego rybaka, który, jak się wydawało, zdecydowanie uczestniczył w głównym nurcie życia społeczności. Co więcej, był on pomocny i chętny do współpracy przy badaniach prowadzonych przez Johnsona. Jednak im dłużej Johnson pracował z tym rybakiem, tym bardziej przekonywał się, że jest on marginalnym członkiem tej rybackiej społeczności.

> „Po pierwsze, był jankesem w południowym miasteczku. Po drugie, otrzymywał emeryturę od marynarki wojennej (a więc nie był uważany za «poważnego rybaka» przez innych członków społeczności). Po trzecie, był znaczącym działaczem republikańskim w miasteczku z przewagą demokratów. Wreszcie, zakotwiczał swoją łódź osobno, z dala od wspólnego portu".
>
> (1990, s. 56)

Marginalna pozycja informatorów nie tylko może zafałszować obraz rzeczywistości, jaki otrzymamy, ale może ona również ograniczać ich (a więc i nasz) dostęp do różnych sektorów społeczności, którą chcemy badać.

Te uwagi powinny dać ci pewne wyczucie problemów, związanych z nieprobabilistycznym doborem próby. Zakończę je przytoczeniem następującego polecenia:

> „Twoim nadrzędnym celem jest zgromadzenie *jak najbogatszego zasobu danych*. W idealnym przypadku oznacza to szeroki i zróżnicowany zakres informacji, zebranych w dość długim czasie. W owych idealnych warunkach można to osiągnąć wchodząc w bezpośredni kontakt z określonym społecznym miejscem i okolicznościami, oraz zanurzając się w nie na dłuższy czas".
>
> (Lofland, Lofland 1995, s. 16)

Innymi słowy, nieprobabilistyczyny dobór próby ma swoje zastosowania, zwłaszcza w badaniach jakościowych. Jednak badacze muszą uważać i uwzględniać ograniczenia tego doboru, zwłaszcza w odniesieniu do poprawnych i precyzyjnych opisów populacji. Ta kwestia stanie się jaśniejsza w miarę omawiania logiki i warsztatu doboru losowego.

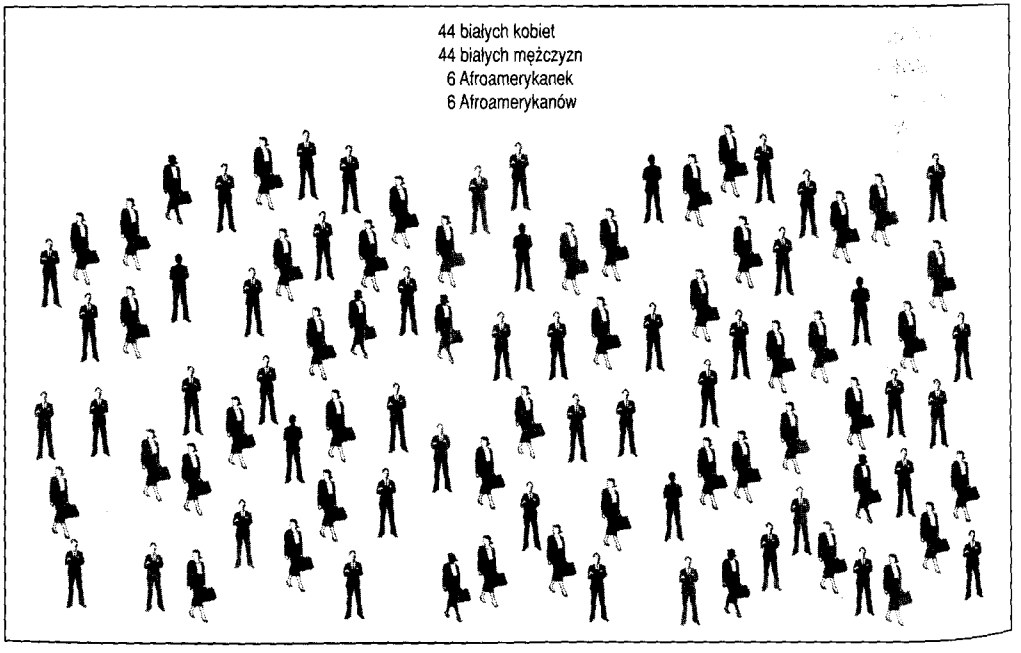
Teoria i logika doboru losowego

Dobór nieprobabilistyczny, choć poprawny przy niektórych celach badawczych, nie może dać gwarancji, że obserwowana przez nas próba jest reprezentatywna dla całej populacji. Gdy badacze chcą uzyskać dokładny opis statystyczny wielkich populacji – na przykład gdy chcą określić, jaki procent danej populacji stanowią bezrobotni, jaki procent zamierza głosować na kandydata X albo jaki procent uważa, że ofiara gwałtu powinna mieć prawo do aborcji – zwracają się ku **losowemu (probabilistycznemu) doborowi próby** (*probability sampling*). Wszystkie duże sondaże korzystają z metod doboru probabilistycznego.

Chociaż zastosowanie doboru losowego wiąże się z pewnymi wyszukanymi zastosowa-

Rysunek 7.1

Populacja 100 osób



niami statystyki, podstawową logikę tego doboru nietrudno zrozumieć. Gdyby wszyscy członkowie jakiejś populacji byli identyczni pod każdym względem – wszystkich cech demograficznych, postaw, doświadczeń, zachowań itd. – nie byłoby żadnej potrzeby stosowania starannych procedur doboru próby. W tym skrajnym przypadku doskonałej homogeniczności *dowolny pojedynczy* przypadek wystarczyłby jako próba do zbadania cech całej populacji.

Tak naprawdę, rzecz jasna, ludzie wchodzący w skład jakiejkolwiek rzeczywistej populacji są dość heterogeniczni, różnią się pod wieloma względami. Rysunek 7.1 ukazuje uproszczony obraz heterogenicznej populacji: 100 członków tej małej populacji różni się pod względem płci i rasy. Będziemy korzystać z tej mikropopulacji dla zilustrowania różnych aspektów doboru próby losowej.

Fundamentalna idea, leżąca u podstaw losowego doboru próby, jest następująca: aby próba jednostek z danej populacji mogła być podstawą dobrego opisu całej populacji, musi ona mieć zasadniczo taką samą zmienność jak populacja. Nie jest to jednak takie proste, jak mogłoby się wydawać. Poświęćmy chwilę na przyjrzenie się niektórym drogom, na których mogą zbłądzić badacze. Potem przekonamy się, jak dochodzi do tego, że próba losowa trafnie odzwierciedla zmienność populacji.

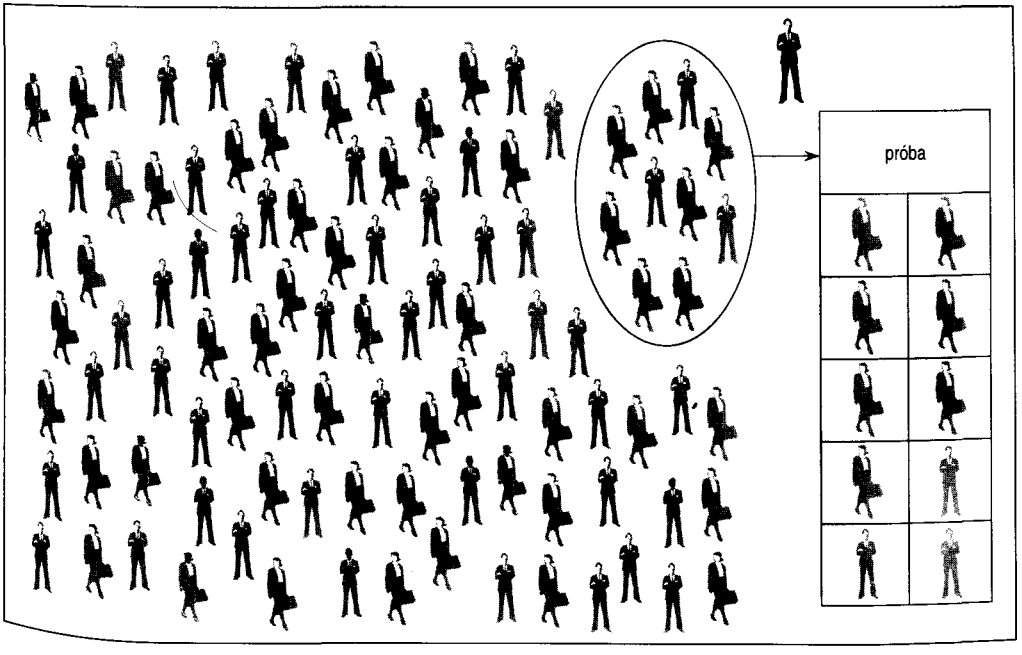
Świadome i nieświadome obciążenie próby

Na pierwszy rzut oka dobór próby może wydawać się całkiem prosty. Aby dobrać próbę 100 studentów, można iść do miasteczka akademickiego i przeprowadzić wywiady z pierwszymi spotkanymi 100 studentami. Taka metoda doboru próby jest często stosowana przez niewprawnych badaczy, ale wiąże się ona z bardzo dużym ryzykiem wprowadzenia do próby obciążeń.

Gdy mówimy o *obciążeniu* (*sampling bias*) w związku z doborem próby, oznacza to po prostu, że wybrane osoby nie są „typowe" czy „reprezentatywne" dla szerszej populacji, z której zostały dobrane. Ten rodzaj obciążenia nie musi być zamierzony. W rzeczy samej,

Rysunek 7.2

Próba oparta na dogodnej dostępności: łatwa, ale niereprezentatywna



właściwie nie można go uniknąć, gdy dobiera się ludzi z najbliższego otoczenia.

Rysunek 7.2 pokazuje, co może się zdarzyć, jeśli po prostu wybiera się osoby, które najłatwiej znaleźć do danych badań. Mimo że kobiety stanowią tylko 50% naszej mikropopulacji, podgrupa najbliżej badacza (w prawym górnym rogu) składa się w 70% z kobiet, a chociaż w populacji jest 12% Afroamerykanów, żadna z tych osób nie znalazła się w próbie.

Poza ryzykiem związanym z badaniem najłatwiej dostępnych osób mogą pojawić się także inne problemy. Zacznijmy od własnych skłonności badacza, które mogą oddziaływać na dobór próby tak bardzo, że nie będzie ona naprawdę reprezentatywna dla całej populacji. Przypuśćmy, że kogoś onieśmielają nieco studenci, którzy wyglądają jak gdyby mogli ośmieszyć jego trud badawczy. Świadomie czy nieświadomie, mógłby unikać przeprowadzania wywiadów z takimi ludźmi. Bądź też mógłby mieć wrażenie, że postawy takich studentów nie mają związku z jego badaniami, a więc unikałby robienia z nimi wywiadów.

Nawet jeśli ktoś starałby się przeprowadzać wywiady ze zrównoważoną grupą studentów, nie znałby dokładnych proporcji różnych typów studentów składających się na tę równowagę i nie zawsze mógłby określić te różne typy po prostu przyglądając się im z zewnątrz.

Gdyby nawet ktoś sumiennie podjął się realizowania wywiadów, powiedzmy, z co dziesiątym studentem wchodzącym do biblioteki uniwersyteckiej, nie mógłby być pewny uzyskania w ten sposób próby *reprezentatywnej*, gdyż różne typy studentów odwiedzają bibliotekę z różną częstotliwością. W takiej próbie byliby nadreprezentowani studenci, którzy odwiedzają ją częściej niż inni.

Podobnie dzieje się w przypadku telefonicznych sondaży „opinii publicznej": słuchacze stacji radiowych czy też czytelnicy gazet mogą dzwonić pod określone numery telefonów i wyrażać swoje opinie, które zostaną automatycznie zarejestrowane. Nie można uważać ich wyników za reprezentatywne dla jakiejś większej populacji. W końcu nawet nie każdy członek tej populacji będzie wiedział o takim głosowaniu. Na tej samej zasadzie nieważne są sondaże przeprowadzane przez gazety i czasopisma, publikujące kupony do wypełnienia i odesłania przez czytelników. Nawet nie wszystkie osoby wiedzące o takim sondażu wyrażą swoją opinię, zwłaszcza jeśli musiałyby zapłacić za znaczek, kopertę czy impuls telefoniczny. Zastrzeżenia te dotyczą też sondaży przeprowadzanych przez Internet.

Jak na ironię, niezdolność takich sondaży do reprezentowania w takim samym stopniu wszystkich opinii została mimowolnie uznana przez Philipa Perinelliego (1986), szefa kadr w serwisie DIAL-IT 900 firmy AT&T Communications, oferującym usługi głosowania na telefon zainteresowanym instytucjom. Perinelli usiłował odpierać krytykę, mówiąc: „50-centowa opłata powoduje, że dzwonią tylko osoby zainteresowane, oraz daje gwarancję, że nikt nie «zapcha» urny". Ale badacze nie mogą określić ogólnej opinii publicznej, jeśli wezmą pod uwagę tylko „osoby zainteresowane". Taka analiza wyklucza tych, dla których kwota 50 centów ma znaczenie, a także tych, którzy nie wierzą w wartość takich sondaży. Oba te typy ludzi mogą mieć własne opinie, a nawet mogą głosować w dniu faktycznych wyborów. Co więcej, twierdzenie Perinelliego, że 50-centowa opłata zapobiega zapchaniu urny, oznacza, że tylko ci, których na to stać, będą brać udział w takich praktykach.

Możliwości nieumyślnego obciążenia próby są nieskończone i nie zawsze oczywiste. Istnieją na szczęście techniki, które pozwolą nam uniknąć obciążeń.

Reprezentatywność i prawdopodobieństwo wyboru

Mimo że pojęcie **reprezentatywności** nie ma ścisłego naukowego sensu, ma ono pewne potoczne znaczenie, które tutaj wykorzystamy. Dla naszych celów uznamy próbę za reprezentatywną dla populacji, z której została dobrana, jeśli zagregowane cechy tej próby ściśle odzwierciedlają te same zagregowane cechy populacji. Jeśli na przykład 50% populacji stanowią kobiety, to próba musi zawierać „w przybliżeniu" 50% kobiet, aby była reprezentatyw-

na. Dalej zajmiemy się szczegółowo owym „przybliżeniem".

Zauważmy, że próby nie muszą być reprezentatywne pod każdym względem; reprezentatywność ogranicza się do tych cech, które są związane z głównymi celami badań. Jednak nie zawsze z góry wiemy, które cechy są istotne.

Podstawowa zasada doboru losowego jest następująca: próba jest reprezentatywna dla populacji, z której jest dobierana, jeśli wszyscy członkowie tej populacji mają takie same szanse, że zostaną wybrani do tej próby (zobaczymy niebawem, że wielkość próby również wpływa na stopień reprezentatywności). Próby mające taką własność są często określane jako próby **EPSEM** (EPSEM to skrót od „metody równego prawdopodobieństwa wyboru", *equal probability of selection method*). Odmiany tej podstawowej zasady doboru losowego omówimy dalej.

Wychodząc od tej podstawowej zasady, musimy zauważyć, że próby – nawet starannie dobrane próby EPSEM – rzadko, jeśli w ogóle kiedykolwiek, reprezentują w sposób doskonały populacje, z których zostały pobrane. Mimo to dobór losowy ma dwie szczególne zalety.

Po pierwsze, próby losowe, choć nigdy nie są doskonale reprezentatywne, są z reguły bardziej reprezentatywne niż inne rodzaje prób, ponieważ są wolne od omówionych powyżej obciążeń. Praktycznie rzecz biorąc, bardziej prawdopodobne jest, że próba z jakiejś populacji będzie dla niej reprezentatywna, gdy jest to próba losowa, niż gdy jest to próba nielosowa.

Po drugie, co ważniejsze, rachunek prawdopodobieństwa pozwala nam oszacować reprezentatywność próby losowej. Można sobie wyobrazić sytuację, w której niezbyt kompetentny badacz mógłby zupełnie przypadkowymi metodami dobrać próbę reprezentującą niemal doskonale jakąś większą populację. Jednak szanse powodzenia takiego przedsięwzięcia są niewielkie. Nie bylibyśmy w stanie oszacować, na ile prawdopodobne jest, że taka próba jest reprezentatywna. Natomiast osoba dobierająca próbę losową może podać dokładne oszacowanie, na ile jej się to udało. Wkrótce przekonamy się, jak można dokonać takiego oszacowania.

Wspominałem, że dobór probabilistyczny zapewnia reprezentatywność próby dla populacji, którą chcemy badać. Jak się za moment przekonamy, dobór probabilistyczny polega na zastosowaniu procedury losowania. Aby rozwinąć tę myśl, potrzebujemy jednak bardziej precyzyjnie określić znaczenie dwóch ważnych terminów: *elementu* i *populacji*.

Element jest tą jednostką, o której zbiera się informacje i która dostarcza podstaw do analiz. W badaniach sondażowych elementy to z reguły ludzie lub pewne typy ludzi. Jednak także inne rodzaje jednostek mogą być elementami w badaniach społecznych: rodziny, kluby towarzyskie czy stowarzyszenia. W danych badaniach elementy i jednostki analizy często są te same, choć te pierwsze odnoszą się do doboru próby, a te drugie do analizy danych.

Używaliśmy dotychczas terminu *populacja* dla oznaczenia grupy lub zbioru, na temat którego chcieliśmy uogólniać nasze wnioski. W bardziej formalnym ujęciu **populacja** to określony teoretycznie zbiór elementów badania. Jeśli tematem badań mogłoby być nieostre pojęcie *Amerykanie*, to wyodrębnienie populacji zawierałoby definicję elementu *Amerykanie* (na przykład poprzez obywatelstwo, zamieszkanie) oraz czasu odnoszącego się do badania (Amerykanie kiedy?). Tłumaczenie abstrakcyjnego pojęcia „dorośli nowojorczycy" na roboczą populację wymagałoby wyszczególnienia wieku określającego *dorosłość* i granic Nowego Jorku. Szczegółowe określenie pojęcia *student college'u* zawierałoby rozważania, dotyczące studiujących w pełnym i niepełnym wymiarze godzin, tokiem kończącym się uzyskaniem stopnia lub nie, studentów poziomu podstawowego (*undergraduate*) i wyższego (*graduate*) itd.

Badana populacja (*study population*) to ten zbiór elementów, z którego próba jest faktycznie pobrana. Praktycznie wygląda to tak, że badacze rzadko mogą mieć pewność, że każdy element pasujący do ustalonych teoretycznych definicji faktycznie miał szansę bycia wybranym do próby. Nawet jeśli istnieją stworzone w celu pobrania próby listy elementów, są one zazwyczaj niekompletne. W spisach studentów mimo najlepszych chęci zawsze brakuje jakichś osób. Niektórzy abonenci telefoniczni

życzą sobie, by nie umieszczano w spisach ich nazwisk i numerów.

Badacze często decydują się ograniczyć swoją badaną populację ściślej, niż pokazałem to w poprzednich przykładach. Ogólnoamerykańskie firmy badawcze mogą ograniczać swoje próby o zasięgu krajowym do 48 sąsiadujących ze sobą stanów, pomijając z przyczyn praktycznych Alaskę i Hawaje. Badacz, chcący dobrać próbę profesorów psychologii, może chcieć ograniczyć badaną populację tylko do tych, którzy są zatrudnieni na wydziałach psychologii, pomijając zatrudnionych na innych wydziałach. Zawsze gdy dokonuje się takiego przedefiniowania badanej populacji, należy wyjaśnić odbiorcom wyniki dokonanej korekty.

Dobór losowy

Teraz, po przedstawieniu definicji, możemy określić ostateczny cel doboru próby: wybranie z populacji takiego zestawu elementów, aby opis tych elementów trafnie oddawał obraz całej populacji, z której zostały one dobrane. Losowy dobór próby zwiększa prawdopodobieństwo osiągnięcia tego celu i dostarcza metod oceny stopnia tego prawdopodobnego sukcesu.

Kluczem do tego procesu jest losowanie. W **losowaniu** każdy element może zostać wybrany z takim samym prawdopodobieństwem, niezależnie od innych zdarzeń w procesie selekcji. Najczęściej cytowanym przykładem jest rzucanie monetą: przy założeniu, że moneta jest doskonała (czyli nie ma obciążeń powodujących częstsze wypadanie orła lub reszki), „wybór" orła czy reszki jest niezależny od poprzednich wyborów orła lub reszki. Niezależnie od tego, ile razy pod rząd wypadnie reszka, prawdopodobieństwo tego, że następnym razem znowu wypadnie reszka wynosi dokładnie 50 do 50. Innym przykładem jest rzucanie doskonałą kostką.

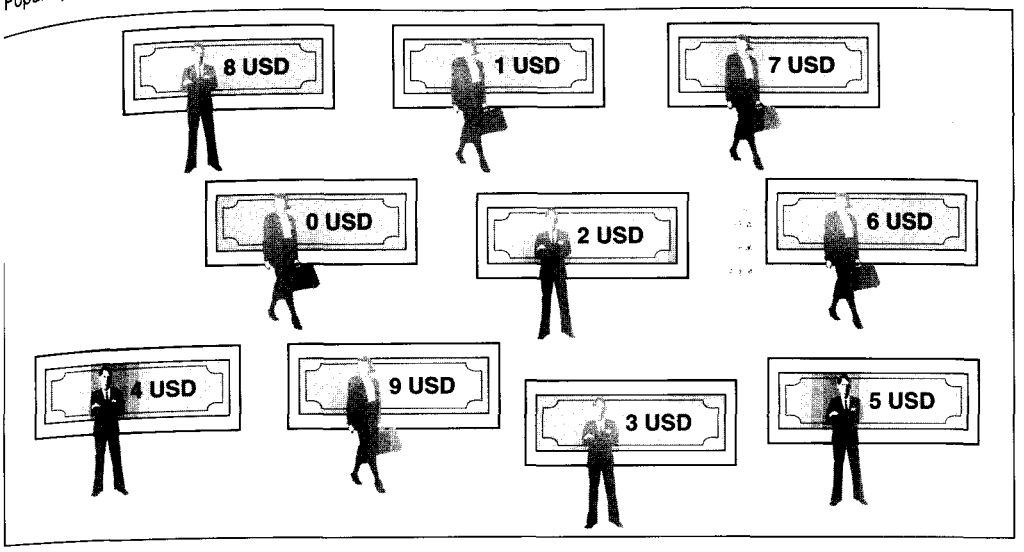
Jednak tego rodzaju obrazy doboru losowego, choć użyteczne, rzadko odnoszą się bezpośrednio do metod doboru próby w badaniach społecznych. Dla badacza społecznego bardziej typowe jest posługiwanie się tabelami liczb losowych lub programami komputerowymi, które zapewniają losowy dobór jednostek do próby. **Jednostka losowania** to ten element lub zbiór elementów, który jest brany do wylosowania na jakimś etapie procesu doboru próby. W rozdziale 9, dotyczącym badań sondażowych, zobaczymy, w jaki sposób używa się komputerów w celu losowego wybierania numerów telefonów do wywiadów telefonicznych.

Przyczyny, dla których stosuje się metody doboru losowego, są dwojakiego rodzaju. Po pierwsze, procedura ta służy kontroli świadomego lub nieświadomego obciążenia próby następującego ze strony badacza. Badacz, dobierający przypadki na zasadzie intuicji, mógłby z powodzeniem dobierać je tak, by uzasadniały jego oczekiwania czy hipotezy związane z badaniami. Dobór losowy zapobiega temu niebezpieczeństwu. Co ważniejsze, dobór losowy daje dostęp do zasobów rachunku prawdopodobieństwa, co stanowi podstawę do oszacowania cech populacji, a także do oszacowania dokładności prób. Przeanalizujmy teraz nieco dokładniej rachunek prawdopodobieństwa.

Rachunek prawdopodobieństwa, rozkład z prób i szacowanie błędu z próby

Rachunek prawdopodobieństwa jest dziedziną matematyki, która dostarcza narzędzi potrzebnych badaczowi do opracowania technik doboru próby zapewniających reprezentatywność i do analizy statystycznej wyników uzyskanych z tak dobranej próby. Ujmując rzecz bardziej formalnie, rachunek prawdopodobieństwa dostarcza podstaw do szacowania parametrów w populacji. **Parametr** to syntetyczny opis jakiejś zmiennej w populacji. Średni dochód wszystkich rodzin w mieście oraz rozkład wieku ludności tego miasta to przykłady parametrów. Kiedy badacze uogólniają wnioski z próby, używają obserwacji uzyskanych na podstawie próby do oszacowania parametrów w populacji. Rachunek prawdopodobieństwa umożliwia im zarówno samo szacowanie, jak i osąd, z jakim prawdopodobieństwem szacunki te będą dokładnie odwzorowywać faktyczne parametry w populacji. I tak na przykład rachunek

Rysunek 7.3
Populacja dziesięciu osób mających od 0 do 9 dolarów



prawdopodobieństwa pozwala badaczom wnioskować na podstawie próby liczącej 2000 osób, jak będzie prawdopodobnie głosować populacja składająca się ze stu milionów wyborców – a także pozwala określić dokładnie prawdopodobny margines błędu tych szacunków.

Rachunek prawdopodobieństwa dokonuje tych pozornie magicznych sztuczek za pomocą pojęcia rozkładu z prób. Pojedyncza próba wylosowana z populacji może dać estymator parametru w populacji. Inne próby mogą dać takie same lub nieznacznie różniące się estymatory. Rachunek prawdopodobieństwa mówi o rozkładach estymatorów uzyskanych z wielkiej liczby takich prób. Aby zobaczyć, jak to działa, przyjrzymy się dwóm przykładom rozkładów z prób, zaczynając od prostego przykładu, w którym nasza populacja będzie składać się jedynie z dziesięciu przypadków.

Rozkład z prób dla dziesięciu przypadków

Wyobraźmy sobie grupę dziesięciu osób, z których każda ma pewną kwotę pieniędzy w kieszeni. Dla uproszczenia załóżmy, że jedna osoba w ogóle nie ma pieniędzy, druga ma jednego dolara, trzecia dwa dolary itd., aż do osoby, która ma dziewięć dolarów. Rysunek 7.3 pokazuje taką dziesięcioosobową populację*.

Naszym zadaniem jest określenie średniej kwoty pieniędzy dla tej populacji: konkretnie, średniej liczby dolarów. Jeśli po prostu zsumujesz pieniądze przedstawione na rysunku 7.3, okaże się, że suma wynosi 45 USD, więc średnia wynosi 4,50 USD. Celem dalszego ciągu tego ćwiczenia jest oszacowanie tej średniej bez rzeczywistej znajomości wszystkich dziesięciu jednostek. Dokonamy tego przez dobranie prób losowych z tej populacji i zastosowanie średnich, obliczonych dla tych prób, do oszacowania wartości średniej dla całej naszej populacji.

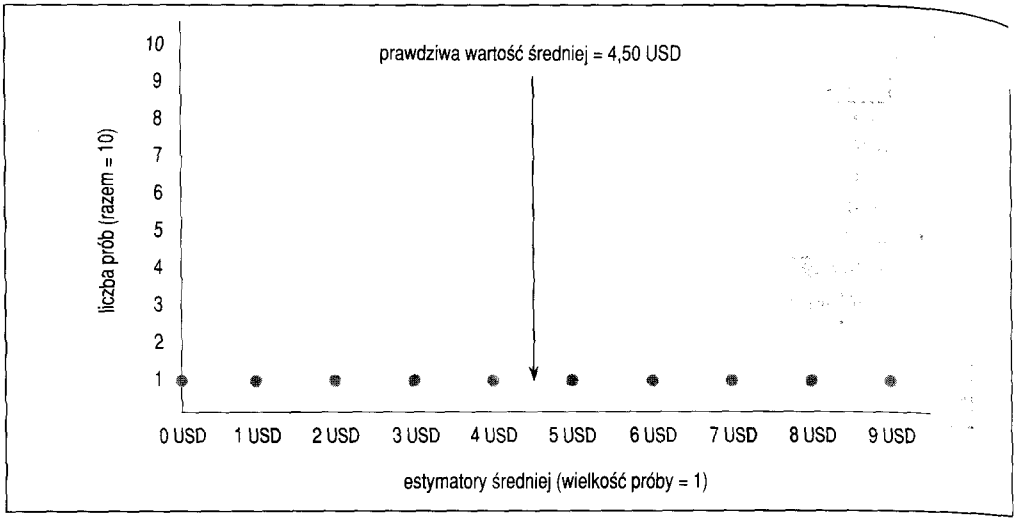
Na początek przypuśćmy, że mamy dobrać – losowo – próbę złożoną z tylko jednej osoby z tych dziesięciu. Tak więc nasze dziesięć możliwych prób składa się z dziesięciu przypadków pokazanych na rysunku 7.3.

Dziesięć kropek na wykresie na rysunku 7.4 przedstawia owe dziesięć prób. Ponieważ losowaliśmy próby tylko jednoosobowe, przedstawiają one również „średnie", które otrzymali-

* Chcę podziękować Hananowi Selvinowi za podpowiedzenie tej metody wprowadzenia do probabilistycznego doboru próby.

Rysunek 7.4

Rozkład z prób dla prób jednoosobowych

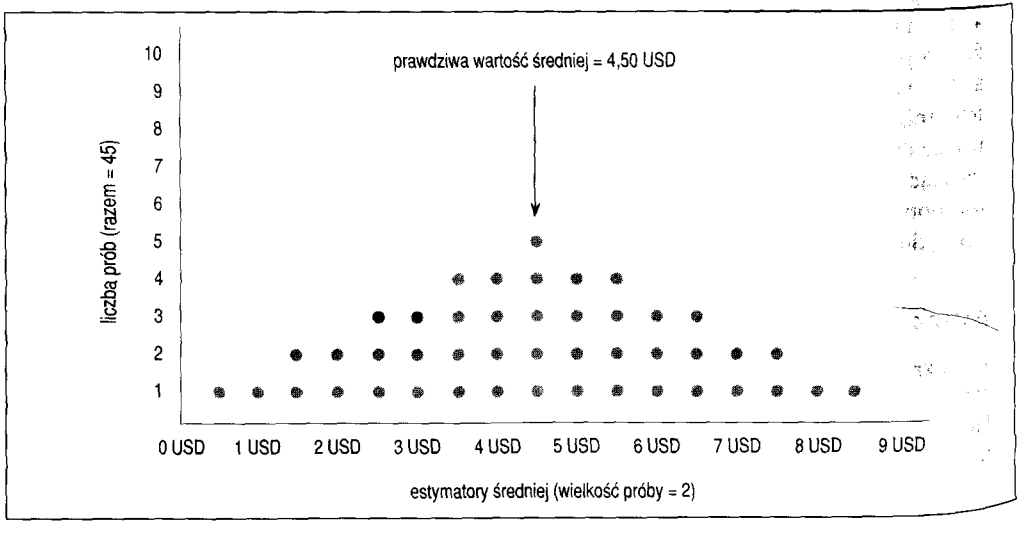


byśmy jako estymatory dla populacji. Rozkład tych kropek na wykresie nazywa się *rozkładem z prób*. Oczywiście w zależności od tego, którą osobę wylosujemy, oszacujemy średnią dla całej grupy jako wartość z zakresu od 0 USD do 9 USD. Jasne zatem, że dobieranie próby tylko jednego przypadku nie jest dobrym pomysłem, gdyż z dużym prawdopodobieństwem rozminęlibyśmy się znacznie z prawdziwą średnią wynoszącą 4,50 USD.

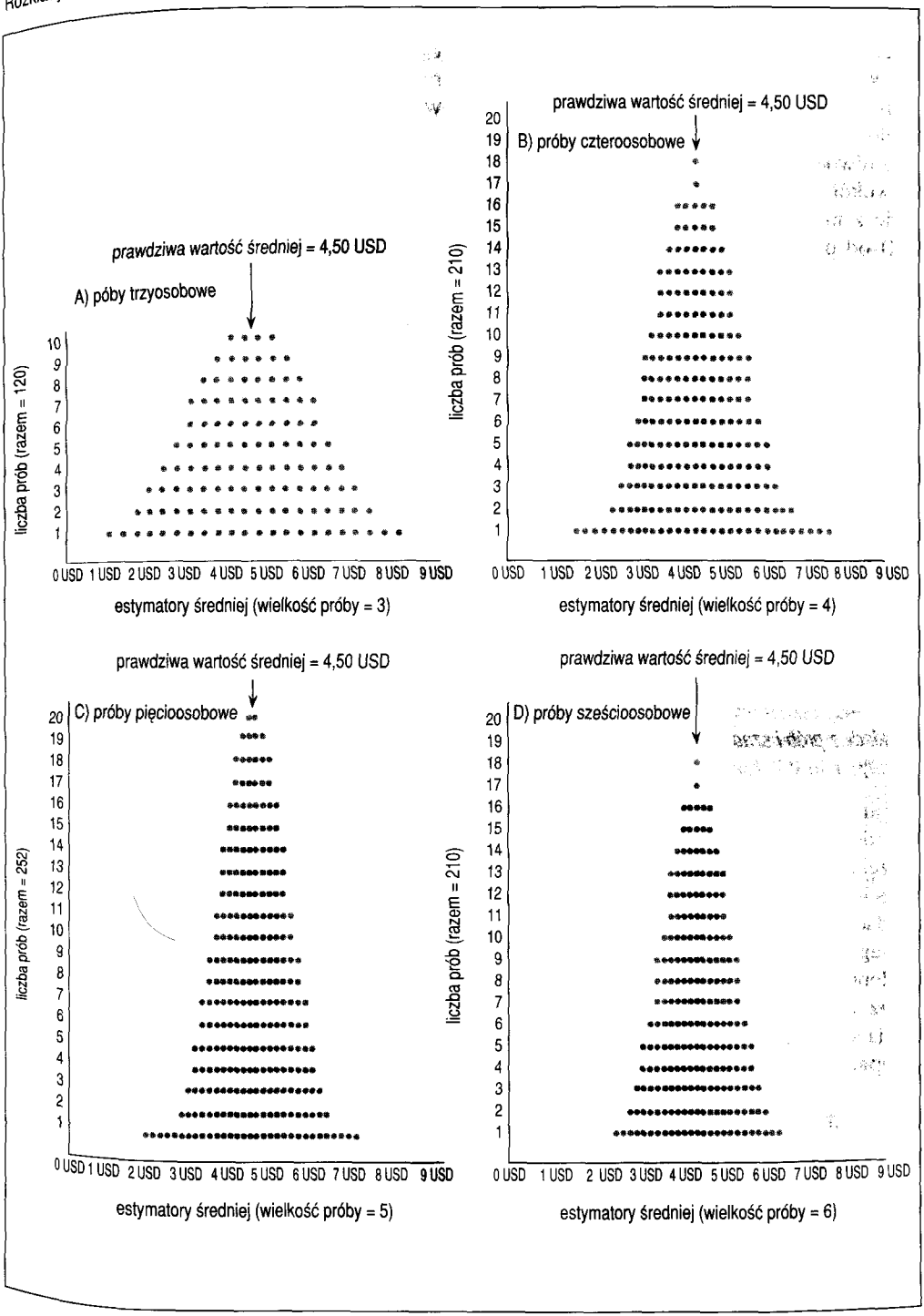
A teraz załóżmy, że losujemy próby dwuosobowe. Jak się można przekonać na podstawie rysunku 7.5, zwiększanie wielkości próby poprawia nasze szacunki. Mamy teraz 45 możliwych prób: [0 USD 1 USD], [0 USD 2 USD], ... [7 USD 8 USD], [8 USD 9 USD]. Co więcej, nie-

Rysunek 7.5

Rozkład z prób dla prób dwuosobowych



Rysunek 7.6
Rozkłady z prób dla prób trzy-, cztero-, pięcio- i sześcioosobowych



które z tych prób mają takie same średnie. Dla przykładu [0 USD 6 USD], [1 USD 5 USD] i [2 USD 4 USD] wszystkie mają średnią 3 USD. Te trzy próby widać na rysunku 7.5 w postaci trzech kropek ponad średnią 3 USD.

W dodatku uzyskane przez nas 45 średnich z prób nie jest tak równo rozłożone, jak miało to miejsce, kiedy nasza wielkość próby była równa tylko 1. Są one raczej jakby skupione wokół prawdziwej wartości 4,50 USD. Tylko dwie z możliwych prób odchylają się aż o 4 USD od prawdziwej wartości ([0 USD 1 USD] i [8 USD 9 USD]), podczas gdy pięć prób daje prawdziwe oszacowanie 4,50 USD; osiem kolejnych prób chybia celu o jedyne 50 centów (na plus lub minus).

Przypuśćmy teraz, że dobieramy jeszcze większe próby. Jak sądzisz, jak wpłynie to na nasze szacunki średniej? Rysunek 7.6 ukazuje rozkłady z prób dla prób o liczebności 3, 4, 5 i 6.

Postęp rozkładów z prób jest ewidentny. Każdy przyrost wielkości próby poprawia rozkład szacunków średniej. Krańcowym przypadkiem w tym przykładzie jest oczywiście dobranie próby 10 osób. Istniałaby tylko jedna możliwa próba (wszyscy) i dałaby nam ona prawdziwą wartość średniej 4,50 USD.

Rozkłady z prób i szacowanie błędu z próby

Przejdźmy teraz do bardziej realistycznej sytuacji dotyczącej znacznie większej populacji i przekonajmy się, jakie zastosowanie ma tutaj pojęcie rozkładu z prób. Załóżmy, że chcemy zbadać populację studentów Uniwersytetu Stanowego (US), aby określić, jaki jest procent studentów, którzy akceptują zaproponowany przez władze uczelni projekt kodeksu zachowania studentów lub nie akceptują go. W tym przypadku badaną populacją będzie zbiór, powiedzmy, 20 000 studentów zarejestrowanych w spisie studentów. Elementami będą pojedynczy studenci US. Dobierzemy próbę losową, na przykład 100 studentów, w celu oszacowania całego gremium studentów. Rozważaną zmienną będą *postawy wobec kodeksu*, zmienna dwuwartościowa: akceptacja lub brak akceptacji. (Logika probabilistycznego doboru próby ma zastosowanie w badaniu innych typów zmiennych, takich jak średni dochód, ale obliczenia są nieco bardziej skomplikowane. Tak więc w niniejszym wprowadzeniu będziemy się skupiać na zmiennych dwuwartościowych).

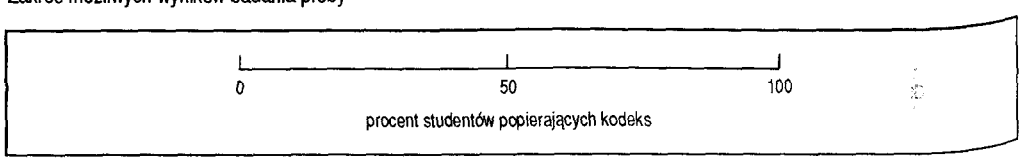
Oś *pozioma rysunku* 7.7 przedstawia wszystkie możliwe wartości owego parametru w populacji, który chcemy oszacować – od 0 do 100% akceptacji. Środkowy punkt osi – 50% – przedstawia akceptację dla kodeksu ze strony połowy studentów i brak akceptacji ze strony drugiej połowy.

By dobrać próbę, przypisujemy każdemu studentowi w spisie studentów numer i wybieramy sto liczb z tablicy liczb losowych. Następnie przeprowadzamy wywiady z owymi stoma studentami, których numery zostały wybrane, i pytamy ich o ich postawy wobec kodeksu studenckiego: czy akceptują go, czy też nie. Przypuśćmy, że w wyniku tego mamy 48 studentów, którzy akceptują kodeks i 52, którzy go nie akceptują. Ten syntetyczny opis zmiennej w próbie jest nazywany **statystyką**. Przedstawiamy tę statystykę, stawiając kropkę na osi *x* w punkcie reprezentującym 48%.

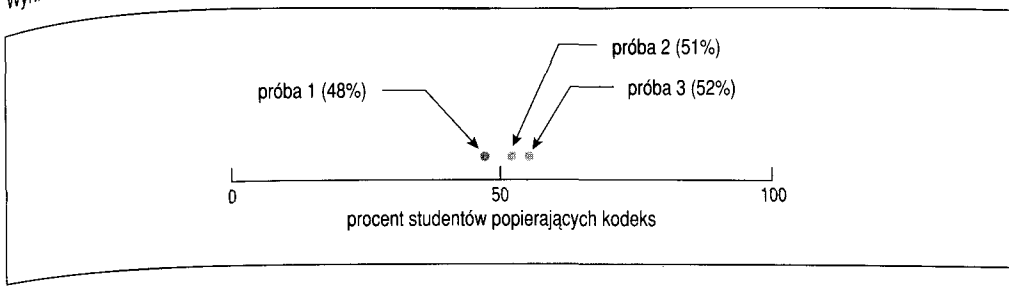
Przypuśćmy teraz, że dobieramy inną próbę 100 studentów dokładnie w ten sam sposób i mierzymy ich akceptację lub brak akceptacji dla kodeksu studenckiego. Może się zdarzyć, że w drugiej próbie zaakceptuje kodeks 51 studentów. Stawiamy następną kropkę w odpowiednim miejscu na osi *x*. Powtarzając po raz kolejny ten proces, możemy przekonać się, że

Rysunek 7.7

Zakres możliwych wyników badania próby



Rysunek 7.8
Wyniki trzech hipotetycznych badań



w trzeciej próbie kodeks ma poparcie 52 studentów.

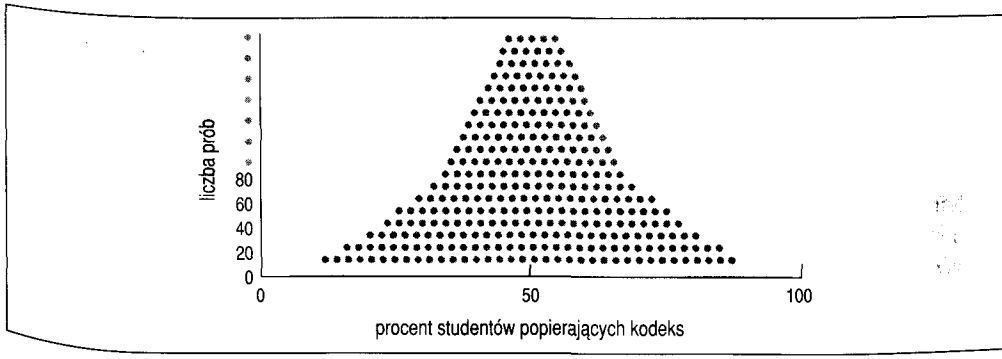
Rysunek 7.8 przedstawia te trzy różne statystyki z prób, reprezentujące odsetki studentów, którzy akceptowali kodeks studencki w każdej z trzech prób losowych. Podstawową regułą losowego doboru próby jest to, że próby pobrane w taki sposób z danej populacji dają oszacowanie parametru, który należy do całej populacji. Zatem każda z tych prób losowych dostarcza nam estymatora odsetka studentów, którzy popierają kodeks wśród wszystkich studentów. Jednak niestety dobraliśmy trzy próby i teraz mamy trzy odrębne estymatory.

Aby poprawić nasze położenie, wyciągnijmy jeszcze więcej prób po 100 studentów każda, w każdej z nich pytając o akceptację lub brak akceptacji dla kodeksu studenckiego, i nanieśmy nowe statystyki z prób na nasz sumaryczny wykres. Wyciągając wiele takich prób stwierdzamy, że niektóre z nowych prób dają takie same estymatory, podobnie jak w przykładzie dotyczącym dziesięciu przypadków. Rysunek 7.9 pokazuje rozkład z wielu takich prób. Ten rozkład w kształcie lit U często określany jest jako *krzywa normaln*

Zauważ, że zwiększając liczbę dobran... i przebadanych prób, zwiększyliśmy również zakres estymatorów uzyskanych w wyniku operacji doboru próby. W pewnym sensie zwiększyliśmy nasz dylemat dotyczący odgadnięcia parametru w populacji. Jednak rachunek prawdopodobieństwa dostarcza nam pewnych ważnych reguł dotyczących rozkładu z prób, przedstawionego na rysunku 7.9.

Po pierwsze, jeśli z jakiejś populacji pobierze się wiele niezależnych prób, statystyki z tych prób będą rozproszone wokół rzeczywistego parametru populacji w pewien znany sposób. A zatem, chociaż rysunek 7.9 ukazuje szeroki zakres estymatorów, w pobliżu 50% jest ich więcej niż gdziekolwiek indziej na tym wykresie. Czyli rachunek prawdopodobieństwa mówi nam, że prawdziwa wartość jest w pobliżu 50%.

Rysunek 7.9
Rozkład z prób



Po drugie, rachunek prawdopodobieństwa dostarcza nam wzoru dla oszacowania, na ile blisko statystyki z prób są skupione wokół prawdziwej wartości. Ujmując rzecz inaczej: rachunek prawdopodobieństwa pozwala nam oszacować **błąd z próby (losowy)** (*sampling error*) – poziom błędu oczekiwany dla danego rodzaju próby. Wzór zawiera trzy czynniki: parametr, wielkość próby i błąd standardowy (miernik błędu z próby):

$$s = \sqrt{\frac{P \times Q}{n}}$$

Symbole P i Q w tym wzorze równają się parametrom populacji w rozkładzie dwuwartościowym: jeśli 60% wszystkich studentów akceptuje kodeks, a 40% nie akceptuje go, to P i Q wynoszą odpowiednio 60% i 40% lub 0,6 i 0,4. Zauważ, że $Q = 1 - P$, a $P = 1 - Q$. Symbol n jest równy liczbie przypadków w każdej próbie, a s to błąd standardowy.

Załóżmy, że rzeczywisty parametr w populacji to 50% studentów akceptujących kodeks i 50% nieakceptujących go. Przypomnijmy, że losowaliśmy próby po 100 przypadków każda. Gdy podstawimy te liczby do wzoru, okaże się, że błąd standardowy równa się 0,05, czyli 5%.

Błąd standardowy jest w rachunku prawdopodobieństwa wartościową informacją, ponieważ określa, na ile blisko estymatory z prób będą rozrzucone wokół parametru populacji. (Jeśli nie jest ci obce pojęcie *odchylenia standardowego* występujące w statystyce, możesz zauważyć, że w tym przypadku błąd standardowy jest odchyleniem standardowym rozkładu z prób). Rachunek prawdopodobieństwa wykazuje w szczególności, że pewne frakcje estymatorów z prób będą mieścić się w określonych odległościach – stanowiących wielokrotności jednego błędu standardowego – od parametru populacji. Około 34% (0,3413) estymatorów z prób mieści się w odległości jednego błędu standardowego od parametru populacji w górę, a kolejne 34% mieści się w odległości jednego błędu standardowego od tego parametru w dół. W naszym przykładzie jeden błąd standardowy wynosi 5%, wiemy więc, że z 34% naszych prób otrzymamy estymatory poparcia studenckiego leżące między 50% (parametr) a 55% (jeden błąd standardowy powyżej); z kolejnych 34% prób otrzymamy estymatory leżące między 50% a 45% (jeden błąd standardowy poniżej parametru). Zatem w sumie wiemy, że z około dwóch trzecich (68%) prób otrzymamy estymatory odchylające się nie więcej niż o ±5% od parametru.

Co więcej, rachunek prawdopodobieństwa głosi, że około 95% prób mieści się w granicach plus minus dwóch błędów standardowych od prawdziwej wartości, a 99,9% prób mieści się w granicach plus minus trzech błędów standardowych. Zatem w odniesieniu do naszego przykładu wiemy, że tylko z jednej próby na tysiąc otrzymalibyśmy wartość estymatora mniejszą niż 35% lub większą niż 65%.

Ów odsetek prób mieszczących się w granicach jednego, dwóch lub trzech błędów standardowych od parametru jest stały dla każdej procedury losowania prób w opisany powyżej sposób, pod warunkiem że pobiera się duże ilości prób. Jednak w każdym takim przypadku wielkość błędu standardowego jest funkcją parametru populacji i wielkości próby. Jeśli wrócimy na chwilę do wzoru, zauważymy, że błąd standardowy będzie wzrastał wraz ze wzrostem wartości P razy Q. Zauważymy też, że ta wartość osiąga swoje maksimum w sytuacji równego podziału populacji. Jeśli $P = 0,5$, $PQ = 0,25$; jeśli $P = 0,6$, $PQ = 0,24$; jeśli $P = 0,8$, $PQ = 0,16$; jeśli $P = 0,9$, $PQ = 0,0099$. Na tej samej zasadzie, jeśli P wyniesie albo 0,0, albo 1,0 (albo 0%, albo 100% poparcia dla kodeksu studenckiego), błąd standardowy wyniesie 0. Oznacza to, że jeśli każdy w populacji ma takie samo zdanie (brak zmienności), wówczas z każdej próby otrzymamy dokładnie ten estymator.

Błąd standardowy jest też funkcją wielkości próby – funkcją odwrotną. Gdy wielkość próby rośnie, błąd standardowy maleje. Innymi słowy, gdy wielkość próby rośnie, próby skupiają się bliżej prawdziwej wartości. Ze wzoru można odczytać jeszcze jedną ogólną wskazówkę: z powodu zawartego w nim pierwiastka kwadratowego błąd standardowy zmniejsza się o połowę, gdy wielkość próby rośnie czterokrotnie. W naszym przykładzie próby o wielkości 100 dają błąd standardowy 5%; żeby zredukować ten błąd do 2,5%, musimy zwiększyć próbę do 400.

Wszystkich tych informacji dostarcza nam ustalony rachunek prawdopodobieństwa w odniesieniu do dużych ilości prób losowych. (Jeśli uzyskałeś zaliczenie ze statystyki, możesz to pamiętać jako „twierdzenie tendencji centralnej"). Jeśli znany jest parametr w populacji i dobiera się wiele prób losowych, można przewidzieć, ile z tych prób będzie się mieścić w granicach określonych przedziałów od parametru.

Należy zdawać sobie sprawę z tego, że te rozważania ilustrują jedynie logikę probabilistycznego doboru próby, nie opisują one sposobu, w jaki faktycznie prowadzi się badania. Zazwyczaj nie znamy wartości parametru: właśnie po to przeprowadza się sondaż na próbie, by oszacować tę wartość. Co więcej, w rzeczywistości nie pobiera się dużych ilości prób: dobiera się tylko jedną próbę. Mimo to powyższe rozważania dotyczące rachunku prawdopodobieństwa dają podstawę do konkluzji na temat typowych sytuacji w badaniach społecznych. Wiedza o tym, jakby to było, gdyby dobrać tysiące prób, pozwala nam na poczynienie założeń co do tej jednej próby, którą w końcu dobieramy i badamy.

Poziomy i przedziały ufności

Skoro rachunek prawdopodobieństwa określa, że 68% wyimaginowanej wielkiej ilości prób dałoby estymatory leżące w granicach jednego błędu standardowego od parametru, odwróćmy tę logikę i wyciągnijmy wniosek, że dowolna próba losowa ma 68-procentową szansę znalezienia się w granicach tego przedziału. Obserwacja ta prowadzi nas do dwóch kluczowych składników szacowania błędu z próby: **poziomu ufności i przedziału ufności**. Dokładność naszych statystyk z prób wyrażamy w kategoriach *poziomu ufności* co do tego, że statystyki te mieszczą się w określonym *przedziale* od parametru. Przykładowo, możemy powiedzieć, że mamy 95-procentową ufność, że nasza statystyka z próby (na przykład 50-procentowe poparcie dla nowego kodeksu studenckiego) mieści się w przedziale ±5% od parametru w populacji. Gdy poszerza się przedział ufności dla danej statystyki, nasza ufność się zwiększa. Możemy na przykład powiedzieć, że mamy 99,9-procentową ufność, że nasza statystyka mieści się w granicach trzech błędów standardowych od prawdziwej wartości parametru.

Mimo że możemy mieć ufność (na danym poziomie), że mieścimy się w określonym przedziale od parametru, zauważyliśmy już, że rzadko znamy rzeczywistą wartość tego parametru. A jednak podstawienie jakiejś wartości tego parametru jest wymagane we wzorze na błąd z próby. W celu rozwiązania tego problemu podstawimy we wzorze nasz estymator z próby w miejsce parametru – a więc, nie znając prawdziwej wartości, zastąpimy ją najlepszą możliwą.

W rezultacie tego wnioskowania i obliczeń możemy oszacować parametr w populacji, a także oczekiwany poziom błędu na podstawie jednej próby dobranej z populacji. Wychodząc od pytania: „Jaki procent wszystkich studentów akceptuje kodeks studencki?", moglibyśmy dobrać próbę losową 100 studentów i przeprowadzić z nimi wywiady. Następnie moglibyśmy obwieścić, że szacunkowo 50% wszystkich studentów akceptuje kodeks i że mamy 95-procentową pewność co do tego, że odsetek akceptujących leży między 40% a 60% (plus minus dwa błędy standardowe). Ów zakres od 40% do 60% to przedział ufności. (Przy poziomie ufności 68% nasz przedział ufności wyniósłby 45–55%).

Logika leżąca u podstaw określania poziomów i przedziałów ufności dostarcza też podstaw do określenia właściwej wielkości próby do badań. Gdy już zdecydujesz się, jaki poziom błędu z próby możesz tolerować, będziesz w stanie obliczyć liczbę przypadków potrzebnych w twojej próbie. Czyli dla przykładu, jeśli chcesz mieć 95-procentową pewność, że wyniki twoich badań są dokładne w granicach ±5% od parametrów w populacji, powinieneś dobrać próbę o liczebności co najmniej 400. (Załącznik H stanowi poręczny przewodnik w tym względzie).

Tak wyglądają podstawy logiki probabilistycznego doboru próby. Dobór losowy pozwala badaczowi na łączenie wyników z próby z całym dorobkiem rachunku prawdopodobieństwa, dzięki czemu może on oszacować trafność tych wyników. Wszelkie twierdzenia dotyczące trafności wyników, związanej z doborem próby, muszą informować zarówno o poziomie ufności, jak i o przedziale ufności.

Badacz musi określić, że ma *x*-procentową pewność, że parametr populacji leży między dwiema konkretnymi wartościami.

Oto jak George Gallup (1984, s. 7) opisał błąd z próby w raporcie prasowym z sondażu Gallupa dotyczącego postaw dzieci i rodziców:

> „Wyniki dotyczące dorosłych oparte są na wywiadach osobistych z 1520 dorosłymi osobami, w wieku 18 i więcej lat, przeprowadzonych w ponad 300 dobranych za pomocą metod naukowych miejscowościach w całym kraju, w okresie 26–29 października. W odniesieniu do wyników uzyskanych na podstawie próby tej wielkości można powiedzieć z 95-procentową pewnością, że błąd wynikający z próby i innych efektów losowych mógłby wynieść trzy punkty procentowe w każdym kierunku".

A oto, co miał do powiedzenia „New York Times" o swoim sondażu na temat poglądów religijnych („Jak przeprowadzono sondaż", 1995, s. 15):

> „Zgodnie z teorią, w 19 na 20 przypadków wyniki oparte na takich próbach nie różnią się o więcej niż trzy punkty procentowe w każdą stronę od rezultatów, które uzyskanoby, gdyby przebadano wszystkich dorosłych Amerykanów".

Gdy czytasz w gazetach stwierdzenia, takie jak przytoczone powyżej, pamiętaj, że są one niekiedy podawane, choć nie ma do nich podstaw. Ze szczególną ostrożnością traktuj te wyniki sondaży, przy których nie podano poziomów i przedziałów ufności – bez tych szczegółowych informacji „wyniki" mają wątpliwą wartość.

Należy uczynić dwa zastrzeżenia, zanim zakończymy omawianie podstawowej logiki rachunku prawdopodobieństwa. Po pierwsze, omówione tu sondażowe zastosowania rachunku prawdopodobieństwa nie zawsze są w pełni uzasadnione technicznie. Teoria rozkładu z prób posługuje się założeniami, które niemal nigdy nie są spełnione w warunkach przeprowadzania sondażu. Na przykład określenie dokładnej proporcji prób zawartych w granicach określonych przedziałów błędu standardowego opiera się na matematycznym założeniu nieskończenie wielkiej populacji, nieskończonej ilości prób oraz losowania ze zwracaniem – czyli założeniu, że każdy wylosowany element zostaje „wrzucony z powrotem do urny" i może być wylosowany ponownie. Po drugie, w naszych rozważaniach zbytniemu uproszczeniu poddaliśmy przeskok dedukcyjny od rozkładu z wielu prób do prawdopodobnych cech jednej próby.

Czynię te zastrzeżenia po to, byś widział we właściwych proporcjach zastosowanie rachunku prawdopodobieństwa do doboru próby losowej. Badacze społeczni często, jak się wydaje, przeceniają dokładność szacunków dokonanych z wykorzystaniem rachunku prawdopodobieństwa. Jak jeszcze wspomnę w innym miejscu tego rozdziału i w innych miejscach tej książki, różne odmiany technik losowania, a także czynniki niezwiązane z losowaniem mogą jeszcze bardziej ograniczać prawomocność takich szacunków. Na przykład osoby wylosowane, które nie mogą lub nie chcą uczestniczyć w badaniach, wpływają na dalsze obniżenie reprezentatywności próby.

Tak czy inaczej, obliczenia omówione w tej części mogą być dla ciebie niezmiernie użyteczne dla zrozumienia i oceny twoich danych. Mimo że oszacowania otrzymane w wyniku tych obliczeń nie są tak dokładne, jak mogliby zakładać niektórzy badacze, mogą być one całkiem uzasadnione dla celów praktycznych. Są one niewątpliwie trafniejsze niż mniej rygorystyczne oszacowania uzyskane na podstawie mniej rygorystycznych metod doboru próby. Co najważniejsze, orientacja w logice, leżącej u podstaw tych obliczeń, pomoże ci świadomie obchodzić się nie tylko z własnymi danymi, ale także z raportami innych badaczy.

Populacje i operaty losowania

Poprzedni podrozdział stanowił wprowadzenie do teoretycznego modelu doboru próby w badaniach społecznych. Chociaż jako studenci, odbiorcy wyników badań czy badacze musimy rozumieć tę teorię, nie mniej ważne jest, byśmy brali również pod uwagę nie zawsze idealne warunki spotykane w terenie. W niniejszym podrozdziale zajmiemy się rozważaniem jednego z aspektów warunków terenowych, który

wymaga kompromisu z wyidealizowanymi warunkami i założeniami teoretycznymi: mianowicie zbieżności bądź niewspółmierności populacji i operatów losowania.

Ujmując rzecz najprościej, **operat losowania** jest to lista lub quasi-lista elementów, z której losuje się próbę. Jeśli próba studentów jest losowana z wykazu studentów, wykaz ten jest operatem losowania. Jeśli pierwotną jednostką losowania złożonej próby z jakiejś populacji jest kwartał spisowy, wykaz takich kwartałów stanowi operat losowania – w formie drukowanej broszurki, zapisu na taśmie magnetycznej lub innego skomputeryzowanego zapisu. Oto kilka informacji na temat operatów losowania, jakie ukazały się w czasopismach badawczych:

> „Dane do tego badania uzyskano z próby losowej *rodziców dzieci klas trzecich szkół publicznych i parafialnych w hrabstwie Yakima w stanie Waszyngton*".
>
> (Petersen, Maynard 1981, s. 92)

> „Przy pierwszym badaniu próba składała się ze 160 nazwisk, wylosowanych z *książki telefonicznej Lubbock w Teksasie*".
>
> (Tan 1980, s. 242)

> „Dane, o których piszemy w tym artykule (...) zostały zgromadzone na podstawie próby losowej *dorosłych w wieku 18 lat i więcej, zamieszkałych w gospodarstwach domowych na terenie 48 przylegających do siebie stanów USA*. Jesienią 1975 roku Centrum Badań Sondażowych Uniwersytetu Michigan przeprowadziło wywiady osobiste z 1914 respondentami".
>
> (Jackman, Senter 1980, s. 345)

W każdym z tych przykładów wyróżniono kursywą faktycznie zastosowane operaty losowania.

Właściwie dobrana próba dostarcza informacje nadające się do opisania populacji elementów, tworzących operat losowania – nic ponadto. Podkreślam to wobec aż nazbyt częstej tendencji występującej wśród badaczy, by dobierać próbę z danego operatu, a następnie formułować twierdzenia dotyczące populacji podobnej do populacji określonej operatem losowania, ale nie *identycznej* z nią.

Dla ilustracji rzućmy okiem na raport, który zajmuje się rozważaniami na temat najczęściej przepisywanych przez lekarzy w USA leków:

> „Niełatwo jest zdobyć informacje na temat sprzedaży leków na recepty. Jednak Rinaldo V. DeNuzzo, profesor farmacji na Albany College of Pharmacy, na Uniwersytecie Union w Albany w stanie Nowy Jork, *śledził sprzedaż leków na recepty* przez 25 lat, robiąc badania ankietowe w pobliskich aptekach. Wyniki zostały opublikowane przez niego w czasopiśmie branżowym *MM&M*.
>
> Najnowszy sondaż DeNuzzo, dotyczący 1980 roku, opiera się na informacjach z 66 aptek w 48 miejscowościach w stanach Nowy Jork i New Jersey. Ponieważ te części kraju nie są w jakiś sposób specyficzne, jego wnioski można uważać za reprezentatywne dla tego, co dzieje się w całym kraju".
>
> (Moskowitz 1981, s. 33)

Główną rzeczą, która powinna cię uderzyć w powyższym fragmencie, jest owa zdawkowa uwaga o tym, czy jest coś specyficznego w stanach Nowy Jork i New Jersey. Otóż jest. Styl życia w tych dwóch stanach raczej nie jest typowy dla pozostałych 48. Nie mamy prawa zakładać, że mieszkańców tych dużych, zurbanizowanych stanów Wschodniego Wybrzeża cechują akurat takie same wzorce korzystania z leków, jak mieszkańców stanów Missisipi, Nebraska czy nawet Vermont.

Czy przynajmniej cytowany sondaż reprezentuje wzorce sprzedaży leków na recepty w stanach Nowy Jork i New Jersey? Aby to stwierdzić, musielibyśmy wiedzieć coś o sposobie, w jaki dobrano owe 48 miejscowości i 66 aptek. Powinniśmy tu zachować czujność, w związku ze wzmianką o „badaniach ankietowych w pobliskich aptekach". Jak zobaczymy, istnieje kilka metod doboru próby, które gwarantują reprezentatywność. Jeśli zostały zastosowane inne metody, nie powinniśmy uogólniać wyników tych badań.

Operat losowania musi być zatem zbieżny z populacją, którą chcemy badać. W najprostszym modelu doboru próby operatem losowania jest lista elementów, tworzących badaną populację. Jednak w praktyce to raczej dostępne operaty losowania definiują badaną popu-

lację niż na odwrót. Czyli często przystępujemy do badań z określonym wyobrażeniem populacji, którą chcemy się zająć, a następnie szukamy możliwych operatów losowania. Po zanalizowaniu i ocenie dostępnych dla naszych celów operatów losowania podejmujemy decyzję, który z tych operatów przedstawia najwłaściwszą dla naszych potrzeb populację.

Najprostsze z punktu widzenia doboru próby są często badania różnych organizacji, gdyż organizacje na ogół mają listy członków. W takich przypadkach listy członków są znakomitymi operatami losowania. Jeśli losuje się próbę z listy członków, dane zebrane na podstawie tej próby mogą być uważane za reprezentatywne dla wszystkich członków – jeśli wszyscy członkowie są uwzględnieni na liście.

Populacje, z których można dobrać próbę na podstawie dobrej listy, to: uczniowie i nauczyciele szkół podstawowych i średnich, studenci i pracownicy uniwersytetów; członkowie wspólnot kościelnych; załogi fabryk; członkowie bractw studenckich; członkowie klubów społecznych czy politycznych oraz członkowie stowarzyszeń zawodowych.

Powyższe uwagi odnoszą się zwłaszcza do organizacji lokalnych. Organizacje ogólnostanowe czy ogólnokrajowe często nie dysponują jedną listą członków. Nie istnieje na przykład jedna lista członków Kościoła Episkopalnego. A jednak nieco bardziej złożony schemat doboru próby pozwala na wykorzystanie lokalnych list członków tego kościoła poprzez dobranie próby najpierw poszczególnych placówek kościelnych, a potem podprób z list członków tych placówek (więcej na ten temat dalej).

Inne rodzaje list jednostek mogą być odpowiednie dla potrzeb badawczych danego projektu. Na przykład instytucje rządowe mają rejestry wyborców, z których można skorzystać[4], gdy chce się przeprowadzić sondaż przedwyborczy lub pogłębione badania zachowań wyborczych – ale należy się wówczas upewnić co do aktualności tych rejestrów. Istnieją też listy właścicieli samochodów, osób korzystających z pomocy społecznej, płatników podatków, osób prowadzących działalność gospodarczą, licencjonowanych przedstawicieli niektórych zawodów itd. Mimo że uzyskanie dostępu do niektórych z nich może okazać się trudne, stanowią one doskonałe operaty losowania dla określonych celów badawczych.

Gdy zdamy sobie sprawę z tego, że elementy losowania nie muszą być pojedynczymi osobami, możemy zauważyć istnienie list, zawierających innego rodzaju elementy: uniwersytety, różnego rodzaju firmy, miasta, czasopisma akademickie, gazety, związki, kluby polityczne, stowarzyszenia zawodowe itd.

Do przeprowadzania błyskawicznych sondaży opinii publicznej często stosowane są książki telefoniczne. Są one niezaprzeczalnie łatwe w użyciu i niedrogie, co bez wątpienia stanowi przyczynę ich popularności. Jeśli celem badań jest formułowanie twierdzeń dotyczących abonentów telefonicznych, książka telefoniczna jest całkiem dobrym operatem losowania. (Oczywiście należy mieć świadomość tego, że dana książka nie będzie zawierała nowych abonentów lub takich, którzy zastrzegli swoje numery telefonów. Losowanie jest dodatkowo utrudnione przez to, że książki telefoniczne zawierają też numery firm). Niestety książki telefoniczne są aż nazbyt często stosowane jako spisy ludności danego miasta lub wyborców. Spośród wielu wad takiego rozumowania najważniejsza dotyczy obciążenia związanego z przynależnością do klasy społecznej. Prawdopodobieństwo posiadania telefonu jest mniejsze wśród ludzi niezamożnych, zamożni natomiast mogą mieć więcej niż jedną linię telefoniczną. Dlatego próba dobrana z książki telefonicznej z dużym prawdopodobieństwem będzie obciążona w kierunku nadreprezentowania przedstawicieli klasy średniej i wyższej.

Obciążenie klasowe właściwe próbom dobieranym z książki telefonicznej często pozostaje ukryte. Sondaże przedwyborcze przeprowadzane w ten sposób są niekiedy zupełnie trafne, być może ze względu na obciążenie klasowe samego głosowania: udział w wyborach osób uboższych jest mniej prawdopodobny. Tak więc często zdarza się, że te dwa obciążenia są zbieżne, wyniki sondażu telefonicznego mogą być zatem bardzo zbliżone do ostatecznych wyników wyborów.

[4] Możliwość ta odnosi się do warunków amerykańskich (przyp. tłum.).

Niestety nigdy nie wie się tego na pewno, aż do samych wyborów. I niekiedy, jak w przypadku sondażu „Literary Digest" z 1936 roku, można przekonać się, że wyborcy nie działali zgodnie z oczekiwanym obciążeniem klasowym. Najistotniejszym ograniczeniem tej metody jest więc niezdolność badacza do oceny poziomu błędu oczekiwanego w wynikach badań.

Do dobierania prostych prób gospodarstw domowych wykorzystywane są często spisy ulic i obszary podatkowe, ale one również nie są doskonałe i mogą być obciążone. Jest na przykład mało prawdopodobne, by w ściśle podzielonych na jednostki obszarach miejskich oficjalne dane wykazywały nielegalne jednostki mieszkalne. Wskutek tego jednostki takie nie mogą być wylosowane, a wyniki badań uzyskanych z takiej próby nie mogą być reprezentatywne dla tych jednostek, które często są biedniejsze i gęściej zaludnione od średniej dla danego obszaru.

Powyższe uwagi odnoszą się jednak do Stanów Zjednoczonych. W niektórych krajach sytuacja jest całkiem inna. Na przykład w Japonii administracja rządowa prowadzi dość dokładne rejestry ludności. Co więcej, prawo nakłada na obywateli obowiązek aktualizowania tych informacji – zgłaszania na przykład zmiany zamieszkania czy urodzeń i zgonów w gospodarstwie domowym. Wyciągnięcie prostej próby losowej ludności jest dzięki temu łatwiejsze w Japonii niż w USA. Tego rodzaju rejestr w Stanach Zjednoczonych stałby w jaskrawej sprzeczności z narodowymi normami dotyczącymi prywatności jednostki.

Przegląd populacji i operatów losowania

Literatura z zakresu badań społecznych poświęca zdumiewająco niewiele uwagi zagadnieniom populacji i operatów losowania, dlatego też w tym miejscu poświęcamy im szczególnie dużo miejsca. Oto streszczenie głównych wskazówek, które warto zapamiętać:

1. Wyniki badań opartych na próbie można uznać za reprezentatywne tylko w odniesieniu do zbioru elementów tworzących operat losowania.
2. Operaty losowania nie zawsze zawierają naprawdę wszystkie elementy, które mogłaby sugerować ich nazwa. Braki są niemal nieuniknione. Dlatego główną troską badacza powinno być określenie zakresu braków i wyrównanie ich, jeśli to możliwe. (Oczywiście badacz może zgodnie z własną intuicją pozwolić sobie na zignorowanie niewielkich, a przy tym trudnych do uzupełnienia braków).
3. Nawet jeśli wyniki są uogólniane tylko na populację tworzącą operat losowania, wszystkie jej elementy muszą mieć taką samą reprezentację w operacie losowania. Najbardziej typowa jest sytuacja, w której każdy element występuje tylko raz. Elementy pojawiające się więcej niż raz będą mieć większe prawdopodobieństwo wyboru, w próbie zaś będą one, ogólnie rzecz biorąc, nadreprezentowane.

W innym miejscu tej książki omówione zostaną bardziej praktyczne zagadnienia związane z populacjami i operatami losowania. Na przykład forma operatu – taka jak drukowana lista, zbiór kart, dyskietki komputerowe czy też taśmy magnetyczne – może mieć wpływ na łatwość jego użycia. A łatwość użycia ma niekiedy znaczenie priorytetowe przed naukowymi dywagacjami: „łatwiejsza" lista może być przedkładana nad „trudniejszą", mimo że ta ostatnia byłaby właściwsza dla danej populacji. Nie należy być zbyt dogmatycznym w tym względzie, jednak każdy badacz powinien starannie rozważyć zalety i wady takich alternatyw.

Typy doboru próby

Do tej pory skupialiśmy się na prostym doborze losowym. I rzeczywiście, statystyka używana na ogół przez badaczy społecznych zakłada taki właśnie dobór próby. Jak się jednak wkrótce przekonamy, badacze mają do wyboru kilka opcji w zakresie metody doboru próby i rzadko, jeśli w ogóle, zdarza im się stosować prosty dobór losowy. Ma to dwie przyczyny. Po pierwsze, poza najprostszymi operatami losowania, jest on niewykonalny. Po drugie, co może zaskaki-

wać, prosty dobór losowy może nie być najwłaściwszą dostępną metodą. Przejdźmy teraz do omówienia prostego doboru losowego i innych dostępnych opcji.

Prosty dobór losowy

Jak przed chwilą zauważyłem, **prosty dobór losowy** (*simple random sampling*) jest podstawową metodą, której użycie jest zakładane w obliczeniach statystycznych w badaniach społecznych. Ze względu na to, że od strony matematycznej próba losowa jest szczególnie złożona, zrezygnujemy z zajmowania się tym jej aspektem na rzecz omówienia sposobów wykorzystania tej metody do realizacji badań.

Po właściwym ustaleniu operatu losowania badacz, który chce zastosować prosty dobór losowy, numeruje kolejno wszystkie elementy listy, nie pomijając żadnych numerów. Następnie używa się tablic liczb losowych (załącznik E), aby dobrać elementy do próby. Ich użycie wyjaśnia ramka zatytułowana „Korzystanie z tablic liczb losowych".

W przypadkach, gdy operat losowania ma formę nadającą się do automatycznego odczytu, jak na przykład dyskietka komputerowa lub taśma magnetyczna, taka prosta próba losowa może zostać dobrana przez komputer. (Działa to tak, że program komputerowy sam numeruje elementy operatu losowania, generuje własne serie liczb losowych, a następnie podaje listę wybranych elementów).

Rysunek 7.10 jest ilustracją graficzną prostej próby losowej. Zauważmy, że członkowie naszej hipotetycznej mikropopulacji zostali ponumerowani od 1 do 100. Przechodząc do załącznika E, decydujemy się użyć ostatnich dwóch cyfr pierwszej kolumny i zacząć od trzeciej liczby od góry. W wyniku tego osoba numer 30 zostaje wybrana do próby jako pierwsza, osoba

Korzystanie z tablic liczb losowych

W badaniach społecznych często wskazane jest wybranie zbioru liczb losowych na podstawie tablic, takich jak ta z załącznika E. Oto jak to się robi.

Przypuśćmy, że chcesz dobrać prostą próbę losową 100 osób (lub innych jednostek) z populacji o liczebności 980.

1. Najpierw ponumeruj członków populacji: w tym przykładzie od 1 do 980. Problemem jest teraz wybór 100 liczb losowych. Gdy już to zrobisz, twoja próba będzie składać się z osób mających te numery, które wybrano. (*Uwaga*: nie jest konieczne faktyczne ponumerowanie jednostek, jeśli wiesz, ile ich jest. Jeśli na przykład dysponujesz listą, możesz zawsze przeliczyć ją po dobraniu liczb losowych).
2. Następny krok to określenie, z ilu cyfr mają składać się dobierane przez ciebie liczby. W naszym przykładzie mamy 980-osobową populację, więc potrzebne ci będą liczby trzycyfrowe, aby każdy element miał szansę znalezienia się w próbie. (Gdyby populacja liczyła 11 825 członków, musiałbyś dobierać liczby pięciocyfrowe). Chcemy zatem dobrać 100 liczb losowych z zakresu od 001 do 980.
3. Wróć teraz do pierwszej strony załącznika E. Zauważ, że zawiera ona wiele rzędów i kolumn liczb pięciocyfrowych, a takich stron jest kilka. Ta tabela przedstawia serie liczb losowych z zakresu od 00001 do 99999. Aby zastosować ją do pobrania twojej hipotetycznej próby, musisz najpierw odpowiedzieć sobie na następujące pytania:
   a. Jak stworzysz liczby trzycyfrowe z pięciocyfrowych?
   b. Według jakiego wzorca będziesz poruszał się po tabeli, wybierając swoje liczby?
   c. W którym miejscu zaczniesz?

   Na każde z tych pytań istnieje więcej niż jedna poprawna odpowiedź. Najważniejsze jest stworzenie planu i trzymanie się go. Oto przykład.
4. Umówmy się, że aby otrzymać liczby trzycyfrowe z pięciocyfrowych wybierzemy z tabeli liczby pięciocyfrowe, ale będziemy brać pod uwagę tylko trzy pierwsze cyfry każdej z tych liczb, licząc od lewej. Gdybyśmy wybrali pierwszą liczbę z pierwszej strony – 10 480 – bralibyśmy pod uwagę tylko 104. (Moglibyśmy się umówić, że będziemy brać trzy ostatnie cyfry, czyli 480, albo trzy środ-

numer 67 jako następna itd. (Osoba numer 100 byłaby wybrana, gdyby na liście wypadło „00").

Dobór systematyczny

W praktyce prosty dobór losowy jest stosowany rzadko. Jak się przekonamy, przeważnie nie jest on najbardziej efektywną metodą, może zaś być pracochłonny, jeśli jest przeprowadzany ręcznie. Z reguły prosty dobór losowy wymaga istnienia listy elementów. Gdy taka lista jest dostępna, badacze na ogół stosują zamiast prostego doboru losowego dobór systematyczny.

W **systematycznym doborze próby** (*systematic sampling*) do próby dobierany jest (systematycznie) co *k*-ty element listy. Gdyby lista zawierała 10 000 elementów, a my chcielibyśmy dobrać próbę wielkości 1000, wybieralibyśmy do próby co dziesiąty element. Aby uniknąć wszelkich możliwych obciążeń, spowodowanych działaniem czynnika ludzkiego, pierwszy element powinien być wylosowany. Zatem w odniesieniu do poprzedniego przykładu nasza procedura rozpoczęłaby się od wylosowania liczby losowej między 1 a 10. Do próby zostałby włączony element o tym numerze, plus każdy co dziesiąty element po nim. Techniczne określenie tej metody to *próba systematyczna z losowym punktem startowym*. W odniesieniu do systematycznego doboru próby stosuje się często dwa terminy. **Interwał (odstęp) losowania** (*sampling interval*) to standardowa odległość między elementami dobieranymi do próby: dla naszego przykładu wynosi on 10. **Proporcja losowania** (*sampling ratio*) to stosunek liczby elementów wylosowanych do liczebności całej populacji: w naszym przykładzie jest to 1/10.

$$\text{interwał losowania} = \frac{\text{wielkość populacji}}{\text{wielkość próby}}$$

$$\text{współczynnik losowania} = \frac{\text{wielkość próby}}{\text{wielkość populacji}}$$

kowe, czyli 048, i każdy taki plan byłby skuteczny). Najważniejsze to zrobić plan i trzymać się go. Dla wygody pozostańmy przy trzech pierwszych cyfrach.

5. Mamy również swobodę wyboru co do sposobu poruszania się w tabeli: kolumnami w dół albo w górę, wierszami od prawej do lewej lub od lewej do prawej, bądź też po przekątnej. Podobnie jak poprzednio, każdy z tych planów będzie skuteczny, jeśli będziemy go konsekwentnie stosować. Umówmy się dla wygody, że będziemy poruszać się kolumnami w dół. Gdy dojdziemy do samego dołu kolumny, przejdziemy do góry następnej; gdy wyczerpiemy całą stronę, zaczniemy od góry pierwszej kolumny na następnej stronie.
6. Zastanówmy się teraz, w którym miejscu zaczniemy? Możesz zamknąć oczy i wycelować ołówkiem w tablicę, wyznaczając na chybił trafił punkt startowy. (Wiem, że nie brzmi to naukowo, ale za to się sprawdza). Bądź też, jeśli boisz się, że zniszczysz książkę lub że w nią nie trafisz, zamknij oczy i wymyśl numer kolumny i numer wiersza. („Wybiorę liczbę w piątym rzędzie drugiej kolumny"). Zacznij od tej liczby.
7. Załóżmy, że zdecydujemy się zacząć od liczby w piątym rzędzie drugiej kolumny. Jeśli spojrzysz na pierwszą stronę załącznika E, przekonasz się, że punkt startowy to liczba 39 975. Wybraliśmy 399 jako naszą pierwszą liczbę losową, a zostało nam ich jeszcze do wybrania 99. Przemieszczając się w dół drugiej kolumny, wybieramy 069, 729, 919, 143, 368, 695, 409, 939 itd. Na dole drugiej kolumny wybieramy liczbę 104 i kontynuujemy procedurę wyboru od góry trzeciej kolumny: 015, 255 itd.
8. Widzisz, jakie to łatwe? Ale zbliża się problem. Gdy dochodzimy do kolumny piątej, pędzimy, wybierając 816, 309, 763, 078, 061, 277, 988... Chwileczkę! Mamy tylko 980 studentów na tym roku. Jak możemy dobrać liczbę 988? Rozwiązanie jest proste: pomijamy ją. Za każdym razem, gdy natkniesz się na liczbę spoza zakresu, pomijaj ją i poruszaj się dalej w ustalony sposób: 188, 174 itd. To samo rozwiązanie stosujemy, gdy jakaś liczba się powtórzy. Gdy na przykład wybierzesz ponownie liczbę 399, po prostu zignoruj ją za drugim razem.
9. To tyle. Powtarzasz tę procedurę tak długo, aż uzyskasz 100 liczb losowych. Wracając do twojej listy – próba będzie się składać z osoby numer 399, osoby numer 69, osoby numer 729 itd.

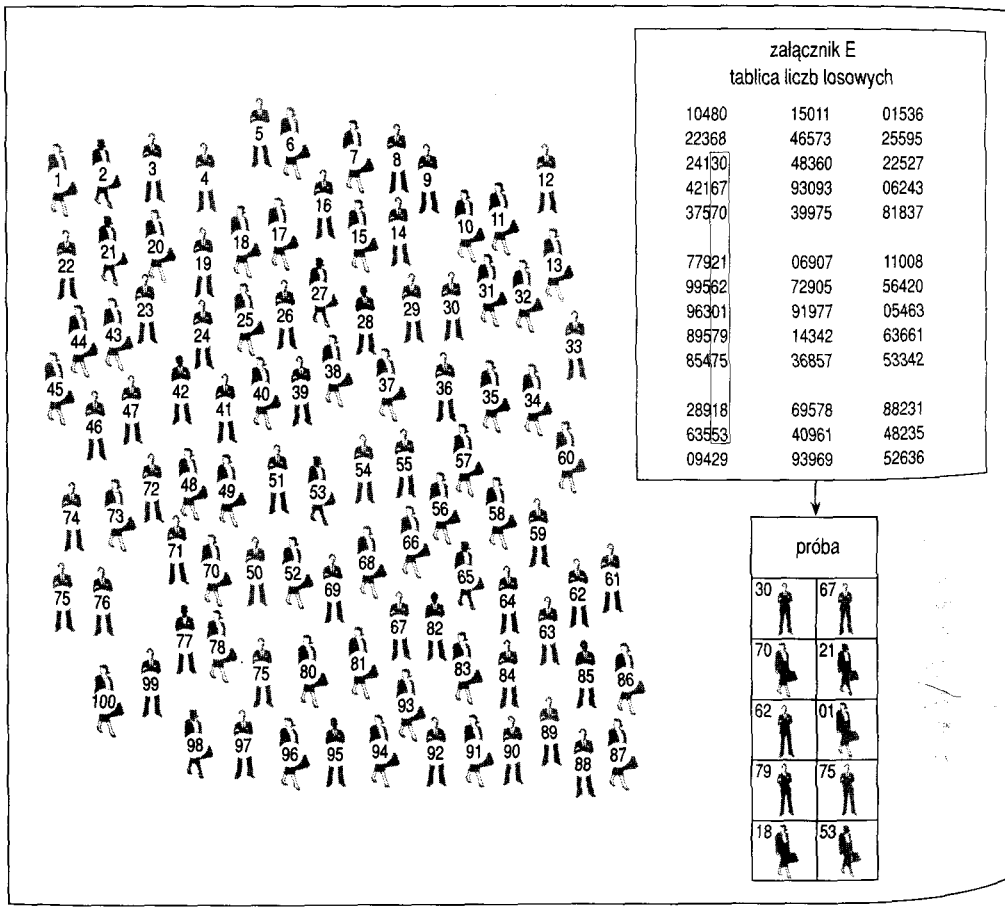
Od strony praktycznej dobór systematyczny jest w sumie identyczny z prostym doborem losowym. Jeśli lista elementów została naprawdę dobrze wymieszana przed losowaniem, można utrzymywać, że próba systematyczna dobrana z tej listy jest w istocie prostą próbą losową. Dotychczasowe spory na temat względnych zalet prostego doboru losowego i doboru systematycznego były przeważnie rozstrzygane na korzyść tej ostatniej, prostszej metody. Od strony empirycznej wyniki są w sumie identyczne. W niektórych zaś przypadkach, jak przekonamy się dalej, dobór systematyczny jest nieco dokładniejszy niż prosty dobór losowy.

Istnieje pewne niebezpieczeństwo, związane z próbą systematyczną. Układ elementów na liście może sprawić, że ten typ doboru będzie *niewskazany*. *Taki układ związany jest na ogół z cyklicznością*. Jeśli lista jest ułożona tak, że jakiś cykliczny wzór powtarza się zgodnie z interwałem losowania, może okazać się, że próba jest mocno obciążona. Oto dwa przykłady ilustrujące to niebezpieczeństwo.

W pewnych klasycznych badaniach żołnierzy podczas II wojny światowej badacze wylosowali próbę systematyczną z wykazów poszczególnych jednostek. Do badań dobierano co dziesiątego żołnierza z wykazu. Jednak spisy były ułożone w postaci tabel organizacji: najpierw sierżanci, potem kaprale i szeregowi, oddział po oddziale. Każdy oddział miał dziesięciu członków. Oznaczało to, że co dziesiąta

Rysunek 7.10

Prosty dobór losowy



osoba na liście była sierżantem oddziału. Próba systematyczna składała się z samych sierżantów. Naturalnie równie dobrze mogło okazać się, że do próby nie wszedł żaden sierżant.

Weźmy inny przykład – przypuśćmy, że dobieramy próbę mieszkań w bloku mieszkalnym. Jeśli próba byłaby dobierana z listy mieszkań, ułożonej według numerów mieszkań (na przykład 101, 102, 103, 104, 201, 202 itd.), zachodziłoby niebezpieczeństwo, że interwał losowania mógłby być zbieżny z liczbą mieszkań na piętrze lub jej wielokrotnością. I wówczas mogłyby zostać wylosowane próby składające się wyłącznie z mieszkań zlokalizowanych w północno-zachodnim rogu, albo też wyłącznie z mieszkań przy windzie. Jeśli te typy mieszkań miałyby jakieś inne wspólne cechy szczególne (na przykład wyższe czynsze), próba byłaby obciążona. Takie samo niebezpieczeństwo pojawiłoby się przy systematycznym losowaniu domów w dzielnicy składającej się z takiej samej liczby domów w każdym kwartale.

Zatem rozważając możliwość zastosowania doboru systematycznego z jakiejś listy, powinno się starannie zbadać charakter tej listy. Jeśli jej elementy są ułożone w jakimś szczególnym porządku, należy sprawdzić, czy porządek ten nie spowoduje przypadkiem obciążenia próby, a następnie należy podjąć działania zapobiegające takim możliwym obciążeniom (na przykład można dobrać prostą próbę losową z cyklicznych fragmentów).

Jednak zazwyczaj dobór systematyczny ma przewagę nad prostym doborem losowym, jeśli nie z innych przyczyn, to choćby z powodu wygody jego stosowania. Problemy związane z uporządkowaniem elementów operatu losowania można na ogół z łatwością rozwiązać.

Dobór warstwowy

Dotychczas zajmowaliśmy się dwiema metodami losowania próby z listy: losowaniem prostym i losowaniem systematycznym. **Warstwowanie** (*stratification*) nie jest alternatywą dla tych metod; daje ono raczej możliwość modyfikacji ich użycia.

Zarówno losowanie proste, jak i systematyczne zapewniają pewien stopień reprezentatywności i pozwalają na oszacowanie występującego błędu. Dobór warstwowy jest metodą pozwalającą na zwiększenie reprezentatywności poprzez obniżenie możliwego błędu z próby. Aby zrozumieć tę metodę, musimy powrócić na chwilę do podstaw teorii rozkładu z próby.

Przypomnijmy, że błąd z próby jest redukowany przez dwa czynniki związane z próbą. Po pierwsze, duża próba daje mniejszy błąd z próby niż mała próba. Po drugie, z populacji homogenicznej otrzymujemy próby obciążone mniejszym błędem niż te, które otrzymujemy z populacji heterogenicznej. Jeśli z jakimś stwierdzeniem zgadza się 99% populacji, to jest skrajnie nieprawdopodobne, że stopień tej zgody będzie znacząco różny od tej wielkości w jakiejkolwiek próbie losowej. Jeśli populacja ta jest podzielona w tej kwestii w proporcjach 50–50%, to błąd z próby będzie dużo większy.

Dobór warstwowy opiera się na tym drugim czynniku w teorii doboru próby. Badacz gwarantuje, że w próbie znajdzie się odpowiednia liczba elementów wylosowanych z homogenicznych podzbiorów tej populacji, zamiast zdawać się na losowanie próby z całej populacji. Na przykład aby otrzymać próbę warstwową studentów uniwersytetu, trzeba najpierw uporządkować tę populację według lat i kierunków studiów, i wtedy wylosować odpowiednią liczbę studentów w każdej z tych warstw. W próbie niewarstwowanej reprezentatywność ze względu na rok czy kierunek studiów będzie podlegać takiemu samemu błędowi z próby jak pozostałe zmienne. W próbie warstwowanej ze względu na rok studiów błąd z próby dla tej zmiennej będzie zmniejszony do zera.

Można również stosować bardziej złożone metody warstwowania. Oprócz warstw wyróżnionych ze względu na rok studiów, można powarstwować populację ze względu na płeć, średnią ocen itd. W ten sposób można zagwarantować, że w próbie znajdzie się właściwa liczba studentów drugiego roku legitymujących się średnią 3,5, studentek drugiego roku ze średnią 4,0 itd.

Ostatecznym celem stratyfikacji jest zatem uporządkowanie populacji w homogeniczne

podzbiory (które między sobą są heterogeniczne), a potem wylosowanie odpowiedniej liczby elementów z każdego z nich. W tym samym stopniu, w jakim podzbiory są homogeniczne pod względem zmiennych użytych do warstwowania, mogą one być homogeniczne także pod względem innych zmiennych. Ponieważ wiek jest związany z rokiem studiów, próba warstwowa, w której warstwy wyodrębniono ze względu na rok studiów, może być bardziej reprezentatywna w porównaniu z próbą niewarstwową także pod względem wieku. W takim stopniu, w jakim aspiracje zawodowe nadal związane są z płcią, próba warstwowa, w której warstwy wyodrębniono ze względu na płeć, może być bardziej reprezentatywna pod względem aspiracji zawodowych.

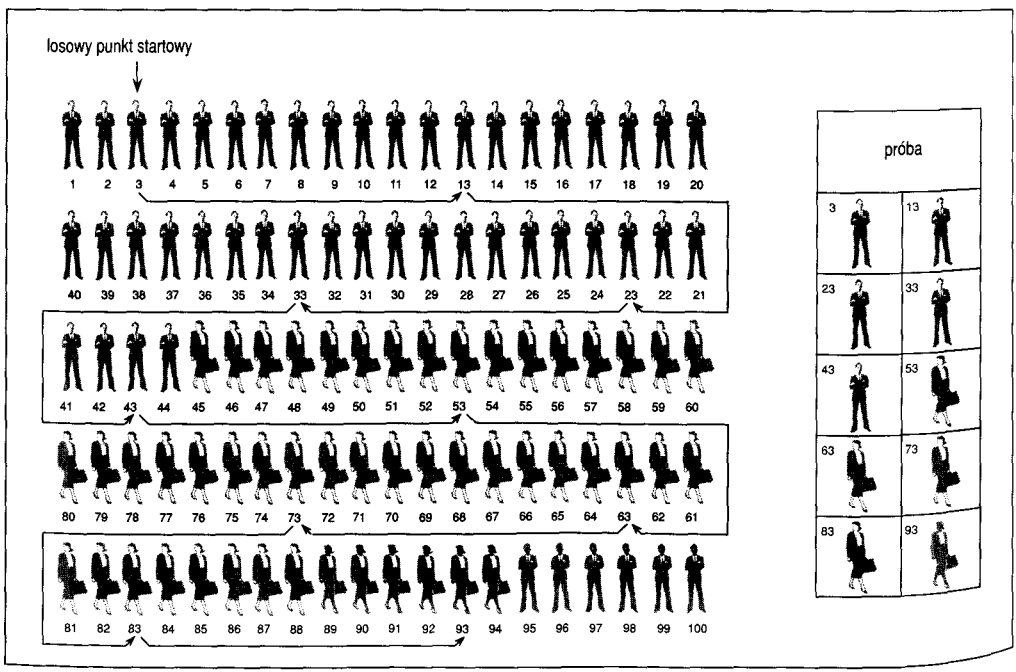
Na ogół wybór zmiennych użytych do warstwowania populacji zależy od ich dostępności. Płeć bywa często określona na listach nazwisk. Listy uniwersyteckie są z reguły ułożone latami studiów. Listy członków wydziałów wskazują na ich przynależność do poszczególnych instytutów. Akta agencji państwowych mogą być poukładane według regionów geograficznych. Listy wyborców są ułożone według obwodów głosowania.

Przy wyborze dostępnych zmiennych do warstwowania należy jednak przede wszystkim brać pod uwagę te, które z założenia są związane ze zmiennymi, pod względem których próba ma być ściśle reprezentatywna. Ponieważ płeć jest związana z wieloma zmiennymi i często jest łatwo dostępną zmienną stratyfikacyjną, jest ona często stosowana do warstwowania. Wykształcenie jest związane z wieloma zmiennymi, ale często nie jest ono zmienną dostępną. Usytuowanie geograficzne w ramach miasta czy kraju jest związane z wieloma rzeczami. Warstwowanie ze względu na położenie geograficzne w ramach miasta z reguły poprawia reprezentatywność próby pod względem klasy społecznej, grupy etnicznej itd. W obrębie kraju zwiększa ono reprezentatywność próby pod względem wielu postaw, a także pod względem klasy społecznej i grupy etnicznej.

Przy pracy ze zwykłą listą wszystkich elementów populacji przeważają dwie metody warstwowania. W pierwszej sortuje się elemen-

Rysunek 7.11

Próba warstwowa systematyczna z losowo dobranym punktem startowym



ty populacji na grupy wyodrębnione ze względu na zastosowane zmienne stratyfikacyjne. Na podstawie udziału procentowego danej grupy w populacji stosuje się dobór losowy – prosty lub systematyczny – takiej ilości elementów z tej grupy, jaka odpowiada takiemu samemu jej udziałowi procentowemu w pożądanej przez nas wielkości próby. Jeśli na przykład studenci drugiego roku studiów ze średnią 4,0 stanowią 1% populacji studentów, a potrzebna jest próba 1000 studentów, należy wylosować 10 studentów drugiego roku ze średnią 4,0.

Druga metoda to pogrupowanie studentów jak poprzednio, a następnie połączenie tych grup w jedną ciągłą listę, począwszy od studentów pierwszego roku ze średnią 5,0, a skończywszy na studentkach ostatniego roku ze średnią 2,0. Można wtedy przeprowadzić losowanie systematyczne z całej listy, z losowo dobranym punktem startowym. Ze względu na układ listy powinniśmy w wyniku doboru systematycznego otrzymać właściwą liczbę jednostek dla każdej podgrupy (z marginesem błędu 1 lub 2 jednostek). (*Uwaga*: prosta próba losowa dobrana z tak ułożonej listy zniwelowałaby warstwowanie).

Rysunek 7.11 jest graficzną ilustracją systematycznej próby warstwowej. Jak widać, mikropopulacja została uszeregowana ze względu na płeć i rasę. Następnie, począwszy od losowo dobranego numeru „3", dobierana jest co dziesiąta osoba: 3, 13, 23, ..., 93.

Dobór warstwowy zapewnia odpowiednią reprezentację zmiennych stratyfikacyjnych, co z kolei polepsza reprezentatywność pod względem innych zmiennych, które są z nimi związane. Zatem ogólnie rzecz biorąc próba warstwowa jest najczęściej bardziej reprezentatywna niż prosta próba losowa pod względem większej liczby zmiennych. Chociaż więc dobranie próby metodą prostego doboru losowego jest nadal traktowane jako świętość, powinno być już teraz jasne, że często można to zrobić jeszcze lepiej.

Ukryty podział na warstwy w doborze systematycznym

Wspomniałem już, że w pewnych okolicznościach dobór systematyczny może być trafniejszy niż prosty dobór losowy. Zachodzi to zawsze wtedy, gdy układ listy tworzy ukryte warstwy. Jak już zauważyliśmy, jeśli lista studentów jest ułożona latami studiów, dobrana z niej próba systematyczna będzie warstwowana ze względu na rok studiów, podczas gdy prosta próba losowa – nie.

W badaniach studentów Uniwersytetu Hawajskiego studenci zostali podzieleni na warstwy według roku studiów, a następnie uporządkowani według swoich studenckich numerów identyfikacyjnych. Te numery były jednak jednocześnie ich numerami ubezpieczenia społecznego. Pierwsze trzy cyfry numeru ubezpieczenia wskazują na stan, w którym nadano ten numer. Zatem w ramach danego roku studiów studenci byli uporządkowani według stanów, w których nadano im numer ubezpieczenia społecznego, co dało przybliżoną stratyfikację ze względu na pochodzenie geograficzne.

Uporządkowana lista elementów może zatem okazać się bardziej użyteczna niż lista nieuporządkowana, wymieszana. Podkreślam to ze względu na nieuprawnione przekonanie, że przed pobraniem próby systematycznej listy powinny zostać wymieszane. Jest to wskazane tylko wówczas, gdy układ list wiąże się ze wspomnianymi wcześniej problemami.

Przykład: dobór próby studentów uniwersytetu

Przenieśmy teraz te zasady na grunt praktyczny. Przyjrzyjmy się rzeczywistej procedurze zastosowanej do dobrania próby studentów uniwersytetu. Celem było przeprowadzenie sondażu za pomocą rozsyłanego pocztą kwestionariusza na reprezentatywnym przekroju studentów w głównym campusie Uniwersytetu Hawajskiego. Poniżej opisane są poszczególne etapy i decyzje związane z doborem próby do tych badań.

Populacja badana i operat losowania

Dostępnym operatem losowania, nadającym się do dobrania tej próby, był oczywiście plik komputerowy używany przez administrację uniwersytetu. Zawierał on nazwiska studen-

tów, adresy stałe i czasowe, numery ubezpieczenia społecznego i różne inne informacje, takie jak kierunek i rok studiów, wiek i płeć.

Jednak owa komputerowa baza danych zawierała dane wszystkich ludzi, których można było nazwać studentami na mocy jakiejkolwiek definicji, chociaż wielu z nich nie odpowiadało celom badania (na przykład funkcjonariusze policji, którzy uczestniczyli w weekendowych warsztatach w ramach treningu wrażliwości). Wskutek tego badacze musieli zdefiniować badaną populację w nieco ściślejszy sposób. Ostateczna definicja obejmowała 15 225 kandydatów, podejmujących studia dzienne w semestrze jesiennym w uniwersyteckim campusie Manoa, z uwzględnieniem wszystkich wydziałów i kierunków, studentów poziomu podstawowego (*undergraduate*) i wyższego (*graduate*), zarówno Amerykanów, jak i obcokrajowców. Tak więc program komputerowy użyty do losowania ograniczał swoje działanie do studentów, którzy pasowali do tej definicji.

Podział na warstwy

Program wykorzystany przy doborze próby miał również możliwość podzielenia studentów na warstwy przed dokonaniem losowania. Badacze zadecydowali, że wystarczający będzie podział na warstwy wyodrębnione ze względu na rok studiów, choć można było dokonać także warstwowania w ramach każdego roku studiów ze względu na płeć, wydział, kierunek studiów itd.

Dobór próby

Gdy uporządkowano studentów według lat studiów, wylosowano próbę systematyczną z całej tak uporządkowanej listy. Ustalona początkowo wielkość próby potrzebnej do tych badań wynosiła 1100. Aby ją uzyskać, podano programowi komputerowemu wartość proporcji losowania wynoszącą 1/14. Program wygenerował liczbę losową z zakresu od 1 do 14; do próby został wylosowany student o tym numerze i dalej co czternasty następny.

Gdy już wylosowano próbę, polecono komputerowi, by wydrukował na naklejkach samoprzylepnych nazwiska i adresy pocztowe wszystkich wylosowanych studentów. Naklejki te następnie przyklejono po prostu na koperty, w których wysyłano kwestionariusze.

Modyfikacja próby

Konieczna była modyfikacja początkowego kształtu próby. Zanim wysłano kwestionariusze, badacze zorientowali się, że niespodziewane wydatki, związane z drukiem kwestionariuszy, uniemożliwiły pokrycie kosztów wysyłki do wszystkich 1100 studentów. Wybrano więc systematycznie jedną trzecią naklejek (z losowo wybranym punktem startowym), w celu wykluczenia ich z próby. Ostateczna próba do tych badań została zatem zmniejszona do liczby 733 studentów.

Wspominam o tej modyfikacji, by ukazać często występującą potrzebę zmiany planu badań w ich toku. Ponieważ studenci, których ostatecznie wykluczono z próby, byli systematycznie wykluczani z początkowej próby systematycznej, można uważać pozostałych 733 studentów za wystarczająco reprezentatywnych dla całej badanej populacji. Zmniejszenie wielkości próby wpłynęło oczywiście na zwiększenie błędu z próby.

Dobór próby wielostopniowy grupowy

Zajmowaliśmy się poprzednio dość prostymi procedurami doboru próby z list elementów. Jest to sytuacja idealna. Jednak na nieszczęście znaczna część interesujących badań społecznych wymaga dobrania próby z populacji, dla których nie można po prostu sporządzić listy w celu wylosowania próby: populacji miasta lub kraju, wszystkich studentów w USA itd. Schemat doboru próby musi być w takich przypadkach znacznie bardziej złożony. Zakłada on na ogół podejście wielostopniowe, z początkowym losowaniem grup elementów – zwanych także gronami czy wiązkami (*clusters*) – a następnie losowaniem elementów w obrębie każdej wylosowanej grupy.

Próba grupowa (gronowa, wiązkowa, *cluster sampling*) może być stosowana wtedy, gdy spo-

rządzenie wyczerpującej listy elementów wchodzących w skład populacji, takiej jak na przykład wierni wszystkich kościołów w USA, jest niemożliwe lub bardzo trudne. Jednak elementy populacji często są już pogrupowane w subpopulacje, a listy takich subpopulacji albo już istnieją, albo można je bez trudu sporządzić. Dla przykładu, członkowie kościołów w USA należą do odrębnych kościołów, których spis albo istnieje, albo można go stworzyć. Zgodnie z zasadami doboru próby grupowej, badacze dobierają w jakiś sposób próbę z listy kościołów (na przykład systematyczną próbę warstwową). Następnie mogliby otrzymać spisy wiernych od każdego z wybranych kościołów. Potem dobrano by próbę z każdej z tych list w celu uzyskania prób wiernych należących do poszczególnych kościołów (zob. na przykład Glock, Ringer, Babbie 1967).

Inna typowa sytuacja związana jest z losowaniem w ramach obszarów zamieszkania, takich jak miasta. Mimo że nie ma jednej listy całej populacji miasta, mieszkańcy mieszkają w odrębnych dzielnicach miasta lub kwartałach spisowych. Badacze mogą zatem dobrać najpierw próbę kwartałów, stworzyć listę ludzi mieszkających w wybranych kwartałach, a potem dobrać podpróby osób mieszkających w każdym z nich.

W bardziej złożonym schemacie badacze mogą wylosować kwartały, spisać gospodarstwa domowe w każdym wybranym kwartale, wylosować próbę tych gospodarstw, spisać ludzi mieszkających w każdym gospodarstwie i w końcu dobrać pewne osoby w każdym wylosowanym gospodarstwie domowym. Schemat doboru wielostopniowego prowadzi ostatecznie do wylosowania próby osób, ale nie wymaga początkowego spisania wszystkich jednostek składających się na populację miasta.

Dobór próby wielostopniowy grupowy wymaga zatem powtarzania dwóch podstawowych kroków: spisywania i losowania. Sporządza się listę pierwotnych jednostek losowania (kościołów, kwartałów) i niekiedy dzieli się je na warstwy przed losowaniem. Potem losuje się próbę tych jednostek. Następnie robi się spis elementów tych wybranych pierwotnych jednostek losowania i niekiedy dzieli się je na warstwy. Kolejnym krokiem jest losowanie z owych list wtórnych jednostek losowania itd.

Wielostopniowy dobór grupowy umożliwia badaczowi prowadzenie badań, których w inny sposób nie można byłoby zrealizować. Przykładem może być sondaż „Święty Mikołaj", opisany w ramce zatytułowanej „Dobór próby fanów świętego Mikołaja".

Schematy doboru wielostopniowego a błąd z próby

Mimo że dobór grupowy jest bardzo efektywny, ceną za tę efektywność jest mniej dokładna próba. Prosta próba wylosowana z populacji cechuje się pojedynczym błędem z próby, ale dwustopniowa próba grupowa cechuje się dwoma błędami z próby. Po pierwsze, początkowa próba grup reprezentuje populację tych grup z uwzględnieniem zakresu błędu z próby. Po drugie, próba elementów dobranych w ramach danej grupy reprezentuje wszystkie elementy w tej grupie tylko w takim zakresie, jaki jest możliwy po uwzględnieniu błędu z próby. Tak więc badacz ryzykuje na przykład, że dobierze próbę, w której nadreprezentowane będą zamożne kwartały miasta, oraz próbę, w której nadreprezentowane będą zamożne gospodarstwa domowe w tych kwartałach. Najlepsze rozwiązanie tego problemu wiąże się z liczbą wylosowanych początkowo grup i z liczbą elementów wylosowanych w ramach tych grup.

Zazwyczaj badacze są ograniczeni całkowitą wielkością próby; można na przykład prowadzić badania na próbie ludności miasta, ograniczonej do 2000 osób. Jednak nawet zakładając to szerokie ograniczenie, ma się do wyboru liczne opcje w doborze próby grupowej. Jedną skrajnością będzie wybranie jednej grupy i w jej ramach losowanie 2000 elementów, drugą zaś skrajnością – wybranie 2000 grup, a w ich ramach po jednym elemencie. Rzecz jasna, żadna z tych dwóch opcji nie jest godna polecenia, ale między nimi leży wiele innych możliwości. Na szczęście logika rozkładu z próby dostarcza nam pewnych ogólnych wskazówek ułatwiających ten wybór.

Przypomnijmy, że błąd z próby jest redukowany przez dwa czynniki: zwiększenie wiel-

kości próby oraz zwiększenie homogeniczności losowanych elementów. Te czynniki działają na każdym poziomie wielostopniowego doboru próby. Próba grup będzie najbardziej reprezentatywna dla wszystkich grup wtedy, gdy dobiera się dużą liczbę grup i gdy wszystkie grupy są do siebie bardzo podobne. Próba elementów będzie najbardziej reprezentatywna dla wszystkich elementów danej grupy wtedy, gdy dobiera się dużą liczbę elementów z danej grupy i gdy wszystkie elementy tej grupy są bardzo do siebie podobne.

Jednak dla danej wielkości próby zwiększenie liczby grup nieodłącznie wiąże się ze zmniejszeniem liczby elementów w każdej z nich. Pod tym względem reprezentatywność grup wzrasta kosztem mniejszej reprezentatywności elementów tych grup, lub na odwrót. Na szczęście aby rozwiązać ten dylemat, można odwołać się do czynnika homogeniczności.

Z reguły elementy wchodzące w skład jakiejś naturalnej grupy w populacji są bardziej homogeniczne niż wszystkie elementy tej populacji razem wzięte. Członkowie jakiegoś jednego kościoła są bardziej do siebie podobni niż wszyscy członkowie wszystkich kościołów; mieszkańcy jakiegoś osiedla w mieście są bardziej podobni do siebie niż wszyscy mieszkańcy całego miasta. Zatem aby jakaś naturalna grupa była odpowiednio reprezentowana w próbie, może wystarczyć relatywnie niewielka liczba elementów, chociaż niezbędna może być większa liczba grup, aby odpowiednio reprezentować ich różnorodność. Ten fakt jest bardziej widoczny w skrajnym przykładzie bardzo zróżnicowanych grup, składających się z identycznych elementów w ramach każdej z nich. W takiej sytuacji duża liczba grup stanowiłaby odpowiednią reprezentację wszystkich ich członków. Mimo że taka skrajna sytuacja nigdy nie zdarza się w rzeczywistości, w większości przypadków jest ona bliższa prawdy niż jej przeciwieństwo: identyczne grupy składające się z bardzo zróżnicowanych elementów.

Ogólna wskazówka przy doborze grupowym polega zatem na tym, by maksymalizować liczbę dobieranych do próby grup, a jednocześnie zmniejszać liczbę elementów wewnątrz nich. Tę naukową wskazówkę należy jednak zrównoważyć ograniczeniem technicznym. Wydajność próby grupowej opiera się na zminimalizowaniu konieczności sporządzania listy elementów populacji. Ponieważ początkowo losowane są gru-

Dobór próby fanów świętego Mikołaja

Z nadejściem świąt Bożego Narodzenia 1985 roku „New York Times" stwierdził, że ciekawe byłoby zbadanie dzieci amerykańskich pod kątem ich wiary w świętego Mikołaja. Ze względu na brak jakiegokolwiek wykazu tych, którzy byli grzeczni i niegrzeczni, wydawcy musieli zastosować pewną sztuczkę. Oto opis tego, co zrobiono:

„Najnowszy sondaż «New York Timesa» jest oparty na wywiadach telefonicznych, przeprowadzonych w dniach 14–18 grudnia z grupą 261 dzieci w wieku od 3 do 10 lat, w całych Stanach Zjednoczonych, z wyjątkiem Alaski i Hawajów.

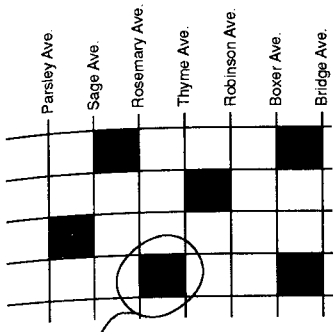
Próba central telefonicznych, pod których numery dzwoniono, została dobrana przez komputerowe losowanie z kompletnej listy wszystkich central w całym kraju. Zostały one wybrane w taki sposób, by każdy region kraju był reprezentowany w próbie proporcjonalnie do jego udziału w całej populacji kraju. Numery telefonów były generowane losowo w ramach każdej centrali, co umożliwiło wejście do próby numerom zarówno zawartym w spisach, jak i zastrzeżonym.

Po przeprowadzeniu wywiadów z 1358 dorosłymi, poproszono rodziców o zgodę na wywiady z dziećmi na temat świąt Bożego Narodzenia. Wyniki zostały przeważone ze względu na wielkość gospodarstwa domowego i liczbę linii telefonicznych oraz ze względu na występujące w próbie zróżnicowania dotyczące regionu, rasy, płci, wieku i wykształcenia.

Przy okazji okazało się, że 87% dzieci wierzy w świętego Mikołaja: od 96% dzieci w wieku 3–5 lat do 69% dzieci w wieku 9–10 lat".

Źródło: Sara Rimer, *Poll Sees Landslide for Santa: Of U.S. Children, 87% Believe*, „New York Times" 24 grudnia 1985.

Rysunek 7.12
Wielostopniowy dobór grupowy



Etap pierwszy: Określić kwartały i dobrać z nich próbę (wybrane kwartały są zacieniowane).

Etap drugi: Iść do każdego z wybranych kwartałów i sporządzić spis wszystkich gospodarstw domowych (przykład jednego kwartału, dla którego sporządzono spis).

1. 491 Rosemary Ave.
2. 487 Rosemary Ave.
3. 473 Rosemary Ave.
4. 455 Rosemary Ave.
5. 437 Rosemary Ave. ●●
6. 423 Rosemary Ave.
7. 411 Rosemary Ave.
8. 403 Rosemary Ave.
9. 1101 4th St.
10. 1123 4th St.
11. 1137 4th St. ●●
12. 1157 4th St.
13. 1169 4th St.
14. 1187 4th St.
15. 402 Thyme Ave.
16. 408 Thyme Ave.
17. 424 Thyme Ave. ●●
18. 446 Thyme Ave.
19. 458 Thyme Ave.
20. 480 Thyme Ave.
21. 498 Thyme Ave.
22. 1186 5th St.
23. 1174 5th St. ●●
24. 1160 5th St.
25. 1140 5th St.
26. 1122 5th St.
27. 1118 5th St.
28. 1116 5th St.
29. 1104 5th St. ●●
30. 1102 5th St.

Etap trzeci: Z każdego spisu dobrać próbę gospodarstw domowych (w tym przykładzie dobierano co szóste gospodarstwo, zaczynając od dobranego losowo numeru 5).

py, trzeba sporządzić jedynie listy elementów wylosowanych grup, a nie wszystkich elementów w populacji ogólnej. Jednak zwiększanie liczby grup działa akurat przeciw czynnikowi wydajności. Stworzenie list dla małej liczby grup jest szybsze i tańsze niż sporządzenie list dla dużej ich liczby. (Należy pamiętać, że konieczne jest stworzenie listy wszystkich elementów danej grupy, nawet jeśli tylko kilka z nich ma zostać wylosowanych do próby).

Schemat doboru próby, na który się ostatecznie zdecydujemy, będzie wypadkową tych dwóch ograniczeń. Zatem prawdopodobnie będziesz dobierał tak wiele grup, ile się da. Nie zostawiaj tej kwestii zbyt otwartej, oto najogólniejsza wskazówka. Z reguły badacze starają się dobierać po pięć gospodarstw domowych z jednego kwartału spisowego. Jeśli wywiady mają być przeprowadzone na próbie 2000 gospodarstw domowych, będziesz potrzebował dobrać 400 kwartałów i zrobić po pięć wywiadów w gospodarstwach domowych w każdym z tych kwartałów. Na rysunku 7.12 zaprezentowano graficzny przegląd tego procesu.

Zanim przejdziemy do innych, bardziej szczegółowych procedur grupowego doboru próby, przypomnijmy jeszcze raz, że metoda ta niemal nieuchronnie prowadzi do strat w dokładności próby. Jednak sposób, w jaki to następuje, jest dość złożony. Po pierwsze, jak zauważyliśmy wcześniej, w doborze wielostopniowym mamy do czynienia z błędem z próby na każdym etapie. Ponieważ wielkość próby, dobieranej na danym etapie, jest z konieczności mniejsza niż całkowita wielkość próby, zatem błąd z próby na danym etapie będzie większy, niż byłby w przypadku losowania jednostopniowego. Po drugie, błąd z próby jest szacowany na podstawie zaobserwowanej zmienności między elementami próby. Gdy elementy te są losowane ze względnie homogenicznych grup, oszacowanie błędu z próby będzie zbyt optymistyczne i trzeba je skorygować z uwzględnieniem grupowego doboru próby.

Warstwowanie w wielostopniowej próbie grupowej

Dotychczas omawialiśmy próbę grupową tak, jak gdyby na każdym etapie doboru stosowano prostą próbę losową. W rzeczywistości dla poprawienia próby można zastosować techniki podziału na warstwy.

Podstawowe dostępne opcje są w zasadzie takie same jak w jednostopniowym losowaniu z listy. Na przykład dobierając ogólnokrajową próbę kościołów można najpierw podzielić listę kościołów na warstwy według wyznania, regionu geograficznego, wielkości, miejskiego lub wiejskiego charakteru, a do pewnego stopnia prawdopodobnie także klasy społecznej.

Gdy już pogrupuje się pierwotne jednostki losowania (kościoły, kwartały) według odpowiednich dostępnych zmiennych stratyfikacyjnych, można zastosować losowanie zwykłe lub systematyczne. Można z każdej grupy czy warstwy wylosować określoną liczbę jednostek, bądź też można sporządzić ciągłą listę podzielonych na warstwy grup i systematycznie losować jednostki z tej listy.

Błąd z próby na tym etapie będzie zredukowany w takim zakresie, w jakim grupy są podzielone na homogeniczne warstwy. Tak jak poprzednio, pierwotnym celem stratyfikacji jest homogeniczność.

Nie ma powodu, dla którego nie można by zastosować podziału na warstwy na każdym etapie losowania. Listę elementów w ramach wybranej grupy można podzielić na warstwy przed przystąpieniem do kolejnego etapu losowania. Na ogół jednak nie robi się tego. (Przypomnijmy tu założenie względnej homogeniczności wewnątrz grup).

Dobór proporcjonalny

Wprowadzimy teraz nieco bardziej wyrafinowaną formę doboru grupowego, która jest stosowana w wielu badaniach sondażowych, prowadzonych na dużą skalę. W poprzednich rozważaniach mowa była o prostym lub systematycznym losowaniu grup, a potem prostym lub systematycznym losowaniu elementów w obrębie każdej wybranej grupy. Zauważmy, że procedura ta stanowi ogólny schemat losowania, w którym każdy element w populacji generalnej ma taką samą szansę znalezienia się w próbie.

Załóżmy, że mamy wylosować gospodarstwa domowe w jakimś mieście. Jeśli w mieście tym jest 1000 kwartałów, a dobieramy początkowo próbę 100, oznacza to, że dla każdego kwartału prawdopodobieństwo znalezienia się w próbie, wynosi 100/1000 lub 0,1. Jeśli następnie losujemy jedno na dziesięć gospodarstw domowych wchodzących w skład wybranych kwartałów, to prawdopodobieństwo

dla każdego gospodarstwa domowego, że znajdzie się w próbie z danego kwartału, wynosi 0,1. Aby obliczyć całkowite prawdopodobieństwo dla gospodarstwa domowego, że znajdzie się w próbie, po prostu mnożymy przez siebie prawdopodobieństwa na poszczególnych etapach losowania. Czyli dla każdego gospodarstwa domowego prawdopodobieństwo, że jego kwartał zostanie wylosowany, wynosi 1/10, i również 1/10 wynosi prawdopodobieństwo, że akurat to gospodarstwo domowe zostanie wylosowane do próby, gdy kwartał, do którego należy, jest jednym z wybranych. W tym przykładzie całkowite prawdopodobieństwo wyboru wynosi dla każdego gospodarstwa domowego $1/10 \times 1/10 = 1/100$. Tak dobrana próba byłaby reprezentatywna dla wszystkich gospodarstw domowych w mieście, gdyż każde z nich miałoby taką samą szansę znalezienia się w próbie.

Jednak ta procedura wiąże się z pewnym niebezpieczeństwem. Problemem jest zwłaszcza zmienność wielkości kwartałów (mierzonej liczbą gospodarstw domowych). Załóżmy, że połowa populacji miasta mieszka w 10 gęsto zaludnionych kwartałach, wypełnionych wieżowcami, a reszta populacji mieszka w domkach jednorodzinnych, usytuowanych na terenie pozostałych 990 kwartałów. Gdy wylosujemy najpierw próbę 1/10 kwartałów, ryzykujemy z dużym prawdopodobieństwem, że do próby nie trafi żaden z dziesięciu kwartałów gęsto zaludnionych. Niezależnie od tego, co się stanie w drugiej fazie losowania, nasza końcowa próba gospodarstw domowych będzie niereprezentatywna dla miasta, gdyż w jej skład wejdą tylko domki jednorodzinne.

Zawsze gdy losowane grupy znacznie różnią się wielkością, właściwe jest zastosowanie zmodyfikowanego schematu doboru próby, nazywanego doborem proporcjonalnym do wielkości (*probability proportionate to size, PPS*). Schemat ten zapobiega opisanemu wyżej problemowi, a mimo to otrzymujemy w jego wyniku próbę, w której każdy element ma taką samą szansę znalezienia się w próbie.

Jak wynika z nazwy, każda grupa ma szansę znalezienia się w próbie proporcjonalną do swojej wielkości. Zatem kwartał mający 200 gospodarstw domowych ma dwa razy większą szansę niż taki, który ma tylko 100 gospodarstw domowych. (Wkrótce przekonamy się, jak to się robi, na przykładzie kobiet chodzących do kościoła). Jednak w ramach każdej grupy losuje się stałą liczbę elementów, powiedzmy pięć gospodarstw domowych w każdym kwartale. Zauważmy, że w wyniku tej procedury każde gospodarstwo domowe ma taką samą całkowitą szansę znalezienia się w próbie.

Przyjrzyjmy się gospodarstwom domowym w dwóch różnych kwartałach miasta. Kwartał A ma 100 gospodarstw domowych, kwartał B tylko 10. W doborze proporcjonalnym kwartał A uzyskałby dziesięciokrotnie większą szansę znalezienia się w próbie niż kwartał B. Zatem jeśli w ogólnym schemacie losowania prawdopodobieństwo znalezienia się w próbie wynosi dla kwartału A 1/20, to dla kwartału B wyniesie ono tylko 1/200. Zauważmy, że oznacza to, iż prawdopodobieństwo dla wszystkich gospodarstw domowych z kwartału A, że ich kwartał zostanie wylosowany, wyniosłoby 1/20; dla wszystkich gospodarstw kwartału B wyniosłoby ono tylko 1/200.

Jeśli wylosowano kwartał A i dobieramy teraz próbę po pięć gospodarstw domowych z każdego wylosowanego kwartału, to prawdopodobieństwo dla gospodarstw domowych z kwartału A, że wejdą do próby gospodarstw domowych z tego kwartału, wynosi 5/100. Ponieważ w takich przypadkach mnożymy prawdopodobieństwa, widzimy, że całkowite prawdopodobieństwo znalezienia się w próbie dla każdego gospodarstwa domowego z kwartału A równa się $1/20 \times 5/100 = 5/2000 = 1/400$.

Gdyby natomiast wylosowano kwartał B, to gospodarstwa domowe z tego kwartału miałyby znacznie większe szanse na to, że znajdą się wśród owych pięciu gospodarstw wylosowanych do próby z tego kwartału: 5/10. Gdy jednak weźmie się pod uwagę mniejsze prawdopodobieństwo wyboru kwartału B w stosunku do kwartału A, całkowite prawdopodobieństwo znalezienia się w próbie będzie dla gospodarstw z kwartału B takie samo jak dla gospodarstw z kwartału A: $1/200 \times 5/10 = 5/2000 = 1/400$.

Dalsze udoskonalanie tej metody czyni ją bardzo wydajną i skuteczną metodą doboru

dużych prób grupowych. Na razie jednak wystarczy zrozumieć jej podstawową logikę.

Dobór nieproporcjonalny i ważenie

W ostateczności próba losowa jest reprezentatywna dla danej populacji, gdy wszystkie elementy tej populacji mają taką samą szansę znalezienia się w próbie. Dlatego w każdym z poprzednich przykładów widzieliśmy, że różne procedury losowania dawały w wyniku taką samą szansę znalezienia się w próbie – nawet jeśli ostateczne prawdopodobieństwo wyboru było wynikiem wielu prawdopodobieństw cząstkowych.

Ujmując jednak rzecz ogólniej, probabilistyczny dobór próby to taki rodzaj doboru, w którym każdy element populacji ma znane, różne od zera prawdopodobieństwo bycia wybranym – nawet jeśli prawdopodobieństwo to jest różne dla poszczególnych elementów. W przypadku kontrolowanych procedur probabilistycznego doboru próby każda taka próba może być reprezentatywna dla populacji, z której została wylosowana, jeśli każdemu pojedynczemu elementowi próby przypisze się wagę równą odwrotności jego prawdopodobieństwa znalezienia się w próbie. Jeśli zatem wszystkie elementy próby miały taką samą szansę znalezienia się w próbie, to każdy z nich ma taką samą wagę: 1. Nazywa się to próbą *samoważącą się* (*self-weighting sample*).

Niekiedy właściwe jest nadanie niektórym przypadkom większych wag niż innym. Proces ten nazywa się **ważeniem** (*weighting*). Nieproporcjonalny dobór próby oraz ważenie stosuje się na dwa podstawowe sposoby. Przede wszystkim można wylosować nieproporcjonalne próby z subpopulacji, w celu zapewnienia wystarczającej liczby przypadków do analizy z każdej z tych subpopulacji. Na przykład w obszarze podmiejskim może zamieszkiwać jedna czwarta populacji jakiegoś miasta. Może cię jednak szczególnie interesować dokładna analiza gospodarstw domowych znajdujących się na tym obszarze i wydaje ci się, że jedna czwarta całkowitej wielkości próby to zbyt mało. W wyniku tych rozważań możesz zdecydować się na wylosowanie takiej samej liczby gospodarstw domowych z tego obszaru podmiejskiego, jak z pozostałej części miasta. Zatem gospodarstwa w obszarze podmiejskim zyskują większe szanse dostania się do próby niż te usytuowane w innych częściach miasta.

Dopóki analizujesz te dwie części próby odrębnie, czy też porównawczo, nie musisz przejmować się zróżnicowanym doborem próby. Jednak gdy zechcesz połączyć obie próby, by stworzyć obraz całego miasta, musisz uwzględnić nieproporcjonalny dobór próby. Jeśli przez n oznaczymy liczbę gospodarstw domowych wylosowanych w każdym z dwóch obszarów, to dla gospodarstw domowych w obszarze podmiejskim prawdopodobieństwo wejścia do próby wyniesie n dzielone przez jedną czwartą całej populacji miasta. Ponieważ cała populacja miasta oraz wielkość próby są takie same dla obu obszarów, więc gospodarstwom domowym w obszarze podmiejskim nadana zostanie waga $1/4n$, a pozostałym gospodarstwom domowym – $3/4n$. Można uprościć tę procedurę ważenia poprzez nadanie każdemu z gospodarstw domowych wylosowanych poza obszarem podmiejskim wagi równej 3. (Wynikiem tej procedury jest proporcjonalna reprezentacja każdego elementu próby. Wartość dla populacji powinna być uwzględniona w ważeniu, jeśli chce się otrzymać szacunki dla populacji).

Oto przykład problemów, jakie mogą wystąpić, gdy nieproporcjonalnemu doborowi próby nie towarzyszy ważenie. Gdy „Harvard Business Review" zdecydowało się na przeprowadzenie wśród swoich prenumeratorów sondażu na temat molestowania seksualnego w pracy, właściwe wydawało się dobranie nadreprezentacji kobiet, gdyż wśród prenumeratorów zdecydowanie przeważali mężczyźni. Oto sposób, w jaki tę kwestię wyjaśnili G. C. Collins i Timothy Blodgett:

> „Wprowadziliśmy również inne zniekształcenie próby: dla zapewnienia reprezentatywności odpowiedzi uzyskanych od kobiet wysłaliśmy kwestionariusze do faktycznie wszystkich kobiet prenumeratorów, aby uzyskać stosunek proporcji mężczyzn do kobiet 68% do 32%.

> W wyniku tego obciążenia otrzymano odpowiedzi od respondentów, z których 52% stanowili mężczyźni, a 44% kobiety (4% nie podało swojej płci) – w porównaniu z ogólnymi proporcjami mężczyzn i kobiet wśród prenumeratorów HBR w USA wynoszącymi 93% i 7%".
>
> (1981, s. 78)

Zwróć uwagę na parę spraw w tym cytacie. Po pierwsze, dobrze by było wiedzieć, co oznacza sformułowanie „faktycznie wszystkich kobiet prenumeratorów". Najwyraźniej autorzy badania nie przesłali kwestionariuszy do wszystkich kobiet, ale nie ma żadnego wyjaśnienia, które z nich pominięto i dlaczego. Po drugie, pojęcie *reprezentatywności* nie zostało użyte w znaczeniu, w jakim zwykle się je stosuje w naukach społecznych. Autorzy badania mieli tu rzecz jasna na myśli to, że chcieli uzyskać znaczącą czy też „odpowiednio dużą" liczbę odpowiedzi od kobiet, oraz że dobranie nieproporcjonalnie dużej próby kobiet jest doskonałym sposobem osiągnięcia tego celu.

Dobierając do próby więcej kobiet, niż dałaby prosta próba losowa, autorzy byli w stanie „dobrać" wystarczającą liczbę kobiet (812), by porównywać je z mężczyznami (960). Dlatego gdy autorzy donoszą na przykład, że 32% kobiet i 66% mężczyzn zgadza się ze stwierdzeniem, że „skala molestowania seksualnego w pracy jest znacznie przeceniana", wiemy, że odpowiedzi kobiet są oparte na znaczącej liczbie przypadków. To dobrze. Ale są też problemy.

Zacznijmy od tego, że sondaże wśród prenumeratorów zawsze są problematyczne. W tym przykładzie badacze mogą co najwyżej mieć nadzieję, że ich wnioski dotyczą tego, „co sądzą prenumeratorzy «Harvard Business Review»". Ujmując rzecz swobodnie, można również traktować tę populację jako reprezentanta bardziej wyrobionej części kadry zarządzającej. Niestety ogólny poziom realizacji próby wyniósł 25%. Mimo że jest to całkiem niezły wynik jak na sondaż wśród prenumeratorów, jest to jednak niski wskaźnik w kategoriach uogólniania z próby losowej.

Poza tym nieproporcjonalny dobór próby stwarza jeszcze inny problem. Gdy autorzy stwierdzają, że 73% respondentów popiera politykę przedsiębiorstw przeciwdziałającą molestowaniu (Collins, Blodgett 1981, s. 78), wartość ta jest niewątpliwie zawyżona, gdyż udział kobiet w tej próbie jest nieproporcjonalnie duży – a kobiety są bardziej skłonne popierać taką politykę. Również wtedy, gdy badacze wykazują większe prawdopodobieństwo, że skargi na molestowanie seksualne są przesadzone w odczuciu kadry zarządzającej wyższego szczebla niż kadry zarządzającej średniego i niższego szczebla (1981, s. 81), można traktować ten wynik z podejrzliwością. Jak donoszą badacze, udział kobiet jest nieproporcjonalnie większy wśród kadry niższego szczebla. Już sam ten fakt może być powodem rzekomych różnic między kadrą różnych szczebli zarządzania. Krótko mówiąc, nieuwzględnienie nadreprezentacji kobiet w próbie wprowadza zamieszanie we wszystkich wynikach tego sondażu, które nie są zestawiane odrębnie według płci. Rozwiązaniem tego problemu mogło być przeważenie uzyskanych odpowiedzi przez płeć, we wcześniej opisany sposób.

Przykład: dobór próby kobiet chodzących do kościoła

Zakończmy omawianie doboru grupowego przykładem – będzie to przykład badań dotyczących kobiet chodzących do kościoła. Przykład ten nie jest tak złożony, jak próby losowe stosowane dla jednostek geograficznych, takich jak miasta, stany czy kraje. Niemniej jednak powinien on unaocznić różne zasady związane z doborem grupowym.

W omawianym projekcie badawczym zostałem poproszony o zbadanie postaw religijnych i społecznych kobiet, należących do wspólnot kościelnych pewnej diecezji Kościoła Episkopalnego. Miałem zbadać reprezentatywną próbę wszystkich kobiet chodzących do kościoła w tej diecezji. Ponieważ nie istniał jeden spis takich kobiet, zaprojektowałem wielostopniowy dobór próby. W pierwszej fazie losowania zastosowałem dobór proporcjonalny, dobierając kościoły w diecezji, a później dobierałem kobiety z każdego z kościołów.

Dobór kościołów

Diecezja, której dotyczyło badanie, wydaje co roku raport, zawierający spis około setki kościołów diecezjalnych z podaniem liczby wiernych. Ten spis stanowił operat losowania w jego pierwszej fazie.

Potrzebna do badania liczba respondentek wynosiła około 500, zatem zdecydowałem się wylosować 25 kościołów proporcjonalnie do ich wielkości, a następnie dobrać po 20 kobiet z każdego z nich. Aby to zrobić, ułożyłem listę kościołów w porządku geograficznym, a potem utworzyłem tabelę podobną do fragmentu pokazanego w tabeli 7.2.

Obok każdego kościoła w tabeli wpisałem liczbę jego członków. Zliczałem następnie te liczby wpisując kolejno skumulowaną liczbę członków. Ostateczna suma wyniosła około 200 000. Teraz celem było wylosowanie 25 kościołów w taki sposób, żeby dla każdego z nich prawdopodobieństwo znalezienia się w próbie było proporcjonalne do liczby wiernych. Aby tego dokonać, posłużyłem się skumulowaną liczbą członków i stworzyłem zakresy liczb dla każdego kościoła równe liczbie jego członków. Kościołowi A w tabeli przypisałem liczby od 1 do 3000; kościołowi B – od 3001 do 8000; kościołowi C – od 8001 do 9000 itd.

Losując teraz 25 liczb z zakresu od 1 do 200 000, mogłem dobrać 25 kościołów do badania. Wylosowałem 25 liczb, stosując dobór systematyczny w następujący sposób: ustaliłem interwał losowania wynoszący 8000 (200 000/25) i dobrałem losowo punkt startowy między 1 a 8000. Powiedzmy, że liczba startowa wyniosła 4538. Ponieważ liczba ta wypadła w zakresie przypisanym kościołowi B (3001–8000), wylosowany został kościół B.

Tabela 7.2

Formularz zastosowany do sporządzenia spisu kościołów

Kościół	Liczba wiernych	Skumulowana liczba wiernych
Kościół A	3000	3000
Kościół B	5000	8000
Kościół C	1000	9000

Następnie do liczby startowej dodawałem wielokrotności liczby 8000 (interwału losowania), dobierając do próby każdy kościół, w którego zakresie liczbowym wypadała kolejna liczba. Widać z tego, że każdy kościół w diecezji miał szansę znalezienia się w próbie bezpośrednio proporcjonalną do jego wielkości, mierzonej liczbą wiernych. Dla kościoła, liczącego 4000 członków, prawdopodobieństwo znalezienia się w próbie było dwukrotnie większe niż dla kościoła mającego 2000 członków, a dziesięciokrotnie większe niż dla kościoła mającego zaledwie 400 członków.

Dobór kobiet chodzących do kościoła

Po wylosowaniu kościołów postarałem się o zdobycie spisów kobiet należących do każdego z nich. Spisy te różniły się między sobą pod względem formy i zawartości. W niektórych przypadkach otrzymałem listy wszystkich członków (mężczyzn i kobiet) i konieczne było wybranie z nich samych kobiet przed rozpoczęciem losowania. Pod względem formy spisy przyjmowały różne postaci – od drukowanych list do prostokątnych kart w formacie naklejek na koperty.

Po otrzymaniu spisu członków danego kościoła obliczałem właściwy dla niego interwał losowania na podstawie liczby kobiet należących do tego kościoła, podzielonej przez przypadającą na niego wielkość próby (20). Gdy dany kościół liczył 2000 kobiet, ustalałem interwał losowania wynoszący 100. W celu wylosowania próby kobiet z tego kościoła dobierałem losowo liczbę startową i dodawałem do niej wielokrotności interwału losowania. Powtarzałem tę procedurę dla każdego kościoła, aż uzyskałem ostateczną próbę kobiet chodzących do kościoła.

Zauważmy, że ten schemat doboru próby dawał każdej kobiecie w tej diecezji taką samą szansę znalezienia się w próbie, ale tylko przy założeniu, że każdy kościół ma wśród swych członków taką samą liczbę kobiet i mężczyzn. Zakładałem w tym przykładzie, że w każdym z kościołów kobiety stanowiły połowę wiernych. Oczywiście w rzeczywistości w niektórych kościołach stanowiły one więcej, a w innych mniej niż połowę wiernych. Biorąc pod

uwagę cele tego konkretnego badania, uznałem te drobne nierówności za nieistotne.

Zastosowanie bardziej wyrafinowanej metody doboru próby w drugiej fazie rozwiązałoby ten ewentualny problem. Ponieważ dla każdego kościoła prawdopodobieństwo znalezienia się w próbie opiera się na zakładanej liczbie kobiet, można było również obliczyć interwał losowania na podstawie tego założenia, zamiast na podstawie faktycznej liczby kobiet w spisie. Gdybym założył w pierwszej fazie losowania, że do danego kościoła należy 1000 kobiet (z ogólnej liczby 2000) wiernych, mógłbym ustalić interwał losowania wynoszący 50 (1000/20). Mógłbym następnie zastosować ten interwał, dobierając respondentów bez uwzględniania faktycznej liczby kobiet wśród członków tego kościoła. Jeśli wśród wiernych byłoby rzeczywiście 1000 kobiet, wówczas prawdopodobieństwo wylosowania do próby tego kościoła byłoby trafnie określone i wylosowanoby z niego 20 kobiet. Jeśli wśród wiernych byłoby 1200 kobiet, oznaczałoby to, że kościół ten miał zbyt małą szansę znalezienia się w próbie, ale zaradzałoby temu wylosowanie 24 kobiet przy zastosowaniu ustalonego wcześniej interwału losowania. Na tej samej zasadzie, jeśli wśród wiernych tego kościoła byłoby tylko 800 kobiet, wylosowanoby ich tylko 16. W ten sposób każda kobieta w diecezji miałaby taką samą szansę znalezienia się w próbie.

Podsumowanie probabilistycznego doboru próby

Dużo miejsca poświęcono w tym rozdziale najważniejszej metodzie doboru próby w kontrolowanych badaniach sondażowych: metodzie probabilistycznej (losowej). W każdym z omawianych wariantów widzieliśmy, że elementy populacji są dobierane poprzez losowanie ze znanym niezerowym prawdopodobieństwem wyboru.

W zależności od konkretnej sytuacji badawczej probabilistyczny dobór próby może być albo bardzo łatwy, albo też bardzo trudny, czasochłonny i kosztowny. Jednak bez względu na wszystko, jest on najbardziej efektywną metodą doboru badanych elementów. Dzieje się tak z dwóch powodów.

Po pierwsze, dobór probabilistyczny pozwala na uniknięcie świadomych bądź nieświadomych obciążeń próby powstałych z winy badacza. Gdy wszystkie elementy populacji mają taką samą (lub nie taką samą, ale potem przeważoną) szansę znalezienia się w próbie, istnieje jak największe prawdopodobieństwo, że dobrana w ten sposób próba dokładnie odzwierciedla populację wszystkich elementów.

Po drugie, dobór probabilistyczny pozwala na oszacowanie błędu z próby. Chociaż żadna próba losowa nie jest nigdy doskonale reprezentatywna pod każdym względem, metody doboru kontrolowanego pozwalają badaczowi na oszacowanie oczekiwanego marginesu błędu.

W tym przydługim rozdziale zajmowaliśmy się podstawowym zagadnieniem znacznej części badań społecznych: doborem obserwacji, które powiedzą nam coś bardziej ogólnego, wykraczającego poza specyficzne aspekty, które faktycznie obserwowaliśmy. Jest to kwestia, z którą są konfrontowani badacze empiryczni, stykający się z większą liczbą działań i aktorów tych działań, niż są w stanie w pełni obserwować i rejestrować, a także badacze zachowań politycznych, którzy chcą przewidywać wyniki wyborów, ale nie mogą przeankietować wszystkich wyborców. W dalszych częściach tej książki poznamy bardziej szczegółowo sposoby radzenia sobie z tym zagadnieniem, opracowane przez badaczy społecznych.

Główne zagadnienia

- Badacze społeczni muszą wybierać obserwacje, pozwalające im dokonywać uogólnień na ludzi i wydarzenia, których nie obserwują. Często wymaga to dobierania próby, wybierania ludzi do obserwacji.
- Niekiedy można i należy stosować próby losowe, oparte na ścisłych technikach statystycznych, ale kiedy indziej właściwsze są nieprobabilistyczne techniki doboru próby.
- Nieprobabilistyczne techniki doboru próby obejmują dobór oparty na dostępności ba-

danych, dobór celowy lub arbitralny, metodę kuli śnieżnej i dobór kwotowy. Ponadto badacze zajmujący się jakąś grupą społeczną mogą korzystać z informatorów. Każda z tych technik ma swoje zastosowania, ale żadna z nich nie daje gwarancji, że dobrana w ten sposób próba jest reprezentatywna dla populacji, z której została dobrana.

- Metody probabilistycznego doboru próby są znakomitym sposobem dobierania prób reprezentatywnych z dużych, znanych populacji. Metody te są w stanie zaradzić problemowi świadomego lub nieświadomego obciążenia próby poprzez zapewnienie każdemu elementowi w populacji znanego (niezerowego) prawdopodobieństwa znalezienia się w próbie.
- Kluczem do probabilistycznego doboru próby jest losowanie.
- Nawet najstaranniej dobrana próba nigdy nie jest doskonałą reprezentacją populacji, z której została dobrana. Zawsze jest jakiś margines błędu z próby.
- Metody probabilistycznego doboru próby umożliwiają dla danej próby oszacowanie oczekiwanego błędu z próby, poprzez przewidywanie rozkładu z prób ze względu na dany parametr.
- Oczekiwany błąd w danej próbie określany jest w kategoriach poziomów ufności i przedziałów ufności.
- Operat losowania jest listą lub quasi-listą członków populacji. Jest wykorzystywany do losowania próby. Reprezentatywność próby zależy bezpośrednio od tego, w jakim stopniu operat losowania uwzględnia wszystkich członków populacji generalnej, którą próba ma reprezentować.
- Badacze mają do wyboru różne schematy losowania próby.
- Prosty dobór losowy jest logicznie najbardziej podstawową techniką doboru probabilistycznego, ale rzadko stosuje się go w praktyce.
- Dobór systematyczny wiąże się z wybieraniem co *k*-tego członka z operatu losowania. Metoda ta jest bardziej praktyczna niż prosty dobór losowy, a przy tym – poza paroma wyjątkami – funkcjonalnie równoważna.
- Stratyfikacja, proces grupowania członków populacji we względnie homogeniczne warstwy przed losowaniem, poprawia reprezentatywność próby, redukując wielkość błędu z próby.
- Wielostopniowy dobór grupowy jest relatywnie złożoną techniką doboru próby, którą stosuje się często wtedy, gdy nie istnieje lista wszystkich członków populacji. Na ogół aby uzyskać potrzebną wielkość próby, badacze muszą określić równowagę między liczbą grup a ich liczebnością. Aby zrównoważyć błąd z próby związany z procedurą wielostopniową, można zastosować podział na warstwy.
- Dobór probabilistyczny proporcjonalny do wielkości to specjalna, efektywna metoda wielostopniowego doboru grupowego.
- Gdy występują różnice w prawdopodobieństwie znalezienia się w próbie między poszczególnymi członkami populacji, to dla zapewnienia reprezentatywnego obrazu całej populacji badacze muszą nadać poszczególnym obserwacjom wagi. Generalnie waga przypisana do konkretnego członka populacji powinna być odwrotnością prawdopodobieństwa jego znalezienia się w próbie.

Kluczowe pojęcia

nieprobabilistyczny dobór próby
dobór celowy (arbitralny)
metoda kuli śnieżnej
kwotowy dobór próby
informator
probabilistyczny dobór próby
reprezentatywność
metoda równego prawdopodobieństwa wyboru
element
populacja
populacja badana
losowanie
jednostka losowania
parametr
statystyka
błąd z próby
poziom ufności
przedział ufności
operat losowania
prosty dobór losowy
dobór systematyczny
interwał (odstęp) losowania
proporcja losowania
dobór warstwowy
dobór grupowy (gronowy, wiązkowy)
dobór proporcjonalny
ważenie

Pytania i ćwiczenia

1. Powróć do omówienia sondażu Gallupa z 1948 roku, który prognozował, że Thomas Dewey pokona Harry'ego Trumana w wyborach prezydenckich. Opisz, w jaki sposób Gallup mógł zmodyfikować zastosowaną przez siebie próbę kwotową, aby uniknąć popełnienia błędu.
2. Przy użyciu załącznika E z tej książki dobierz prostą próbę losową 10 liczb z zakresu od 1 do 9876. Opisz każdy etap tego procesu.
3. Znajdź w bibliotece albo w Internecie raport z badań z zastosowaniem probabilistycznego doboru próby. Co było operatem losowania? Jaki rodzaj próby losowej został zastosowany? Czy procedury doboru próby były właściwe?
4. W jednym lub dwóch akapitach opisz etapy wielostopniowego grupowego doboru próby studentów I roku studiujących język angielski w amerykańskich college'ach i uniwersytetach.
5. Generalny Sondaż Społeczny (GSS, którego odpowiednikiem jest w Polsce PGSS – Polski Generalny Sondaż Społeczny, http://pgss.iss.uw.edu.pl) jest ostoją nowoczesnych badań społecznych, stosujących staranne metody doboru próby. Przy użyciu strony www sprawdź pierwotną wielkość próby z trzech ostatnich sondaży, a także liczbę zrealizowanych wywiadów. Rozpocznij poszukiwania na stronie GSS (http://www.icpsr.umich.edu:8080/GSS/homepage.htm) i przejdź potem do załączników.

Dodatkowe lektury

Frankfort-Nachmias Chava, Leon-Guerrero Anna, *Social Statistics for a Diverse Society*, Pine Forge Press, Thousand Oaks, Cal. 1997. Przestudiuj zwłaszcza rozdział 11. Ten podręcznik do statystyki obejmuje wiele tematów, które omawialiśmy w tym rozdziale, ale w bardziej statystycznym kontekście. Pokazuje związki między probabilistycznym doborem próby a analizą statystyczną.

Kalton Graham, *Introduction to Survey Sampling*, Sage, Newbury Park, Cal. 1983. Kalton bardziej zagłębia się w matematyczne szczegóły doboru próby, niż my robiliśmy to w tym rozdziale, ale bez usiłowania osiągnięcia takiej doskonałości jak opisany poniżej Kish.

Kish Leslie, *Survey Sampling*, Wiley, New York 1965. Bezapelacyjnie najlepsze dzieło w dziedzinie doboru próby w badaniach społecznych. Praca ta obejmuje materiał od spraw najprostszych do najbardziej złożonych i zmatematyzowanych, zarówno zaawansowanych teoretycznie, jak i całkowicie praktycznych. Fragmenty przystępne i trudne przenikają się nawzajem, gdyż Kish wyczerpuje wszystko, czego możesz chcieć lub potrzebować się dowiedzieć na temat każdego aspektu doboru próby.

Sudman Seymour, *Applied Sampling*, s. 145–194, w: *Handbook of Survey Research*, Peter H. Rossi, James D. Wright, Andy B. Anderson (red.), Academic Press, New York 1983. Doskonały, praktyczny przewodnik w zakresie doboru próby w badaniach sondażowych.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Część 3

Typy obserwacji

8. Eksperyment

9. Badania sondażowe

10. Jakościowe badania terenowe

11. Badania niereaktywne

12. Badanie ewaluacyjne

o tym, jak zgłębiliśmy strukturę samych badań, jesteśmy gotowi na odkrycie różnorodności technik obserwacyjnych dostępnych badaczom społecznym.

Zwykle myślimy o eksperymentach w kontekście nauk przyrodniczych. W rozdziale 8 zobaczymy, jak używają eksperymentów badacze społeczni. Metoda ta jest najściślej kontrolowalna spośród wszystkich, którymi będziemy się zajmować. Zrozumienie eksperymentu może ci także pomóc zrozumieć ogólną logikę badań społecznych.

Rozdział 9 opisuje badania sondażowe, jedną z najbardziej popularnych metod w naukach społecznych. Ten typ badań obejmuje gromadzenie danych poprzez zadawanie ludziom pytań w formie kwestionariuszy do samodzielnego wypełniania lub podczas wywiadów, które można z kolei przeprowadzać osobiście lub przez telefon.

Rozdział 10, zajmujący się jakościowymi badaniami terenowymi, opisuje najbardziej być może naturalną spośród stosowanych w naukach społecznych metod gromadzenia danych – bezpośrednią obserwację zjawisk społecznych w warunkach naturalnych. Jak zobaczysz, niektórzy badacze wychodzą poza czystą obserwację i stają się uczestnikami zjawisk, które badają, ponieważ zależy im na bardziej osobistym oglądzie tych zjawisk i pełniejszym ich zrozumieniu.

Rozdział 11 omawia trzy formy niereaktywnego gromadzenia danych, wykorzystujące niektóre dane dostępne wokół nas. Na przykład analiza treści jest metodą gromadzenia danych społecznych poprzez dokładne rozróżnianie i liczenie wytworów społecznych, takich jak książki, piosenki, przemówienia czy obrazy. Nie wchodząc w osobisty kontakt z ludźmi, możesz w ten sposób badać wiele różnych zjawisk społecznych. Analiza istniejących danych statystycznych jest inną metodą badania ludzi bez konieczności rozmawiania z nimi. Różne organizacje rządowe i prywatne zbierają regularnie wiele danych, które często można wykorzystać – bez zmian albo z niewielkimi modyfikacjami – do udzielenia odpowiedzi na właściwie zadane pytania. Wartościowym źródłem dla analizy socjologicznej są też dokumenty historyczne.

Rozdział 12, dotyczący badań ewaluacyjnych, zajmuje się szybko rozwijającą się dziedziną nauk społecznych, obejmującą zastosowanie modeli eksperymentalnych i quasi-eksperymentalnych do testowania interwencji społecznej w realne życie. Badania ewaluacyjne można stosować, aby sprawdzić na przykład efektywność programu odwykowego czy też efektywność nowej stołówki szkolnej. W tym samym rozdziale zajmiemy się krótko zastosowaniem wskaźników społecznych do oceny szerszych procesów społecznych.

Zanim przejdziemy do opisu powyższych metod badawczych, trzeba uczynić dwie uwagi. Po pierwsze, prawdopodobnie dojdziesz do wniosku, że przez całe swoje życie od czasu do czasu używałeś tych metod naukowych na co dzień. W pewnym sensie codziennie stosujesz jakieś formy badań terenowych. Za każdym razem, gdy próbujesz ocenić motywację lub poglądy autora na podstawie jego pism, wykorzystujesz prymitywną formę analizy treści. Często uczestniczysz przynajmniej w przypadkowych eksperymentach. Część 3 pomoże ci używać tych metod w taki sposób, by uniknąć pułapek codziennej, niekontrolowanej obserwacji.

Po drugie, żadna z powyższych metod gromadzenia danych nie stosuje się do wszystkich badanych przedmiotów i sytuacji. Na początku każdego rozdziału próbuję przedstawić kilka sugestii co do możliwych zastosowań danej metody. Nie mogę jednak przewidzieć wszystkich tematów, które mogą cię kiedyś zainteresować. Na ogół najlepiej w każdym przypadku stosować różne techniki. Każda metoda ma swoje wady, zatem stosowanie wielu metod pozwala wypełnić wszelkie luki. Jeśli zaś różne, niezależnie stosowane metody doprowadzą cię do tego samego wniosku, uzyskasz pewną formę powtórzenia badań. ■

Rozdział 8

ksperyment

Przegląd holograficzny

Eksperyment jest rodzajem obserwacji, który pozwala na badanie związków przyczynowych. Wiele eksperymentów w badaniach społecznych przeprowadzanych jest w kontrolowanych warunkach laboratoryjnych, choć mogą one również wykorzystywać naturalne wydarzenia do zbadania ich skutków społecznych.

Wprowadzenie

Tematy odpowiednie dla eksperymentu

Eksperyment klasyczny
Zmienne niezależne i zależne
Pretest i posttest
Grupy eksperymentalne i kontrolne
Eksperyment podwójnie ślepy

Dobór uczestników
Dobór losowy
Randomizacja
Dobieranie grup przez dopasowanie
Dobieranie grup przez dopasowanie czy randomizacja?

Odmiany schematów eksperymentalnych
Preeksperymentalne schematy badawcze
Trafność w badaniach eksperymentalnych

Przykłady eksperymentów

Eksperymenty „naturalne"

Mocne i słabe strony metody eksperymentalnej

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

W tym rozdziale przedstawiona zostanie metoda badawcza najczęściej kojarzona z nauką w ogóle: eksperyment. Skoncentrujemy się tutaj na eksperymencie jako formie naukowej obserwacji w naukach społecznych. Zasadniczo eksperymenty wymagają: (1) podjęcia działania i (2) obserwowania skutków tego działania. Badacze społeczni na ogół wybierają grupę osób badanych, coś z nimi robią i obserwują skutek tego, co zostało zrobione. W tym rozdziale omówione zostaną zarówno logika, jak i niektóre techniki eksperymentów w naukach społecznych.

Warto zauważyć na początku, że eksperymenty są również często stosowane w ludzkich przedsięwzięciach nienaukowych. Dla przykładu: przygotowując gulasz dodajemy sól, próbujemy, dodajemy więcej soli, ponownie próbujemy. Rozbrajając bombę przecinamy drucik, patrzymy czy bomba wybuchnie, przecinamy następny i...

Nieustannie eksperymentujemy również w naszych próbach wypracowania uogólnionego wyjaśnienia świata, w którym żyjemy. Wszystkich umiejętności nabywamy eksperymentując: umiejętności jedzenia, chodzenia, [illegible]ienia, jazdy na rowerze, pływania itd. Eksperymentując studenci dowiadują się, jak dużo muszą się uczyć, żeby odnieść sukces na uczelni. Przez eksperymentowanie profesorowie dowiadują się, jak długo muszą przygotowywać się do dobrego wykładu. Ten rozdział omawia niektóre sposoby, dzięki którym badacze społeczni używają eksperymentów do wypracowania uogólnionego wyjaśnienia. Przekonamy się, że – tak jak inne metody dostępne badaczom społecznym – eksperymentowanie ma zarówno mocne, jak i słabe strony.

Tematy odpowiednie dla eksperymentu

Eksperymenty bardziej nadają się do jednych, mniej zaś do innych celów badawczych. Są szczególnie odpowiednie w projektach wymagających relatywnie ograniczonych i dobrze zdefiniowanych pojęć oraz tez. W kontekście tradycyjnego wizerunku nauki, omówionego wcześniej, model eksperymentalny jest szczególnie odpowiedni przy testowaniu hipotez. Ponieważ eksperyment koncentruje się na ustalaniu związków przyczynowych, bardziej nadaje się do celów wyjaśniających niż opisowych.

Przypuśćmy, że chcemy znaleźć sposoby ograniczenia uprzedzeń wobec Afroamerykanów. Zakładamy, że uzyskanie wiedzy o ich wkładzie w historię Stanów Zjednoczonych może zmniejszyć poziom uprzedzeń i postanawiamy przetestować tę hipotezę eksperymentalnie. Najpierw badamy grupę eksperymentalną, żeby określić jej poziom uprzedzeń. Następnie pokazujemy tej grupie film dokumentalny o istotnym wkładzie Afroamerykanów w naukowy, literacki, polityczny oraz społeczny rozwój narodu. Na koniec ponownie mierzymy poziom uprzedzeń wobec Afroamerykanów, żeby określić, czy film faktycznie wpłynął na zredukowanie uprzedzeń.

Eksperymentowanie jest też z powodzeniem wykorzystywane do badania interakcji w małej grupie. Można zebrać małą grupę eksperymentalną i wyznaczyć jej członkom zadanie, takie jak przygotowanie rekomendacji dla spopularyzowania idei wspólnego użytkowania samochodów w celu ograniczenia ich liczby na drogach. Następnie obserwujemy, w jaki sposób grupa organizuje się i radzi sobie z problemem. W trakcie kilku takich eksperymentów możemy systematycznie różnicować rodzaj zadania albo nagród oferowanych za poradzenie sobie z nim. Poprzez obserwowanie różnic w sposobach organizowania się grup oraz ich działania w różnych warunkach możemy bardzo wiele dowiedzieć się o istocie interakcji w małej grupie oraz czynnikach na te interakcje wpływających. Czasami adwokaci przedstawiają dowody w różny sposób kilku pozorowanym ławom przysięgłych, żeby przekonać się, który sposób jest najbardziej efektywny.

Potocznie myśli się, że eksperymenty przeprowadzane są w laboratoriach. Faktycznie większość przykładów z tego rozdziału wymaga takiego otoczenia, jednak nie musi być to regułą. Badacze społeczni często badają eks-

perymenty naturalne: „eksperymenty", które dokonują się w normalnym toku społecznych wydarzeń. Będzie ich dotyczyć ostatnia część tego rozdziału.

Eksperyment klasyczny

Najczęściej stosowany rodzaj eksperymentu, zarówno w naukach przyrodniczych, jak i społecznych, składa się z trzech głównych par składowych: (1) zmiennych niezależnych i zależnych, (2) pretestu i posttestu, (3) grupy eksperymentalnej i kontrolnej. Ta część rozdziału będzie poświęcona każdej z tych składowych oraz sposobowi, w jaki są połączone ze sobą w czasie przeprowadzania eksperymentu.

Zmienne niezależne i zależne

Istotą eksperymentu jest zmierzenie wpływu zmiennej niezależnej na zmienną zależną. Z reguły zmienna niezależna przyjmuje formę bodźca eksperymentalnego, który albo występuje, albo nie. Znaczy to, że jest ona *zmienną dychotomiczną*, przyjmującą dwie możliwe wartości: obecna lub nieobecna. W tym typowym schemacie eksperymentator porównuje sytuację, w której oddziałuje bodźcem, z sytuacją, w której brak bodźca.

W przykładzie dotyczącym uprzedzeń wobec Afroamerykanów *uprzedzenia* są zmienną zależną, a *zetknięcie z historią Afroamerykanów* jest zmienną niezależną. Hipoteza badacza zakłada, że uprzedzenia zależą, przynajmniej częściowo, od braku wiedzy na temat historii Afroamerykanów. Celem eksperymentu jest zbadanie trafności tej hipotezy poprzez oddziaływanie na grupę osób odpowiednim bodźcem, takim jak film dokumentalny. Innymi słowy, zmienna niezależna to przyczyna, a zależna to skutek. Możemy więc powiedzieć, że obejrzenie filmu było przyczyną zmiany uprzedzeń, albo że redukcja uprzedzeń była skutkiem obejrzenia filmu.

Liczba zmiennych, które można wykorzystać jako zmienne zależne i niezależne w eksperymencie, jest praktycznie nieograniczona. Co więcej, dana zmienna może służyć jako zmienna niezależna w jednym eksperymencie i jako zmienna zależna w innym. Uprzedzenia są zmienną zależną w wyżej omawianym przykładzie, ale mogą być zmienną niezależną w eksperymencie badającym wpływ uprzedzeń na zachowania wyborcze.

Zarówno zmienna niezależna, jak i zależna, które chcemy wykorzystać w eksperymencie, muszą zostać operacyjnie zdefiniowane. Uzyskanie definicji operacyjnych może wymagać różnych metod obserwacji. Odpowiedzi na pytania kwestionariusza mogą być na przykład podstawą zdefiniowania uprzedzeń. Rozmawianie z Afroamerykanami lub ignorowanie ich, zgadzanie się lub nie zgadzanie z nimi to możliwe elementy operacyjnej definicji interakcji z Afroamerykanami w układzie małej grupy.

W modelu eksperymentalnym zmienne niezależne i zależne powinny być zdefiniowane operacyjnie przed rozpoczęciem eksperymentu. Czasem jednak, co będzie widoczne w powiązaniu z sondażem i innymi metodami, lepiej jest najpierw zgromadzić zróżnicowane obserwacje w trakcie zbierania danych, a później, podczas dalszych analiz, określić najbardziej użyteczne definicje operacyjne zmiennych. Jednak w końcu eksperymentowanie, podobnie jak inne metody ilościowe, wymaga ścisłych i standaryzowanych pomiarów i obserwacji.

Pretest i posttest

W najprostszym schemacie eksperymentalnym przeprowadza się pomiar zmiennej zależnej (**pretest**), poddaje uczestników wpływowi bodźca, czyli zmiennej niezależnej i ponownie mierzy się zmienną zależną (**posttest**). Odnotowana różnica między pierwszym a ostatnim pomiarem zmiennej zależnej przypisywana jest wpływowi zmiennej niezależnej.

W przykładzie dotyczącym uprzedzeń i zetknięcia z historią Afroamerykanów zaczęlibyśmy od pretestu mierzącego poziom uprzedzeń wśród uczestników eksperymentu. Dzięki użyciu kwestionariusza, w którym pytalibyśmy na przykład o postawy wobec Afroamerykanów,

można by zmierzyć poziom uprzedzeń u każdej badanej osoby oraz średni poziom uprzedzeń w całej grupie. Po pokazaniu uczestnikom filmu o historii Afroamerykanów ponownie zastosowalibyśmy ten sam kwestionariusz. Odpowiedzi uzyskane w tym poststeście pozwoliłyby nam zmierzyć końcowy poziom uprzedzeń u każdego z uczestników oraz średni poziom uprzedzeń dla całej grupy. Zaobserwowanie niższego poziomu uprzedzeń po drugim wypełnieniu kwestionariusza doprowadziłoby nas do wniosku, że film faktycznie wpłynął na zmniejszenie uprzedzeń.

W eksperymentalnych badaniach postaw, takich jak uprzedzenia, stajemy wobec szczególnego problemu praktycznego związanego z trafnością. Jak już zapewne się domyślasz, uczestnicy eksperymentu mogli odpowiadać inaczej w drugim kwestionariuszu, nawet jeśli ich postawy pozostały niezmienione. Przy pierwszym wypełnianiu kwestionariusza uczestnicy mogli być nieświadomi jego celu. Do czasu wypełniania drugiego kwestionariusza mogli się jednak domyślić, że badacze zainteresowani są poziomem ich uprzedzeń. Ponieważ nikt nie chce być postrzegany jako osoba uprzedzona, za drugim razem uczestnicy mogli „oczyścić" swoje odpowiedzi. Mogłoby się zatem wydawać, że film zredukował uprzedzenia, chociaż faktycznie wcale się tak nie stało.

Jest to przykład związany z bardziej ogólnym problemem, który nęka przeróżne formy badań w naukach społecznych: sam fakt badania może zmienić badany przedmiot. W różnych miejscach tego rozdziału zostaną omówione techniki radzenia sobie z tym problemem w przypadku eksperymentu. Pierwszą z tych technik jest zastosowanie grupy kontrolnej.

Grupy eksperymentalne i kontrolne

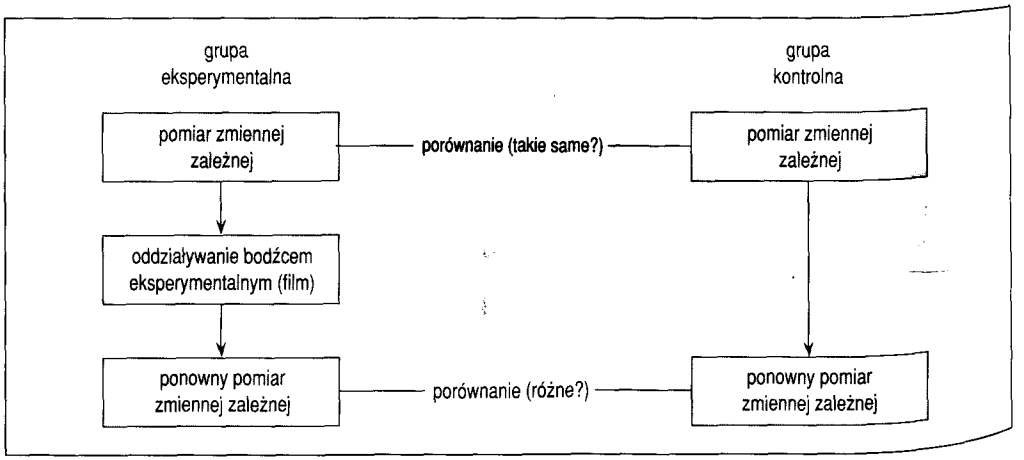
Eksperymenty laboratoryjne bardzo rzadko polegają na obserwacji samej tylko **grupy eksperymentalnej**, na którą oddziałuje się bodźcem. Badacze obserwują również **grupę kontrolną**, na którą bodźcem się nie oddziałuje.

W przykładzie o uprzedzeniach i historii Afroamerykanów można by badać dwie grupy. Na początku każda grupa wypełnia kwestionariusz skonstruowany do pomiaru uprzedzeń wobec Afroamerykanów. Następnie pokazuje się film tylko grupie eksperymentalnej. Na koniec badacze przeprowadzają posttest dotyczący uprzedzeń w obydwu grupach. Ilustracją tego podstawowego schematu eksperymentalnego jest rysunek 8.1.

Zastosowanie grupy kontrolnej pozwala badaczowi wykryć wszelki wpływ samego eksperymentu. Jeśli posttest wykaże, że ogólny poziom uprzedzeń charakteryzujący grupę kontrolną zmniejszył się w tym samym stopniu, jak w grupie eksperymentalnej, to pozorna redukcja

Rysunek 8.1

Podstawowy schemat eksperymentalny



uprzedzeń musi być raczej funkcją eksperymentu albo jakichś czynników zewnętrznych, nie zaś obejrzenia filmu. Z kolei jeżeli uprzedzenia zmalały tylko w grupie eksperymentalnej, to ten spadek będzie postrzegany jako konsekwencja obejrzenia filmu, ponieważ jest to jedyna różnica między obydwiema grupami. Albo też, jeśli uprzedzenia zmalały w obu grupach, ale w grupie eksperymentalnej bardziej niż w kontrolnej, to również może być to podstawą do założenia, że film zredukował uprzedzenia.

Potrzeba stosowania grupy kontrolnej w badaniach społecznych stała się oczywista w związku z serią badań nad satysfakcją pracowników, przeprowadzonych przez F. J. Roethlisbergera i W. J. Dicksona (1939) w późnych latach 20. i wczesnych 30. XX wieku. Badacze ci zajmowali się warunkami pracy w hali okablowywania central telefonicznych w Western Electric Works, na przedmieściach Chicago – Hawthorne, w stanie Illinois. Starali się oni odkryć, jakie zmiany w warunkach pracy zwiększyłyby satysfakcję zatrudnionych i ich produktywność.

Ku wielkiemu zadowoleniu badaczy okazało się, że polepszanie warunków konsekwentnie wpływa na wzrost satysfakcji i produktywności. Na przykład jaśniejsza, lepiej oświetlona hala powodowała, że produktywność szła w górę. Jeszcze lepsze oświetlenie, produktywność jeszcze wyższa. Wówczas, pragnąc mocniej potwierdzić swoje naukowe konkluzje, badacze *przyciemnili* światło. O! Produktywność znów wzrosła!

Stało się jasne, że robotnicy odpowiadali raczej na zainteresowanie okazywane im przez badaczy niż na ulepszone warunki pracy. W rezultacie tego zjawiska, często zwanego *efektem Hawthorne*, badacze społeczni stali się bardziej wyczuleni i bardziej ostrożni wobec możliwych skutków samych eksperymentów. Zastosowanie odpowiedniej grupy kontrolnej – intensywnie badanej bez jakiejkolwiek zmiany warunków pracy – wskazałoby na istnienie tego efektu.

Nigdzie potrzeba grupy kontrolnej w eksperymentowaniu nie jest tak duża jak w badaniach medycznych. Wielokrotnie pacjenci uczestniczący w eksperymentach medycznych czują się lepiej i nie wiadomo, na ile poprawa ich samopoczucia jest efektem eksperymentalnego leczenia, a na ile samego eksperymentu. W badaniach wpływu nowego leku badacze często podają więc grupie kontrolnej *placebo* – „lekarstwo" nie mające żadnego faktycznego wpływu, na przykład tabletki z cukru. W efekcie pacjenci z grupy kontrolnej są przekonani, że – tak jak grupa eksperymentalna – otrzymują testowane lekarstwo. Często czują się lepiej. Jeśli jednak nowy środek jest efektywny, ci, którzy go otrzymują, powinni czuć się jeszcze lepiej niż ci, którzy dostają *placebo*.

W eksperymentach przeprowadzanych w naukach społecznych grupy kontrolne zabezpieczają nie tylko przed skutkami samego eksperymentu, lecz również różnych wydarzeń zewnętrznych wobec laboratoryjnej sytuacji, w której przeprowadza się eksperyment. W przykładzie z badaniami uprzedzeń można sobie wyobrazić, że popularny przywódca Afroamerykanów został zamordowany w samym środku trwającego tydzień eksperymentu. Takie zdarzenie może przerazić uczestników badania i zmusi ich do zastanowienia się nad ich własnymi postawami wobec Afroamerykanów, co doprowadzi do zredukowania uprzedzeń. Ponieważ wpływ ten powinien ujawnić się w stopniu mniej więcej równym wśród członków grupy eksperymentalnej i kontrolnej, większe zredukowanie uprzedzeń w grupie eksperymentalnej wskazywałoby na wpływ bodźca: filmu dokumentalnego.

Czasami schemat eksperymentalny wymaga zastosowania więcej niż jednej grupy eksperymentalnej bądź kontrolnej. Poza filmem dokumentalnym można badać wpływ przeczytania książki o historii Afroamerykanów. W takim przypadku można jednej grupie pokazać film oraz kazać przeczytać książkę; inna grupa zobaczy tylko film; jeszcze inna przeczyta tylko książkę; grupa kontrolna nie zrobi nic z tych rzeczy. W takim schemacie można by było określić wpływ każdego z bodźców osobno, a także ich łączny efekt.

Eksperyment podwójnie ślepy

Podobnie jak pacjenci, którzy czują się lepiej, kiedy wydaje im się, że otrzymują nowe lekarstwo, również eksperymentatorzy niekiedy mają skłonność do przeszacowywania wyni-

ków. W badaniach medycznych eksperymentatorzy mogą częściej „widzieć" poprawę wśród pacjentów otrzymujących testowany środek niż wśród pacjentów otrzymujących *placebo* (najbardziej prawdopodobne jest to zapewne w przypadku badacza, który jest twórcą leku). **Eksperyment podwójnie ślepy** wyklucza tę możliwość, ponieważ ani badani, ani eksperymentatorzy mający kontakt z badanymi nie wiedzą, która grupa jest eksperymentalna, a która kontrolna. W przykładzie medycznym badacze odpowiedzialni za podanie leku i notowanie poprawy nie wiedzieliby, które osoby dostają lek, a które placebo. I odwrotnie, badacz wiedzący, do której grupy należą badani, nie przeprowadzałby eksperymentu.

W eksperymentach w naukach społecznych, podobnie jak w medycznych, zagrożenie wpływem badacza jest dodatkowo minimalizowane w takim stopniu, w jakim definicje operacyjne zmiennych zależnych są jasne i ścisłe. Lekarze w mniejszym stopniu mogą nieświadomie pomylić się przy odczytywaniu temperatury danego pacjenta, niż gdy oceniają, w jakim stopniu dany pacjent jest ospały. Podobnie badacz małych grup w mniejszym stopniu może pomylić się przy stwierdzaniu, kto mówił i do kogo, niż stwierdzając, czy komentarz danej osoby był nastawiony na współpracę, czy współzawodnictwo, ponieważ jest to o wiele bardziej subiektywny osąd, trudny do precyzyjnego zdefiniowania.

Jak już kilkakrotnie wskazywałem, rzadko można znaleźć definicje operacyjne i pomiary, które byłyby całkowicie precyzyjne i niedwuznaczne. Jest to jeszcze jeden powód, dla którego korzystne może być stosowanie eksperymentu podwójnie ślepego również w badaniach społecznych.

Dobór uczestników

W rozdziale 7 omówiliśmy logikę doboru próby, która wymaga, by próba była reprezentatywna dla określonej populacji. Podobne zasady odnoszą się do eksperymentu. Ponieważ większość badaczy społecznych pracuje na uczelniach, najczęstszą sytuacją jest przeprowadzanie laboratoryjnych eksperymentów wśród studentów. Eksperymentator z reguły prosi studentów uczęszczających na jego zajęcia o udział w eksperymencie albo daje ogłoszenie w uczelnianej gazecie. Uczestnicy mogą za udział otrzymać wynagrodzenie bądź nie (w rozdziale 18 będzie mowa o etycznych kwestiach związanych z proszeniem studentów o udział w tego typu badaniach).

W odniesieniu do normy uogólniania w nauce taka tendencja wyraźnie wskazuje na potencjalną wadę badań społecznych. Ujmując rzecz prosto, studenci uniwersytetu nie są typowi dla ogółu społeczeństwa. Stąd zagrożenie, że możemy dowiedzieć się wiele o postawach i działaniach studentów, ale nie o postawach i działaniach społecznych w ogóle.

Ta potencjalna wada ma jednak mniejsze znaczenie w przypadku badań wyjaśniających niż w badaniach opisowych. To prawda, że mierząc poziom uprzedzeń w grupie studentów w naszym preteście nie mielibyśmy pewności, że podobny poziom charakteryzuje całe społeczeństwo. Gdyby jednak okazało się, że film dokumentalny zredukował uprzedzenia wśród studentów, mielibyśmy większą pewność – choć nie całkowitą – że miałby podobny wpływ na całe społeczeństwo. Procesy społeczne oraz wzorce związków przyczynowych wydają się łatwiejsze do uogólnienia i stabilniejsze niż konkretne cechy, takie jak jednostkowy poziom uprzedzeń.

Poza kwestią uogólniania wyników, podstawową zasadą doboru uczestników w eksperymencie jest porównywalność grup eksperymentalnej i kontrolnej. W idealnej sytuacji grupa kontrolna powinna wyglądać tak, jak wyglądałaby grupa eksperymentalna, gdyby *nie* była wystawiona na działanie bodźca. Logika eksperymentu wymaga zatem, by grupy eksperymentalna i kontrolna były do siebie jak najbardziej podobne. Jest kilka sposobów, którymi można to osiągnąć.

Dobór losowy

Omówienie logiki i technik losowego doboru próby, przedstawione w rozdziale 7, pokazało metodę doboru dwóch podobnych grup ludzi. Z ope-

ratu losowania, składającego się z wszystkich ludzi w badanej populacji, badacz może dobrać losowo dwie próby. Jeśli obie próby są podobne do całej populacji, z której zostały dobrane, będą również podobne do siebie nawzajem.

Trzeba jednak pamiętać, że podobieństwo (reprezentatywność) osiągnięte przez losowanie jest w dużej mierze funkcją wielkości próby. Ogólnie rzecz biorąc, próby losowe mniejsze od 100 nie są szczególnie reprezentatywne, eksperymenty zaś w naukach społecznych rzadko angażują tak dużo osób do grupy eksperymentalnej czy kontrolnej. Dobór losowy jest więc rzadko stosowany w eksperymentach dla doboru jednostek z większej populacji. Logika losowego doboru jest jednak wykorzystywana przez badaczy przy przydzielaniu uczestników do grup.

Randomizacja

Eksperymentator, dokonawszy w jakikolwiek sposób wyboru osób, może losowo przydzielić je do grupy eksperymentalnej albo kontrolnej. Taką **randomizację** (czyli „ułosowienie") można przeprowadzić numerując kolejno wszystkich uczestników, a później wybierając numery za pomocą tablicy liczb losowych. Eksperymentator może też przydzielić osoby z numerami nieparzystymi do grupy eksperymentalnej, a z parzystymi do kontrolnej.

Wracając do podstawowej zasady doboru losowego: *jeśli eksperymentator wybrał 40* osób, dając na przykład ogłoszenie w gazecie, nie mamy żadnych podstaw, by wierzyć, że tych 40 osób może reprezentować całą populację, z której pochodzi. Nie możemy też założyć, że 20 osób losowo przyporządkowanych do grupy eksperymentalnej może być grupą reprezentatywną dla całej populacji. W większym stopniu możemy być jednak pewni, że 20 osób losowo przyporządkowanych do grupy eksperymentalnej będzie podobnych do 20 osób przyporządkowanych do grupy kontrolnej.

Nawiązując do logiki naszych wcześniejszych uwag dotyczących doboru losowego: możemy potraktować te 40 osób jako populację, z której losowo dobieramy dwie próby – każdą składającą się z połowy populacji. Ponieważ każda próba odzwierciedla cechy całej *populacji, to te dwie próby odzwierciedlają siebie* nawzajem.

Jak widzieliśmy w rozdziale 7, założenie o podobieństwie tych dwóch grup częściowo zależy od liczby osób. W przypadku skrajnym, jeśli wybralibyśmy tylko dwie osoby, poprzez rzut monetą określając jedną jako jednostkę eksperymentalną, a drugą jako kontrolną, nie byłoby podstaw, by sądzić, że te dwie osoby są do siebie podobne. Jednak wraz ze wzrostem liczby badanych osób randomizacja ma sens.

Dobieranie grup przez dopasowanie

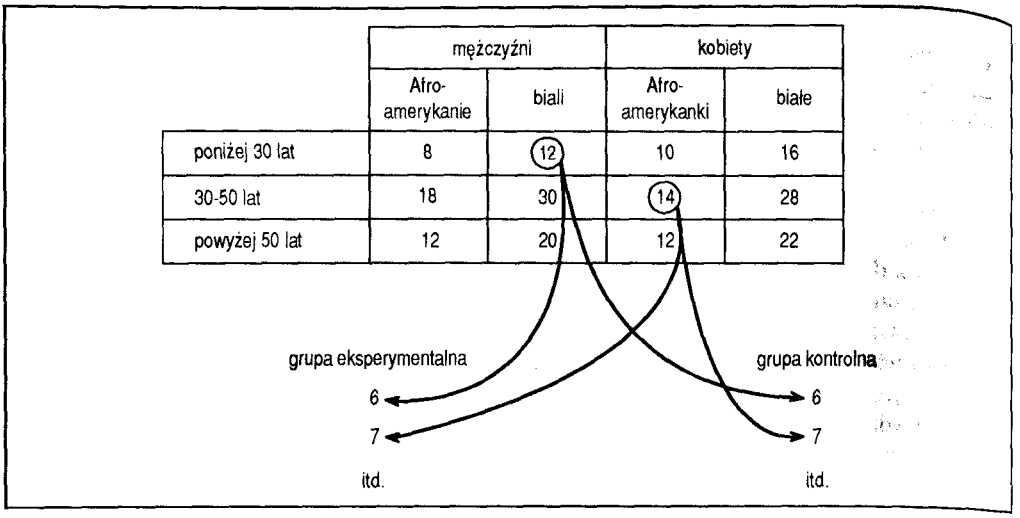
Innym sposobem uzyskiwania porównywalności między grupą eksperymentalną a kontrolną jest **dopasowywanie** (*matching*). Proces dopasowywania podobny jest do doboru kwotowego, który został przedstawiony w rozdziale 7. Jeśli naszych 12 uczestników to młodzi biali mężczyźni, to można przydzielić losowo 6 do grupy eksperymentalnej i 6 do kontrolnej. Mając 14 Afroamerykanek w średnim wieku, możemy przydzielić po 7 do każdej grupy. Możemy powtarzać ten proces dla każdego istotnego z punktu widzenia badań grupowania uczestników.

Proces dopasowywania będzie bardziej efektywny, jeśli przygotujemy macierz kwotową na podstawie wszystkich istotnych właściwości. Rysunek 8.2 pokazuje uproszczony *przykład takiej macierzy.* W przykładzie tym eksperymentator zadecydował, że istotne cechy to rasa, wiek i płeć. Idealną sytuacją jest zbudowanie macierzy w taki sposób, aby w każdej komórce znajdowała się parzysta liczba przypadków. Wówczas połowa osób z każdej komórki trafia do grupy eksperymentalnej, a druga połowa do kontrolnej.

Można również zrekrutować więcej osób, niż tego wymaga nasz schemat eksperymentu. Następnie analizuje się wiele cech osób należących do grupy początkowej. W przypadku znalezienia pary podobnych do siebie osób, jedną osobę przyporządkowujemy losowo do grupy eksperymentalnej, a drugą do kontrolnej. Osoby niepodobne do innych nie biorą udziału w eksperymencie.

Rysunek 8.2

Przykład macierzy kwotowej



Niezależnie od tego, którą metodę wybierzemy, wynik, do którego dążymy, jest taki sam: ogólny, średni opis grupy eksperymentalnej powinien być taki sam jak grupy kontrolnej. Na przykład średnia wieku w obydwu grupach powinna być mniej więcej taka sama, taki sam powinien być układ płci i rasy itd. Taki test porównywalności powinien być stosowany bez względu na to, czy dwie grupy skonstruowano przez dobór losowy, czy przez randomizację.

Dotychczas odwoływałem się do „istotnych" zmiennych bez jasnego określenia, które to zmienne. Oczywiście zmiennych tych nie można określić w żaden definitywny sposób, bardziej definitywny niż mogłem to zrobić w rozdziale 7, pisząc o zmiennych, które powinny być użyte w warstwowym doborze próby. O tym, które zmienne są istotne, decyduje ostatecznie charakter i cel eksperymentu. Ogólną zasadą jest jednak, że grupa eksperymentalna i kontrolna muszą być porównywalne ze względu na takie zmienne, które najprawdopodobniej związane są z poddawaną badaniu zmienną zależną. Na przykład w badaniach nad uprzedzeniami obie grupy powinny być podobne między innymi pod względem wykształcenia, pochodzenia etnicznego i wieku. Co więcej, w niektórych przypadkach można opóźnić przypisanie osób do grup eksperymentalnej i kontrolnej do czasu początkowego pomiaru zmiennej zależnej. Można więc na przykład przeprowadzić pomiar uprzedzeń na podstawie kwestionariusza, a następnie utworzyć grupę eksperymentalną i kontrolną, tak aby mieć pewność, że obydwie grupy wykazują ten sam ogólny poziom uprzedzeń.

Dobieranie grup przez dopasowanie czy randomizacja?

Przyporządkowując badanych do grupy eksperymentalnej i kontrolnej należy mieć na względzie dwa argumenty na rzecz przewagi randomizacji nad dopasowywaniem. Po pierwsze, badacz może nie wiedzieć z góry, które zmienne będą istotne w procesie dopasowywania. Po drugie, większość statystyk stosowanych przy analizie wyników eksperymentu zakłada randomizację. Zatem jeżeli nie uda się zaprojektować eksperymentu w ten sposób, zastosowanie później tych statystyk będzie miało mniejsze znaczenie.

Jednak randomizacja ma sens tylko wtedy, gdy mamy do dyspozycji wystarczająco dużą grupę osób i można zastosować prawa doboru losowego. W sytuacji, gdy mamy tylko kilka osób, lepszą procedurą będzie dopasowywanie.

Czasami badacze mogą połączyć randomizację i dopasowywanie. Na przykład J. Milton Yinger ze współpracownikami (1977), przeprowadzając eksperyment nad wzbogaceniem edukacji nastolatków, musiał przypisać dużą liczbę uczniów 13- i 14-letnich do kilku różnych grup eksperymentalnych i kontrolnych w taki sposób, aby mieć pewność co do porównywalności uczniów tworzących każdą z grup. Cel ten osiągnięto w następujący sposób.

Mając do dyspozycji pewną grupę osób, badacze najpierw skonstruowali warstwy uczniów bardzo do siebie podobnych pod względem około 15 zmiennych. Z każdej warstwy uczniowie byli następnie losowo przyporządkowywani do różnych grup eksperymentalnych i kontrolnych. W ten sposób badacze udoskonalili tradycyjną randomizację. W gruncie rzeczy zastosowali procedurę losowania warstwowego (rozdział 7), z tą różnicą, że użyli znacznie więcej zmiennych stratyfikacyjnych, niż czyni się to w przypadku typowego doboru w badaniach sondażowych.

Do tej pory mowa była o eksperymencie klasycznym – schemacie, który najlepiej przedstawia logikę analizy związków przyczynowych w warunkach laboratoryjnych. Jednak w praktyce badacze społeczni stosują dużo bardziej zróżnicowane schematy eksperymentalne. Przyjrzyjmy się teraz kilku z nich.

Odmiany schematów eksperymentalnych

Donald Campbell i Julian Stanley (1963) w klasycznej książce o projektowaniu badań przedstawiają 16 schematów eksperymentalnych i quasi-eksperymentalnych. W tej części zostanie opisanych kilka z nich, dla lepszego pokazania potencjału eksperymentów w badaniach społecznych.

Preeksperymentalne schematy badawcze

Campbell i Stanley przedstawili najpierw trzy schematy „preeksperymentalne", nie dlatego, by je polecać, lecz dlatego, że są one często stosowane w mniej profesjonalnych badaniach. Schematy te nazywane są „preeksperymentalnymi" w celu zasygnalizowania, że nie spełniają one naukowych standardów schematów eksperymentalnych. W pierwszym z tych schematów – studium przypadku z pojedynczym pomiarem (*one-shot case study*) – badacz dokonuje pomiaru jednej grupy badanych pod względem zmiennej zależnej po wprowadzeniu jakiegoś bodźca eksperymentalnego. Załóżmy, że pokazujemy wspominany wcześniej film o historii Afroamerykanów grupie osób, a następnie dajemy im do wypełnienia kwestionariusz, który ma mierzyć uprzedzenia wobec Afroamerykanów. Załóżmy, że odpowiedzi w kwestionariuszu wskazują na niski poziom uprzedzeń. Może nas to skłaniać do wniosku, że film obniżył uprzedzenia. Jednak – nie mając wyników pretestu – nie możemy być tego pewni. Być może kwestionariusz nie jest w rzeczywistości zbyt czułym miernikiem uprzedzeń albo grupa poddana badaniu była w niewielkim stopniu uprzedzona od samego początku. W obu przypadkach film mógł nie spowodować żadnej różnicy, choć wyniki eksperymentu mogły skłaniać nas do wniosku, że to uczynił.

Drugi schemat preeksperymentalny omówiony przez Campbella i Stanleya dodaje pretest dla grupy eksperymentalnej, ale nie ma w nim grupy kontrolnej. W tym przypadku – nazwanym przez autorów schematem z dwukrotnym pomiarem w jednej grupie (*one-group pretest-posttest design*) – istnieje możliwość, że jakiś czynnik inny niż zmienna niezależna mógł spowodować zmianę między pretestem a posttestem, tak jak na przykład zabójstwo szanowanego afroamerykańskiego przywódcy. Zatem chociaż widzimy spadek uprzedzeń, nie możemy być pewni, że to film był tego przyczyną.

Campbell i Stanley opisując możliwości schematów preeksperymentalnych zwrócili uwagę, że niektóre badania opierają się na grupie eksperymentalnej i kontrolnej, ale nie mają pretestu. Ten typ nazwali statycznym porównaniem międzygrupowym (*static-group comparison*). Można na przykład pokazać film o Afroamerykanach jednej grupie, drugiej zaś nie, a potem zmierzyć uprzedzenia w obydwu gru-

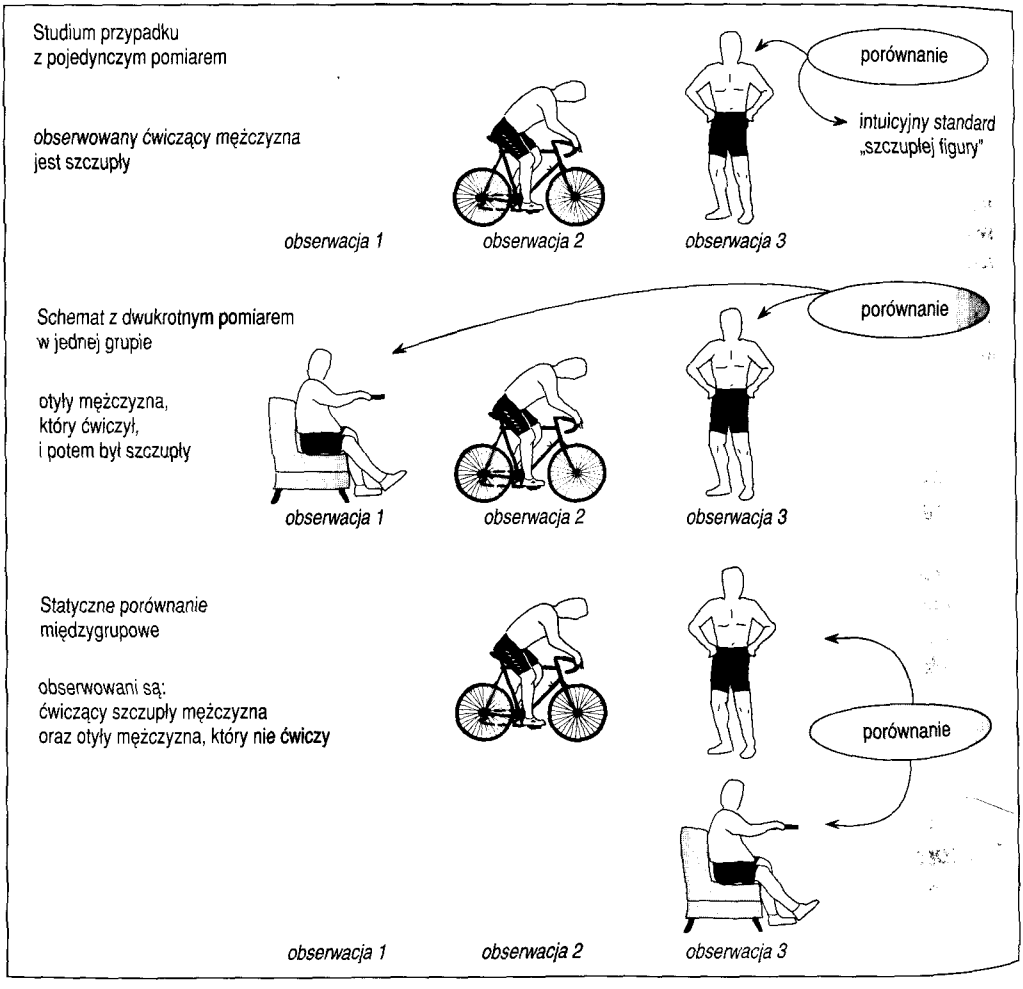
pach. Jeżeli na zakończenie eksperymentu grupa eksperymentalna wykazałaby się niższym poziomem uprzedzeń, można by wnioskować, że przyczyną był film. Jeżeli jednak nie dokonalibyśmy na początku randomizacji badanych, nie wiedzielibyśmy, czy na pewno obie grupy miały na początku ten sam stopień uprzedzeń; być może grupa eksperymentalna od początku była mniej uprzedzona.

Rysunek 8.3 przedstawia graficznie trzy schematy preeksperymentelne z tym samym pytaniem badawczym: czy ćwiczenia fizyczne powodują utratę wagi? Dla przejrzystości przykładów rysunki przedstawiają pojedyncze osoby, a nie grupy, lecz ta sama logika odnosi się do porównań między grupami. Przyjrzyjmy się raz jeszcze trzem schematom preeksperymentalnym na podstawie nowego przykładu.

Studium przypadku z pojedynczym pomiarem jest przykładem popularnej formy logicznego wnioskowania w życiu codziennym. Gdy ktoś zapyta nas, czy ćwiczenia powodują spadek wagi, może nam przyjść do głowy jakiś przykład, który, jak się wydaje, potwierdza tę tezę: *ktoś, kto ćwiczy i jest szczupły*. Z takim rozumowaniem jest jednak kilka problemów. Być może dana osoba była szczupła długo przed tym, zanim zaczęła ćwiczyć; może stała

Rysunek 8.3

Trzy schematy preeksperymentalne



się szczupła z innego powodu, na przykład zaczęła mniej jeść albo zachorowała. Obserwacje pokazane na rysunku nie zabezpieczają przed tymi innymi możliwościami. Co więcej, obserwacja, że mężczyzna na rysunku jest szczupły, zależy od naszego intuicyjnego rozumienia szczupłej i otyłej figury. Reasumując, jest to bardzo słaby dowód na sprawdzanie związku między ćwiczeniami a utratą wagi.

Schemat z pretestem i posttestem w jednej grupie daje nieco mocniejsze dowody na to, że ćwiczenia powodują spadek wagi. Przede wszystkim wykluczyliśmy możliwość, że mężczyzna był szczupły przed rozpoczęciem ćwiczeń. Wciąż jednak nie mamy pewności, że to ćwiczenia spowodowały spadek wagi.

I wreszcie, statyczne porównanie międzygrupowe eliminuje problem naszej wątpliwej definicji tego, czym jest szczupłość i otyłość. W tym przypadku porównujemy figurę mężczyzny, który ćwiczy, oraz mężczyzny niećwiczącego. Taki schemat jednak ponownie otwiera możliwość, że mężczyzna ćwiczący był szczupły już wcześniej.

Trafność w badaniach eksperymentalnych

W tym miejscu chciałbym przedstawić w bardziej systematyczny sposób czynniki, które wpływają na trafność badań eksperymentalnych. Na początek zostaną przedstawione źródła *nietrafności wewnętrznej*, nazwane tak przez Campbella i Stanleya, zrewidowane i rozwinięte przez Thomasa Cooka i Donalda Campbella (1979). Następnie omówię problem *nietrafności zewnętrznej*, odnoszący się do możliwości uogólniania wyników eksperymentu na rzeczywistość pozaeksperymentalną. Po ich przeanalizowaniu będziemy w stanie docenić zalety niektórych spośród bardziej wyrafinowanych schematów eksperymentalnych i quasi-eksperymentalnych, stosowanych czasami przez badaczy społecznych.

Źródła nietrafności wewnętrznej

Problem **nietrafności wewnętrznej** odnosi się do możliwości, że wnioski wyciągnięte z wyników eksperymentu mogą nie być dokładnym odbiciem tego, co działo się w samym eksperymencie. Zagrożenie nietrafnością wewnętrzną zachodzi zawsze, gdy cokolwiek innego niż bodziec mogło wpłynąć na zmienną zależną.

Campbell i Stanley (1963, s. 5–6) oraz T. Cook i Campbell (1979, s. 51–55) wymienili wiele źródeł nietrafności wewnętrznej. Oto 12 z nich:

1. *Historia*. W trakcie przeprowadzania eksperymentu mogą zachodzić wydarzenia historyczne, które zakłócą wyniki. Jednym z przykładów może być zamordowanie afroamerykańskiego przywódcy podczas trwania eksperymentu nad redukowaniem uprzedzeń wobec Afroamerykanów; innym zaś aresztowanie afroamerykańskiego przywódcy za jakieś haniebne przestępstwo, co może spowodować wzrost uprzedzeń.
2. *Dojrzewanie*. Ludzie nieustannie rozwijają się i zmieniają, i takie zmiany mogą oddziaływać na wyniki eksperymentu. W długotrwałym eksperymencie fakt, że badani starzeją się (i mądrzeją?), może mieć znaczenie. W krótszym eksperymencie badani mogą zmęczyć się, znużyć, znudzić, zgłodnieć bądź w inny sposób zmienić się, co wpłynie na ich zachowanie w tym eksperymencie.
3. *Pomiar*. Przeprowadzanie i powtarzanie pomiaru często wpływa na ludzkie zachowanie, tym samym zniekształcając wyniki eksperymentu. Załóżmy, że jako sposób pomiaru uprzedzeń wybieramy wypełnianie kwestionariusza. Następnie oddziałujemy bodźcem eksperymentalnym i ponownie mierzymy uprzedzenia. W czasie przeprowadzania posttestu uczestnicy będą prawdopodobnie bardziej wyczuleni na kwestię uprzedzeń i udzielą bardziej przemyślanych odpowiedzi. Mogą nawet domyślać się, że badamy, na ile są uprzedzeni, a ponieważ mało kto chce uchodzić za osobę uprzedzoną, mogą odpowiadać w sposób, jakiego ich zdaniem oczekujemy, bądź jaki pozwoli im dobrze się zaprezentować.
4. *Narzędzia*. Proces pomiaru w preteście i posttescie związany jest z niektórymi zagadnieniami, dotyczącymi konceptualizacji

i operacjonalizacji, które omówiliśmy w tej książce wcześniej. Jeśli użyjemy różnych miar zmiennej zależnej w preteście i posttescie (na przykład różnych kwestionariuszy dotyczących uprzedzeń), czy możemy być pewni, że są one porównywalne? Może nam się wydawać, że uprzedzenia zmalały po prostu dlatego, że pretest był bardziej wrażliwy niż posttest. Możliwe jest też, jeśli pomiar przeprowadzał eksperymentator, że to jego standardy bądź umiejętności zmieniły się w trakcie badań.

5. *Regresja statystyczna*. Czasem wskazane jest przeprowadzenie eksperymentu na osobach z krańcowymi początkowymi wartościami zmiennej zależnej. Testując nową metodę nauczania matematyki wyjątkowo niezdolnych matematycznie osób, do eksperymentu wybierzemy uczniów, którzy mieli bardzo słabe oceny z matematyki. Rozważmy jednak, co może się stać z matematycznymi osiągnięciami tych osób bez żadnej eksperymentalnej ingerencji. Były tak słabe, że mogą tylko albo stać w miejscu, albo poprawić się: gorzej być nie może. Nawet bez bodźca grupa będzie wykazywać więc jakąś poprawę w czasie. Używając pojęcia *regresji do średniej*, statystycy zwracają często uwagę, że wyjątkowo wysokie osoby jako grupa często mają dzieci niższe od siebie, a osoby wyjątkowo niskie jako grupa mają dzieci od siebie wyższe. Istnieje więc niebezpieczeństwo, że zmiany, które zaszły ze względu na dobór osób przystępujących do eksperymentu ze skrajnych pozycji, zostaną błędnie przypisane wpływowi bodźca eksperymentalnego.
6. *Obciążenia doboru*. Obciążony dobór przedstawiony został wcześniej, przy omawianiu różnych sposobów doboru osób do eksperymentu oraz przyporządkowywania do grupy eksperymentalnej i kontrolnej. Porównania nie będą miały sensu, jeśli grupy nie były porównywalne od samego początku.
7. *Wymieranie grup*. Choć, być może, niektóre eksperymenty są zabójcze, pojęcie *wymierania* odnosi się do bardziej ogólnej i mniej skrajnej kwestii. Zdarza się często, że badani wypadają z eksperymentu przed jego zakończeniem, co wpływa na statystyczną porównywalność oraz konkluzje. Wyobraźmy sobie, że w klasycznym eksperymencie z grupą eksperymentalną i kontrolną, z pretestem i posttestem, fanatycy z grupy eksperymentalnej są tak urażeni filmem o historii Afroamerykanów, że rezygnują z udziału w badaniach. Osoby, które wytrzymały do czasu przeprowadzenia posttestu, od samego początku były w małym stopniu uprzedzone, więc wyniki grupy pokażą wyraźny „spadek" uprzedzeń.
8. *Przyczynowy porządek czasowy*. Może zajść – choć w badaniach społecznych to rzadkość – dwuznaczność dotycząca porządku czasowego bodźca oraz zmiennej zależnej. Gdy tak jest, wniosek z badań o wpływie bodźca na zmienną zależną zderza się z wyjaśnieniem, iż faktycznie to zmienna „zależna" spowodowała zmiany w bodźcu (patrz rozdział 3).
9. *Rozpowszechnianie lub naśladowanie*. Gdy członkowie grupy eksperymentalnej i kontrolnej mogą się ze sobą kontaktować, osoby z grupy eksperymentalnej mogą przekazać jakieś elementy bodźca grupie kontrolnej. Wyobraźmy sobie, że między pokazaniem filmu o historii Afroamerykanów a przeprowadzeniem posttestu poprzez wypełnienie kwestionariuszy jest pewien odstęp czasu. Członkowie grupy eksperymentalnej mogą opowiedzieć osobom z grupy kontrolnej o filmie. Wówczas również grupa kontrolna poddana jest wpływowi bodźca i nie jest już faktyczną grupą kontrolną. Czasem mówi się w tym kontekście o grupie kontrolnej, że została „skażona".
10. *Rekompensata*. W rozdziale 12 zobaczymy, że w eksperymentach w rzeczywistych sytuacjach – takich jak specjalny program edukacyjny – osoby z grupy kontrolnej często pozbawione są czegoś, co postrzegane jest jako wartościowe. W takich sytuacjach może pojawić się nacisk na zaoferowanie im jakiejś formy rekompensaty. Personel szpitalny na przykład, współczując

pacjentom z grupy kontrolnej, może otaczać ich wyjątkowo ciepłą i serdeczną opieką. W takich przypadkach grupa kontrolna przestaje być prawdziwą grupą kontrolną.

11. *Rywalizacja o rekompensatę*. W eksperymentach przeprowadzanych w rzeczywistych sytuacjach osoby pozbawione bodźca eksperymentalnego mogą starać się zrekompensować jego brak cięższą pracą. Jeśli bodźcem jest eksperymentalny program nauczania matematyki, grupa kontrolna może przykładać się do nauki matematyki bardziej, niż to robiła wcześniej, po to, żeby pokonać grupę eksperymentalną.
12. *Zniechęcenie*. Z kolei poczucie straty w grupie kontrolnej może spowodować wśród jej członków załamanie. W eksperymentach edukacyjnych zniechęceni uczniowie z grupy kontrolnej mogą przestać się uczyć, zacząć się źle zachowywać, stać się agresywni.

Takie zatem są niektóre źródła nietrafności wewnętrznej eksperymentu. Mając ich świadomość, eksperymentatorzy opracowali sposoby radzenia sobie z nimi. Klasyczny schemat eksperymentalny, z prawidłowo przeprowadzonym doborem badanych i ich przydzielaniem, radzi sobie z każdym z powyższych problemów. Spójrzmy raz jeszcze na schemat badań przedstawiony na rysunku 8.4.

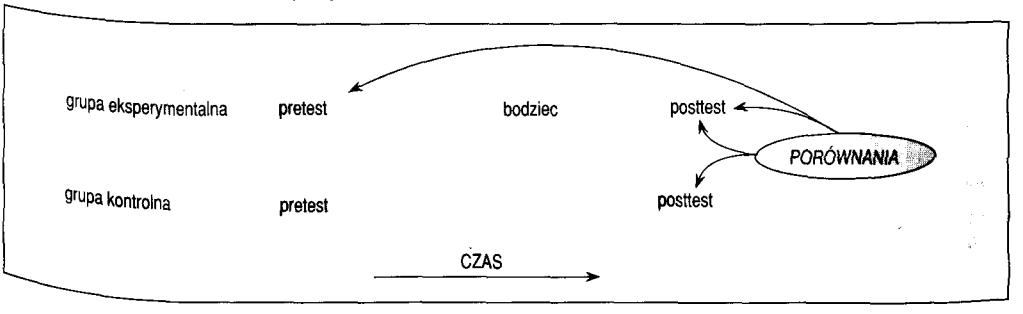
Gdybyśmy zastosowali ten schemat do badań dotyczących uprzedzeń wobec Afroamerykanów, należałoby oczekiwać dwóch odkryć. W grupie eksperymentalnej poziom uprzedzeń wykazany w postteście powinien być niższy od wykazanego w preteście. Ponadto porównanie wyników obydwu posttestów powinno wykazać niższy poziom uprzedzeń w grupie eksperymentalnej niż w kontrolnej.

Schemat ten również dobrze sobie radzi z problemem historii, ponieważ wszystko co dzieje się poza eksperymentem, a co może oddziaływać na grupę eksperymentalną, powinno oddziaływać również na grupę kontrolną. Powinna zatem nadal istnieć różnica w wynikach obydwu testów. Rozwiązany jest również problem dojrzewania, pod warunkiem że badani byli losowo przydzieleni do obydwu grup. Pomiar oraz narzędzia także nie są problemem, ponieważ osoby zarówno z grupy eksperymentalnej, jak i z kontrolnej są poddawane tym samym testom i temu samemu wpływowi eksperymentatora. Jeśli badani zostali przydzieleni losowo, regresja powinna dotyczyć tak samo obu grup, nawet jeśli badaniu poddawane są osoby o skrajnym poziomie uprzedzeń (bądź jakiejś innej zmiennej zależnej). Obciążeniom doboru zapobiega losowy przydział badanych. Trudniejszym do rozwiązania problemem jest wymieranie, ale schemat ten przynosi kilka sposobów wyjścia z takiej sytuacji. Subtelne zmiany w projekcie – na przykład oferowanie grupie kontrolnej *placebo* (powiedzmy filmu, który nie ma nic wspólnego z Afroamerykanami) – ułatwiają rozwiązanie problemu.

Pozostałych pięciu problemów nietrafności wewnętrznej można uniknąć poprzez staranne realizowanie kontrolowanego schematu eksperymentalnego. Omawiany schemat ułatwia jasne określenie zmiennych zależnych i nieza-

Rysunek 8.4

Ponowne spojrzenie na schemat klasyczny



leżnych. Osoby z grupy eksperymentalnej i kontrolnej można izolować od siebie, zmniejszając w ten sposób ryzyko rozpowszechniania bądź naśladownictwa. Ścisła kontrola pozwala zapobiec rekompensatom dawanym grupie kontrolnej, a rywalizacja może być obserwowana i wzięta pod uwagę przy ocenianiu wyników eksperymentu, podobnie jak problem zniechęcenia.

Źródła nietrafności zewnętrznej

Nietrafność wewnętrzna wyjaśnia tylko niektóre trudności, na jakie natyka się eksperymentator. Innym problemem jest, nazwana tak przez Campbella i Stanleya, **nietrafność zewnętrzna**, która odnosi się do możliwości uogólniania wyników eksperymentu na rzeczywistość pozaeksperymentalną. Nawet jeśli wyniki eksperymentu faktycznie odpowiadają temu, co wydarzyło się podczas badania, to czy mówią nam cokolwiek o życiu w społeczeństwie?

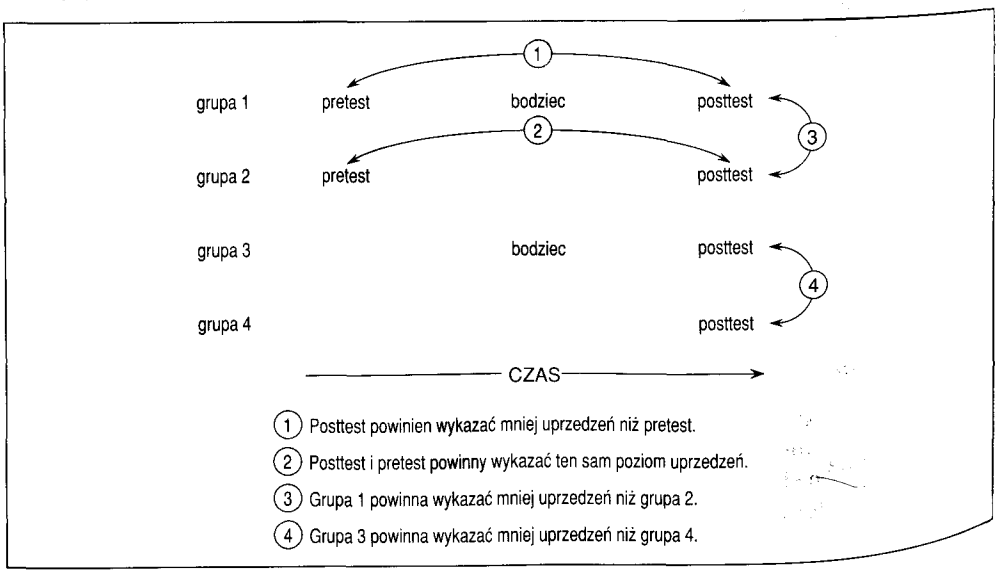
Campbell i Stanley opisali cztery rodzaje tego problemu; skoncentrujemy się na jednym z nich jako przykładzie. Jak wskazują autorzy (1963, s. 18), możliwość uogólniania wyników jest zagrożona, jeśli zachodzi interakcja między testowaną sytuacją a bodźcem eksperymentalnym.

Zostańmy przy przykładzie badań dotyczących uprzedzeń i filmu o historii Afroamerykanów. Wyobraźmy sobie, że grupa eksperymentalna – w schemacie klasycznym – była mniej uprzedzona w postteście niż w preteście i że jej posttest wykazuje mniejsze uprzedzenia niż posttest grupy kontrolnej. Możemy być pewni, że film faktycznie wpłynął na obniżenie poziomu uprzedzeń. Ale czy film miałby taki sam wpływ, gdybyśmy pokazali go w kinach lub telewizji? Nie możemy być tego pewni, gdyż film mógł być skuteczny tylko dlatego, że ludzie zostali uwrażliwieni na kwestię uprzedzeń, na przykład poprzez uczestniczenie w preteście. To właśnie przykład interakcji testowania i bodźca. Klasyczny schemat eksperymentalny nie daje kontroli nad taką sytuacją. Na szczęście badacze eksperymentalni opracowali inne schematy, które sobie z nią radzą.

Czterogrupowy schemat Solomona (Campbell, Stanley 1963, s. 24–25) zajmuje się problemem interakcji testowania z bodźcem. Jak sama nazwa wskazuje, wymaga on czterech grup, do których losowo przydzielono badanych. Schemat ten przedstawiony jest na rysunku 8.5.

Rysunek 8.5

Czterogrupowy schemat Solomona



Zauważmy, że grupy 1 i 2 na rysunku 8.5 składają się na eksperyment klasyczny, w którym grupa 2 jest grupą kontrolną. Grupę 3 poddaje się wpływowi bodźca bez pretestu, w grupie 4 przeprowadza się tylko posttest. Ten schemat pozwala na cztery znaczące porównania, które opisane zostały na rysunku. Jeśli film o historii Afroamerykanów faktycznie powoduje spadek poziomu uprzedzeń – nie wyjaśniony problemem trafności wewnętrznej i nie wyjaśniony interakcją między pomiarem a bodźcem – należy oczekiwać czterech wniosków:

1. W grupie 1 posttest powinien wykazać niższy poziom uprzedzeń niż pretest.
2. W postteście grupa 1 powinna ujawnić mniej uprzedzeń niż grupa 2.
3. Grupa 3 w postteście powinna ujawnić mniej uprzedzeń niż grupa 2 w preteście.
4. Posttest powinien wykazać niższy poziom uprzedzeń w grupie 3 niż w grupie 4.

Zauważmy, że wnioski (3) i (4) wykluczają jakąkolwiek interakcję między przeprowadzeniem pomiaru a bodźcem. Trzeba też pamiętać, że porównania będą miały sens tylko wtedy, gdy badani zostaną przydzieleni do różnych grup losowo, dzięki czemu zapewnimy taki sam początkowy poziom uprzedzeń w poszczególnych grupach, nawet jeśli w preteście uprzedzenia mierzone są tylko w grupach 1 i 2.

Autorzy wskazują na istnienie pewnej ubocznej korzyści z zastosowania tego schematu badań. Schemat czterech grup Solomona nie tylko wyklucza interakcje między pomiarem a bodźcem, lecz również dostarcza danych dla porównań, które ujawniają wielkość takiej interakcji w schemacie klasycznym. Wiedza ta pozwala badaczowi oszacować wartość wcześniejszych badań, z zastosowaniem prostszych schematów.

Ostatni ze schematów eksperymentalnych, który zostanie tu przedstawiony, Campbell i Stanley (1963, s. 25–26) nazwali *schematem bez pretestu z grupą kontrolną*; składa się on z drugiej części – grup 3 i 4 – schematu Solomona. Autorzy przekonująco argumentują, że – z odpowiednią randomizacją – tylko grupy 3 i 4 wystarczą do eksperymentu, który kontroluje trafność wewnętrzną oraz interakcję pomiaru z bodźcem. Stosując losowe przyporządkowanie do grup eksperymentalnej i kontrolnej (co odróżnia ten schemat od omawianego wcześniej statycznego porównania międzygrupowego), otrzymujemy grupy od początku porównywalne ze względu na zmienną zależną – porównywalne w stopniu wystarczającym do zastosowania zwyczajowych testów statystycznych, używanych do oceny wyników – nie ma więc potrzeby przeprowadzania pretestu. Campbell i Stanley twierdzą, że jedynym uzasadnieniem pretestu jest w tej sytuacji tradycja. Eksperymentatorzy są po prostu przyzwyczajeni do pretestów i czują się bezpieczniej przy schematach badawczych, które je zawierają. Powyższe uwagi odnoszą się jednak tylko i wyłącznie do eksperymentów, w których badani zostali przyporządkowani do grupy eksperymentalnej i kontrolnej w sposób losowy, ponieważ tylko to usprawiedliwia założenie, że grupy są równoważne bez faktycznego mierzenia, czy jest tak naprawdę.

Powyższe rozważania wprowadzają w zawiłości schematów eksperymentalnych, ich problemy i niektóre z rozwiązań. Oczywiście stosuje się znacznie więcej schematów eksperymentalnych. Niektóre wymagają więcej niż jednego bodźca bądź zawierają kombinację bodźców. Inne charakteryzuje kilkakrotny pomiar zmiennej zależnej w czasie oraz oddziaływanie bodźcem w różnych momentach na różne grupy. Jeśli interesuje cię ta problematyka, możesz zajrzeć do książki Campbella i Stanleya.

Przykłady eksperymentów

Eksperymenty stosowane są w naukach społecznych do badania bardzo zróżnicowanych zjawisk. Część eksperymentów przeprowadza się w warunkach laboratoryjnych; inne realizowane są w „realnym świecie". Poniższe omówienie pozwala przyjrzeć się obu typom.

Zacznijmy od przykładu z rzeczywistości. W cenionej sztuce George'a Bernarda Shawa *Pigmalion*, która stała się podstawą musicalu *My Fair Lady*, wielokrotnie wystawianego na

Broadwayu, Eliza Doolittle mówi o wpływie innych ludzi na określanie naszej społecznej tożsamości. Oto jak odróżnia ona sposób, w jaki traktowana jest przez swego opiekuna, profesora Higginsa, od sposobu, w jaki traktuje ją przyjaciel Higginsa, pułkownik Pickering:

> „Widzi pan, mówiąc szczerze, poza tym wszystkim, czego każdy może się nauczyć – jak się ubierać, poprawnie mówić itd. – różnica między damą a kwiaciarką nie polega na tym, jak się zachowują, tylko, jak są traktowane. Dla profesora Higginsa pozostanę kwiaciarką, bo tak mnie traktuje i zawsze będzie traktował. Ale wiem, że dla pana mogę być damą, bo pan mnie traktował jak damę i zawsze będzie się tak do mnie odnosił[1]".
>
> (1981 s. 152)

Odczucia, którym Eliza daje tutaj wyraz, to jedna z podstaw nauk społecznych, bardziej precyzyjnie wyrażona przez takich socjologów, jak Charles Horton Cooley („jaźń odzwierciedlona") i George Herbert Mead („uogólniony inny"). Ogólnie rzecz biorąc chodzi o to, że sposób, w jaki myślimy o sobie – nasze wyobrażenie samych siebie – oraz sposób, w jaki się zachowujemy, jest w znacznej mierze funkcją tego, jak widzą nas inni i w jaki sposób się do nas odnoszą. Z kolei sposób, w jaki inni nas postrzegają, jest warunkowany ich, z góry przyjmowanymi, oczekiwaniami. Jeśli powiedziano komuś o nas, że jesteśmy głupi, osoba ta będzie nas prawdopodobnie w ten sposób widzieć – my sami możemy się tak zacząć postrzegać, a nawet naprawdę zachowywać się głupio.

Prawidłowość ta nazywana często *syndromem Pigmaliona* dobrze pasuje do kontrolowanych eksperymentów. W jednym z najbardziej znanych eksperymentów dotyczących *syndromu Pigmaliona* Robert Rosenthal i Lenore Jacobson (1968) przeprowadzili na uczniach szkoły w West Coast badania, które określali jako „harvardzki test nabywania umiejętności fleksyjnych". Wyniki testu pokazali następnie nauczycielom, wskazując jednocześnie niektórych uczniów jako tych, którzy na podstawie wyniku testu powinni w nadchodzącym roku wykazać nagły wzrost zdolności naukowych.

Wyniki przeprowadzonego później testu IQ pokazały, że badacze poczynili słuszne przewidywania. Uczniowie, wskazani jako szczególnie szybko rozwijający się, po upływie roku znacznie wyprzedzili swoich klasowych kolegów, co wskazywałoby, że test był bardzo trafny. W rzeczywistości jednak test był żartem! Swoje przewidywania badacze poczynili losowo, zarówno w odniesieniu do uczniów dobrych, jak i złych. To, co powiedzieli nauczycielom, w ogóle nie znajdowało odbicia w wynikach testu. Postęp poczyniony przez „rozwijających się" uczniów był po prostu wynikiem tego, że nauczyciele oczekiwali poprawy u tych właśnie uczniów, poświęcali im więcej uwagi, zachęcali ich do pracy i nagradzali za postępy (zwróć uwagę na podobieństwo między tą sytuacją a efektem Hawthorne omówionym w tym rozdziale wcześniej).

Badania Rosenthala i Jacobson przyciągnęły wielką uwagę społeczeństwa i naukowców. Późniejsze badania skoncentrowały się na poszczególnych aspektach tego, co stało się znane jako *proces atrybucji* lub *komunikacyjny model oczekiwań*. Badania te, przeprowadzane głównie przez psychologów, podobne są do badań, jakie realizują socjologowie, których zainteresowania są nieco odmienne, często określane zbiorczą nazwą *teorii oczekiwań*. Psychologowie koncentrują się na sytuacjach, w których oczekiwania jednostki dominującej wpływają na zachowania osób podległych – tak jak w przypadku nauczyciela i uczniów, szefa i pracowników. Badania socjologiczne bardziej interesują się rolą wzajemnych oczekiwań równych sobie członków małych grup, mających do zrealizowania jakieś zadanie. W przypadku jury na przykład, w jaki sposób jurorzy oceniają siebie nawzajem na początku, i jak te początkowe oceny wpływają na ich późniejsze interakcje?

Oto przykład eksperymentu, którego celem było zbadanie sposobu, w jaki nasze postrzeganie własnych umiejętności oraz umiejętności innych osób wpływa na gotowość zaakceptowania pomysłów innych osób. Martha

[1] George Bernard Shaw, *Pigmalion*, tłum. Kazimierz Piotrowski, PIW, Warszawa 1981 (przyp. tłum.).

Foschi, G. Keith Warriner i Stephen Hart (1985) byli szczególnie zainteresowani rolą „standardów" w tym względzie:

> „Ogólnie, przez «standardy» rozumiemy to, na ile dobrze bądź źle dana osoba musi wypaść, żeby jakieś umiejętności były jej przypisane bądź by ich jej odmówiono. Standardy są naszym zdaniem kluczową zmienną wpływającą na to, jak zachodzi ocenianie i jakimi się kończy oczekiwaniami. Na przykład w zależności od tego, jakie standardy zostaną zastosowane, ten sam poziom sukcesu może zostać zinterpretowany jako wielkie osiągnięcie bądź zostać odrzucony jako nieistotny".
>
> (1985, s. 108–109)

Badając rolę standardów, badacze skonstruowali eksperyment z czterema grupami eksperymentalnymi i jedną kontrolną. Uczestnikom powiedziano, że eksperyment dotyczy „umiejętności rozpoznania wzorców", którą część ludzi posiada jako wrodzoną zdolność, część zaś umiejętności tej nie ma. Badacze powiedzieli uczestnikom, że będą pracować w parach nad problemami rozpoznawania wzorców.

W rzeczywistości nie ma oczywiście niczego takiego jak umiejętność rozpoznawania wzorców. Przedmiotem eksperymentu było określenie, na ile informacja o tej rzekomej umiejętności wpłynęła na późniejsze zachowanie uczestników.

Pierwszym krokiem eksperymentu było „zbadanie" zdolności rozpoznawania wzorców u każdej osoby. Uczestnikom eksperymentu pokazywano przez 8 sekund wzorzec geometryczny, następnie kolejne dwa, podobne do wcześniejszego, ale nie identyczne z nim. Zadaniem było określenie, który z wzorców pokazanych później jest najbardziej podobny do pierwszego wzorca. Pytano o to 20-krotnie, a komputer pokazywał „wynik". Połowie osób powiedziano, że trafiły 14 razy; drugiej połowie powiedziano, że udało się im trafić tylko 6 razy – niezależnie od tego, który wzór osoby te połączyły z którym. W zależności od wyniku uczestnicy myśleli, że poszło im raczej dobrze lub raczej źle. Zauważmy, że nie istniał standard pozwalający na ocenę wyniku – być może zdobycie czterech prawidłowych odpowiedzi można by uznać za doskonały wynik.

W tym samym momencie, w którym uczestnik otrzymywał swój wynik, otrzymywał też „wynik partnera", chociaż zarówno sami partnerzy, jak i ich wyniki byli jedynie komputerową fikcją (badanym zapowiedziano, że będą porozumiewać się z partnerami przez terminal komputera, ale nie będą mogli siebie widzieć). Osobie, która zdobyła 14 punktów, mówiono, że jej partner zdobył 6; temu kto zdobył 6, mówiono, że wynik jego partnera to 14.

Taka procedura oznaczała, że uczestnicy wchodzili w drużynową fazę eksperymentu przekonani, że (1) poszło im lepiej niż partnerowi lub (2) poszło im gorzej niż partnerowi. Ta informacja była częścią „standardu", którym uczestnicy posługiwali się w czasie eksperymentu. Dodatkowo, połowie każdej grupy powiedziano, że wynik między 12 a 20 oznacza, że dana osoba *niewątpliwie* ma umiejętność rozpoznawania wzorców; pozostałym uczestnikom powiedziano, że 14 to wynik, który nie jest wystarczająco wysoki, żeby można było cokolwiek definitywnie udowodnić. Oznacza to, że w końcu każdy uczestnik mógł być przekonany, że:

1. Jest *zdecydowanie lepszy* w rozpoznawaniu wzorców niż jego partner.
2. Jest *prawdopodobnie lepszy* od partnera.
3. Jest *prawdopodobnie gorszy* od partnera.
4. Jest *zdecydowanie gorszy* od partnera.

Członkom grupy kontrolnej w tym eksperymencie nie mówiono nic ani o ich własnych umiejętnościach, ani o umiejętnościach partnerów. Innymi słowy, nie mieli oni żadnych oczekiwań.

Końcowym krokiem eksperymentu było przydzielenie każdej „drużynie" zadania. Tak jak poprzednio, każdej parze dano początkowy wzorzec, a potem dwa kolejne do porównania i wyboru. Tym razem jednak, gdy uczestnik podawał swój wybór, mówiono mu, jakiego wyboru dokonał partner, i proszono o ponowny wybór. Przy wyborze końcowym można było pozostać przy wcześniejszym albo zmienić go. Wybór „partnera" był oczywiście przygotowany przez komputer i – jak można się domyślić

– często zdarzały się w drużynach odmienne zdania: w sumie 16 na 20 przypadków.

Zmienną zależną w tym eksperymencie była częstość, z jaką uczestnicy będą zmieniać swój wybór tak, aby pasował do wyboru partnera. Hipotezą badaczy było, że członkowie grupy *zdecydowanie lepszej* będą najrzadziej zmieniać wybór. Nieco częściej zmieniać go będą badani z grupy *prawdopodobnie lepsi*, jeszcze częściej z *grupy kontrolnej*, jeszcze częściej z grupy *prawdopodobnie gorsi*, a najczęściej zmiany dokonywać będą badani z grupy *zdecydowanie gorsi*.

Poniżej przedstawiono, jak często uczestnicy zmieniali wybór odpowiedzi w każdej z pięciu grup. Zauważ, że każdy miał 16 okazji, by tego dokonać. Dane dowodzą, że wszystkie przewidywania badaczy sprawdziły się, z wyjątkiem porównania między grupą *prawdopodobnie gorsi* i *zdecydowanie gorsi*. Choć ostatnia grupa faktycznie najczęściej zmieniała odpowiedź, różnica była zbyt mała, by można ją uznać za potwierdzenie hipotezy (w rozdziale 17 omówione będą testy statystyczne, które pozwoliły badaczom podjąć taką decyzję).

Grupa	Średnia liczba zmian
zdecydowanie lepsi	5,05
prawdopodobnie lepsi	6,23
grupa kontrolna	7,95
prawdopodobnie gorsi	9,23
zdecydowanie gorsi	9,28

Po bardziej szczegółowych analizach okazało się, że ta sama podstawowa prawidłowość dotyczy zarówno mężczyzn, jak i kobiet, choć była ona nieco wyraźniejsza w przypadku kobiet. Oto konkretne dane:

	Średnia liczba zmian	
	kobiety	mężczyźni
zdecydowanie lepsi	4,50	5,66
prawdopodobnie lepsi	6,34	6,10
grupa kontrolna	7,68	8,34
prawdopodobnie gorsi	9,36	9,09
zdecydowanie gorsi	10,00	8,70

Ponieważ niekiedy konkretne wysiłki badawcze koncentrują się, jak się wydaje, wyłącznie na swojej własnej tematyce, można zastanawiać się nad ich odniesieniem do czegokolwiek. Jednak, jako część większych badań, również takie badania jak omawiane wnoszą konkretne wyniki do naszego rozumienia ogólniejszych procesów społecznych.

Warto zastanowić się przez chwilę nad niektórymi z życiowych sytuacji, kiedy „oczekiwania" mogą mieć realne i ważne konsekwencje. Wspominałem już o przypadku obrad jury. A co z wszelkimi przejawami uprzedzeń i dyskryminacji? Można też rozważyć sytuację, kiedy oczekiwania pojawiają się przy rozmowach kwalifikacyjnych lub przed spotkaniem z rodzicami ukochanej osoby. Kiedy o tym pomyślisz, przyjdą ci niewątpliwie do głowy inne sytuacje, w których te laboratoryjne pojęcia znajdują zastosowanie w rzeczywistości.

Eksperymenty „naturalne"

Mimo tendencji do zrównywania pojęć *eksperymentu* i *eksperymentu laboratoryjnego* wiele ważnych eksperymentów społecznych odbywa się poza otoczeniem kontrolowanym, często w trakcie normalnych wydarzeń społecznych. Czasem natura projektuje i przeprowadza eksperymenty, które możemy obserwować i analizować; niejednokrotnie tę naturalną funkcję spełniają społeczni i polityczni decydenci.

Wyobraźmy sobie na przykład, że huragan zniszczył jakąś miejscowość. Część mieszkańców doświadczyła poważnych strat materialnych, innym udało się tego szczęśliwie uniknąć. Jakie, zapytajmy, skutki w zachowaniu ludzi wynikają z doświadczenia klęski żywiołowej? Czy ci, którzy bardziej ucierpieli, bardziej będą też starali się zabezpieczyć przed przyszłymi klęskami niż ci, którzy ucierpieli w mniejszym stopniu? Żeby odpowiedzieć na te pytania, możemy, jakiś czas po huraganie, przeprowadzić wywiady z mieszkańcami. Możemy z nimi porozmawiać na temat środków ostrożności, jakie stosowali przed huraganem i jakie stosują teraz, porównując osoby, które znacznie ucierpiały, z tymi, które ucierpiały względnie mało. W ten sposób możemy wykorzystać naturalny

eksperyment, którego nie moglibyśmy sami zaaranżować, choćbyśmy bardzo chcieli.

Podobny przykład znajdujemy w kronikach badań społecznych dotyczących II wojny światowej. Po zakończeniu wojny badacze społeczni przeprowadzili retrospektywne badania sondażowe nad nastrojami w czasie wojny pośród cywili w kilku niemieckich miastach. Chcieli między innymi określić wpływ masowych bombardowań na morale cywili. Porównali wyniki wojennego morale mieszkańców miast, które były bardzo bombardowane, z wynikami otrzymanymi z miast relatywnie mało bombardowanych (bombardowanie nie pogarsza nastrojów wśród cywili).

Ponieważ badacz musi brać rzeczy tak, jak się one zdarzają, w przypadku naturalnych eksperymentów pojawia się wiele omówionych wcześniej problemów związanych z trafnością. Dlatego też, kiedy Stanislav Kasl, Rupert Chisolm i Brenda Eskenazi (1981) postanowili zbadać, jaki wpływ na robotników miała katastrofa nuklearna w Pensylwanii (Three Mile Island), musieli być szczególnie uważni przy projektowaniu badań:

> „Badania katastrof muszą być siłą rzeczy okazjonalne, quasi-eksperymentalne i przeprowadzane post factum. Według terminologii klasycznych analiz schematów badawczych Campbella i Stanleya nasze badania zaliczają się do kategorii «statycznych porównań międzygrupowych», uważanej za jeden ze słabszych schematów badawczych. Słabości są jednak potencjalne, a o ich faktycznej obecności decydują niepowtarzalne okoliczności każdych badań".
>
> (1981, s. 474)

Podstawą tych badań był sondaż przeprowadzony na osobach, które pracowały w Three Mile Island (TMI) 28 marca 1978 roku, kiedy zepsuł się system chłodniczy w reaktorze numer 2 i zaczął się topić uranowy rdzeń reaktora. Sondaż zrealizowano pięć, sześć miesięcy po wypadku. Kwestionariusz mierzył między innymi postawy robotników wobec pracy w elektrowni jądrowej. Gdyby zbadano postawy tylko pracowników TMI po wypadku, badacze nie wiedzieliby, czy postawy zmieniły się w wyniku wypadku. Schemat badawczy został jednak ulepszony dzięki wybraniu innej – bliskiej i prawdopodobnie porównywalnej – elektrowni jądrowej (w skrócie PB), której pracowników potraktowano jako grupę kontrolną: stąd odniesienie do statycznego porównania międzygrupowego.

Nawet mając grupę eksperymentalną i kontrolną badacze byli ostrożni wobec potencjalnych problemów związanych z ich schematem. W szczególności oparty był on na pomyśle, że obie grupy robotników są sobie równoważne z wyjątkiem katastrofy. Badacze mogliby dokonać takiego założenia, gdyby możliwe było przydzielenie pracowników do obydwu elektrowni losowo, co w rzeczywistości było niemożliwe. Zamiast tego musieli porównać cechy obu grup i wywnioskować, czy grupy są równoważne. Ostatecznie badacze stwierdzili, że obie grupy są bardzo podobne, a to, w której elektrowni robotnicy byli zatrudnieni, zależało wyłącznie od miejsca zamieszkania.

Nawet przyjmując, że obie grupy robotników były równoważne, badacze wciąż mieli inny problem związany z porównywalnością. Nie mogli nawiązać kontaktu ze wszystkimi robotnikami zatrudnionymi w TMI w czasie wypadku. Przedstawili ten problem w następujący sposób:

> „W badaniach tych szczególny problem stanowiły ubytki z badanej populacji. Związane to było z możliwością, że wśród potencjalnych respondentów, z którymi nie udało się nawiązać kontaktu, mogli być tacy pracownicy TMI (lecz nie PB), którzy na stałe wyprowadzili się z tej okolicy ze względu na wypadek. Jest wielce prawdopodobne, że takie nieprzypadkowe ubytki mogły wpłynąć na zmniejszenie wartości szacowanego wpływu wypadku na postawy pracowników. Korzystając z rejestru wycofanych lub odłączonych numerów telefonicznych, oceniliśmy, że zakłócenie to było bez znaczenia (1%)".
>
> (Kasl i in. 1981, s. 475)

Przykład TMI wskazuje zarówno na szczególne problemy napotykane w naturalnych eksperymentach, jak i na możliwości uwzględnienia tych problemów. Badania społeczne generalnie wymagają pomysłowości i zrozumienia; eksperymenty naturalne wymagają ich nieco więcej niż inne rodzaje badań.

We wcześniejszej części tego rozdziału omawiany był hipotetyczny przykład badań dotyczących wpływu filmu o historii Afroamerykanów na zmniejszenie poziomu uprzedzeń. Sandra Ball-Rokeach, Joel Grube i Milton Rokeach (1981) mieli okazję zajmować się tym problemem w rzeczywistości poprzez eksperyment naturalny. W 1977 roku telewizyjna adaptacja *Korzeni* Alexa Haleya, historycznej sagi o Afroamerykanach, była pokazywana przez telewizję ABC przez osiem kolejnych wieczorów. Zebrała największą publiczność w dotychczasowej historii telewizji. Ball-Rokeach ze współpracownikami postanowiła sprawdzić, czy *Korzenie* zmieniły postawy białych Amerykanów wobec Afroamerykanów. Możliwość pojawiła się w 1979 roku, kiedy w telewizji została pokazana druga część serialu – *Korzenie: następne pokolenie*. Choć miło byłoby (z punktu widzenia badacza) wylosować próbę Amerykanów oglądających program i nie oglądających go, było to niemożliwe. Zamiast tego badacze dobrali cztery próby ze stanu Waszyngton i wysłali pocztą kwestionariusze, mające mierzyć postawy wobec Afroamerykanów. Po wyemitowaniu ostatniego odcinka, zadzwoniono do respondentów z pytaniem, ile odcinków obejrzeli. Później wysłano respondentom kwestionariusze, które ponownie mierzyły ich postawy wobec Afroamerykanów.

Porównując postawy przed i po wśród tych, którzy oglądali serial, i tych, którzy go nie widzieli, badacze doszli do kilku wniosków. Odkryli na przykład, że osoby od początku bardziej egalitarne znacznie częściej oglądały serial niż osoby bardziej uprzedzone: zjawisko autoselekcji. Co więcej, porównanie uprzedzeń przed i po wśród tych, którzy oglądali serial, wykazało, że wpływ samego filmu był mały albo żaden. Ci, którzy go oglądali, nie byli bardziej egalitarni po obejrzeniu filmu niż przed obejrzeniem go.

Ten przykład wyprzedza opisane w rozdziale 12 badania ewaluacyjne, które można postrzegać jako szczególną odmianę eksperymentu naturalnego. Jak zobaczymy, badania ewaluacyjne wykorzystują logikę eksperymentu do obserwacji i ewaluacji wpływu bodźców w realnym życiu. Ponieważ jest to rodzaj badań społecznych nabierający coraz większego znaczenia, poświęcony mu będzie cały rozdział.

Mocne i słabe strony metody eksperymentalnej

Eksperymenty są podstawowym narzędziem badania związków przyczynowych. Jak każda metoda badawcza, mają jednak zarówno słabe, jak i mocne strony.

Główną zaletą kontrolowanego eksperymentu jest izolacja zmiennej eksperymentalnej i jej wpływu w czasie. Jest to najbardziej widoczne w podstawowym modelu eksperymentalnym. Wybiera się grupę eksperymentalną, która na początku badań ma określoną cechę; po oddziałaniu na nią bodźcem eksperymentalnym okazuje się, że ma ona inną cechę. W zależności od tego, czy uczestnicy nie byli poddani wpływowi innych bodźców, możemy wywnioskować, że zmiana cech spowodowana była wpływem bodźca eksperymentalnego.

Pojedyncze eksperymenty mają dość ograniczony zakres, ale ponieważ wymagają relatywnie niskich nakładów czasowych i finansowych oraz małej liczby uczestników, często możemy powtarzać dany eksperyment kilkakrotnie na różnych grupach ludzi (nie zawsze jest to zasadą, jednak zazwyczaj łatwiej jest powtórzyć eksperyment niż na przykład sondaż). Jak we wszystkich rodzajach badań naukowych, powtarzalność odkryć badawczych potwierdza nasze wnioski dotyczące trafności i możliwości uogólniania tych odkryć.

Największą słabością laboratoryjnych eksperymentów jest ich sztuczność, prowadząca do artefaktów[2]. Procesy społeczne zachodzące w warunkach laboratoryjnych niekoniecznie pojawiają się w bardziej naturalnym społecznym kontekście. Na przykład film o historii Afroamerykanów mógł naprawdę zmniejszyć uprzedzenia wśród członków grupy eksperymentalnej. Nie musi to jednak oznaczać, że ten sam film wyświetlony w kinach całego kraju zredukowałby uprzedzenia panujące pośród ogółu społeczeństwa. Sztuczność i wynikające z niej artefakty są oczywiście problemem raczej dla eksperymentów laboratoryjnych niż dla eksperymentów naturalnych.

[2] Artefakty to zachowania, będące wynikiem samego faktu badania lub też konkretnie zastosowanej metody (przyp. red. nauk.).

Omawiając kilka źródeł nietrafności wewnętrznej i zewnętrznej, przedstawionych przez Campbella, Stanleya i Cooka, przekonaliśmy się, że można skonstruować schemat eksperymentalny, który umożliwia logiczną kontrolę takich sytuacji. Ta możliwość wskazuje na jedną z największych zalet eksperymentów: poddają się one logicznemu rygorowi, co często jest znacznie trudniej osiągnąć w przypadku innych typów obserwacji.

Główne zagadnienia

- Eksperymenty doskonale sprawdzają się w kontrolowanych badaniach procesów przyczynowych.
- Eksperyment klasyczny mierzy wpływ bodźca eksperymentalnego (zmiennej niezależnej) na zmienną zależną poprzez pretest i posttest przeprowadzane w grupach eksperymentalnej i kontrolnej.
- Ważniejsze jest, żeby grupy eksperymentalna i kontrolna były podobne do siebie, niż żeby grupa eksperymentalna była reprezentatywna dla jakiejś większej populacji.
- Eksperyment podwójnie ślepy zabezpiecza przed wpływem eksperymentatora, ponieważ ani eksperymentator, ani uczestnicy eksperymentu nie wiedzą, które osoby są w grupie eksperymentalnej, a które w kontrolnej.
- Dobór losowy, randomizacja i dopasowywanie są metodami uzyskiwania porównywalności grup eksperymentalnej i kontrolnej; preferowaną metodą jest na ogół randomizacja, którą w niektórych schematach można połączyć z dopasowywaniem.
- Campbell i Stanley opisali trzy typy preeksperymentów: studium przypadku z pojedynczym pomiarem, schemat z dwukrotnym pomiarem w jednej grupie, statyczne porównanie międzygrupowe. W żadnym z tych typów nie występują wszystkie metody kontroli dostępne w prawdziwym eksperymencie.
- Campbell i Stanley wyróżnili między innymi 12 źródeł nietrafności wewnętrznej w schematach eksperymentalnych. Klasyczny schemat eksperymentalny z losowo przydzielonymi uczestnikami radzi sobie z każdym z tych 12 źródeł.
- Eksperymentów dotyczy również problem nietrafności zewnętrznej: odkrycia eksperymentalne mogą nie odzwierciedlać rzeczywistości.
- Przykładem nietrafności zewnętrznej, z którą eksperyment klasyczny nie daje sobie rady, jest interakcja między pomiarem a bodźcem.
- Przed nietrafnością zewnętrzną może zabezpieczyć czterogrupowy schemat Solomona oraz inne odmiany schematu klasycznego.
- Campbell i Stanley dowodzili, że przy starannej randomizacji przydziału uczestników do grupy eksperymentalnej i kontrolnej nie ma potrzeby przeprowadzania pretestu.
- W rzeczywistym życiu społecznym często zachodzą naturalne eksperymenty, które badacze mogą wykorzystać w sposób podobny do tego, w jaki zaprojektowaliby i przeprowadzili eksperyment laboratoryjny.
- Jak wszystkie metody badawcze, eksperyment ma swoje mocne i słabe strony. Jego podstawową wadą jest prowadząca do artefaktów sztuczność: to co ma miejsce w eksperymencie, nie musi odzwierciedlać tego, co zachodzi w świecie zewnętrznym. Do zalet eksperymentu można zaliczyć: izolację zmiennej niezależnej, co pozwala na wnioskowanie przyczynowe; względną łatwość powtarzania badań; naukowy rygor.

Kluczowe pojęcia

pretest
randomizacja
posttest
grupa eksperymentalna
grupa kontrolna
eksperyment podwójnie ślepy
dopasowywanie
nietrafność wewnętrzna
nietrafność zewnętrzna

Pytania i ćwiczenia

1. Znajdź w bibliotece lub w Internecie przykład raportu z badań eksperymentalnych. Określ zmienną zależną oraz bodziec.

2. Wybierz 6 spośród 12 źródeł nietrafności wewnętrznej omówionych w tym rozdziale i podaj przykłady (nie omawiane wyżej) dla każdego z nich.
3. Przygotuj hipotetyczny schemat eksperymentalny, który byłby przykładem jednego z problemów nietrafności zewnętrznej.
4. Pomyśl o niedawnej klęsce żywiołowej, której byłeś świadkiem, lub o której czytałeś. Przygotuj pytania badawcze, na które można odpowiedzieć, traktując tę katastrofę jako eksperyment naturalny. Przedstaw projekt badań w dwóch, trzech akapitach.
5. W powyższym rozdziale omówiono krótko problem „efektu placebo". Znajdź w sieci przykład eksperymentu, w którym efekt placebo odegrał ważną rolę. Napisz krótki raport z tych badań, włączając źródło, z którego pochodzą twoje informacje (wskazówka: w polu „szukaj" można wpisać „placebo").

Dodatkowe lektury

Campbell Donald, Stanley Julian, *Experimental and Quasi-Experimental Designs for Research*, Rand McNally, Chicago 1963. Doskonała analiza metod i logiki eksperymentu w naukach społecznych. Książka jest szczególnie wartościowa dzięki zastosowaniu logiki eksperymentu do innych metod badań społecznych. Mimo że dość stara, książka zyskała sobie już status klasycznej i nadal bywa często cytowana.

Cook Thomas D., Campbell Donald T., *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Rand McNally, Chicago 1979. Rozszerzona i unowocześniona wersja Campbella i Stanleya.

Jones Stephen, *Worker Independence and Output: The Hawthorne Studies Reevaluated*, „American Sociological Review" 1990, 55, s. 176–190. Autorzy dokonują przeglądu klasycznych prac i kwestionują tradycyjną interpretację (przedstawioną w tym rozdziale).

Martin David W., *Doing Psychology Experiments*, Brooks/Cole, Monterey 1996. Gruntowne wyjaśnienie logiki stojącej za metodami badawczymi, często w stylu humorystycznym. Książka uwyraźnia idee szczególnie istotne dla początkujących badaczy, na przykład poszukiwanie tematu eksperymentu czy przegląd literatury.

Ray William J., *Methods Toward a Science of Behavior and Experience*, Wadsworth, Belmont 2000. Zwięzły przegląd metod badań społecznych, ze szczególnym naciskiem na eksperyment. Szczególnie mocną stoną książki jest zaplecze z zakresu filozofii nauki.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 9

Badania sondażowe

Przegląd holograficzny

Badacze mają do dyspozycji wiele metod zbierania danych sondażowych – od ankiet pocztowych, poprzez wywiad osobisty, aż po sondaże przeprowadzane za pomocą Internetu. Badacze społeczni muszą wiedzieć, jak dobrać właściwą metodę i jak ją skutecznie zastosować.

Wprowadzenie

Tematy odpowiednie dla badań sondażowych

Wskazówki do zadawania pytań
Wybierz właściwą formę pytania
Spraw, by pytania były jasne
Unikaj podwójnych pytań
Respondenci muszą być w stanie udzielić odpowiedzi
Respondenci muszą chcieć odpowiadać
Pytania powinny być adekwatne
Najlepsze są krótkie pytania
Unikaj form przeczących
Unikaj obciążonych pytań i pojęć

Budowa kwestionariusza
Ogólny format kwestionariusza
Formaty odpowiedzi
Pytania warunkowe
Pytania tabelaryczne
Kolejność pytań w kwestionariuszu
Instrukcje do kwestionariusza
Pilotaż kwestionariusza
Całościowy przykład

Kwestionariusz do samodzielnego wypełniania
Dystrybucja pocztowa i zwroty
Monitorowanie zwrotów
Listy przypominające
Wystarczający poziom zwrotów
Opis przypadku

Kwestionariusz wypełniany przez ankietera
Rola ankietera w badaniach sondażowych
Ogólne zasady pracy ankieterskiej
Koordynacja i kontrola

Sondaże telefoniczne
Wywiad telefoniczny wspomagany komputerowo (CATI)

Nowe technologie a badania sondażowe

Porównanie różnych metod badań sondażowych

Mocne i słabe strony badań sondażowych

Analiza wtórna

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Sondaże są bardzo starą techniką badawczą. Na przykład w Starym Testamencie znajdujemy następujący fragment:

> „Po tej pladze rzekł Pan do Mojżesza i do Eleazara, syna kapłana Aarona: «Dokonajcie obliczenia całej społeczności Izraelitów według rodów, od dwudziestu lat wzwyż (...)»[1]".
>
> (Księga Liczb 26, s. 1–2)

Starożytni władcy Egiptu przeprowadzali spisy, które pomagały im zarządzać ich włościami. Jezus urodził się poza domem, gdyż Józef i Maria byli w podróży do domu przodków Józefa, gdzie wybrali się ze względu na przeprowadzany właśnie rzymski spis ludności.

Mało znaną próbę przeprowadzenia sondażu podjęto wśród francuskich robotników w 1880 roku. Pewien niemiecki socjolog polityki rozesłał pocztą do robotników około 25 tysięcy kwestionariuszy, chcąc zbadać, w jakim stopniu są oni wykorzystywani przez pracodawców. Ten przydługi kwestionariusz zawierał takie pytania, jak na przykład:

> „(...) Czy pracodawca względnie jego urzędnicy stosują wybiegi, żeby Was okpić z części zarobku?
>
> Jeżeli płaca jest od sztuki, czy wykorzystuje się jakość wyrobu jako pretekst do oszukańczych potrąceń z płacy?"[2]

Badaczem przeprowadzającym ten sondaż nie był George Gallup, lecz Karol Marks ([1880] 1956). Mimo że rozesłano 25 tysięcy kwestionariuszy, nie ma żadnej wzmianki o tym, by któryś z nich został zwrócony.

W dzisiejszych czasach sondaż jest często stosowaną techniką obserwacyjną w badaniach społecznych. W typowym badaniu sondażowym badacz dobiera próbę respondentów i rozdaje im standaryzowane kwestionariusze. W rozdziale 7 omówione zostały dokładniej techniki doboru próby. Niniejszy rozdział poświęcony jest zagadnieniom związanym z przygotowaniem kwestionariusza i opisuje różne jego warianty, służące uzyskaniu możliwie adekwatnych odpowiedzi od respondentów.

Rozdział ten kończy krótkie omówienie *analizy wtórnej*, czyli analizy danych sondażowych, zebranych przez kogoś innego. Takie wykorzystanie wyników badań sondażowych stało się w ostatnich latach ważną ich częścią i jest szczególnie przydatne studentom i innym osobom dysponującym niewielkimi funduszami badawczymi.

Zacznijmy od przyjrzenia się, jakie typy tematów nadają się do badań metodą sondażową.

Tematy odpowiednie dla badań sondażowych

Badania sondażowe mogą być zastosowane do celów opisowych, wyjaśniających i eksploracyjnych. Są one stosowane głównie w takich projektach badawczych, w których jednostkami analizy są pojedynczy ludzie. Mimo że metoda ta może być użyta także w przypadku innych jednostek analizy, takich jak grupy czy interakcje, to jednak jakieś pojedyncze osoby muszą posłużyć jako **respondenci** lub informatorzy. Możemy zatem podjąć się przeprowadzenia badań sondażowych, w których jednostką analizy są rozwody, ale kwestionariusze musielibyśmy rozdać osobom biorącym udział w rozwodzie (lub jakimś innym informatorom).

Badania sondażowe są chyba najlepszą dostępną metodą dla tych badaczy, którzy chcą zbierać oryginalne dane w celu opisywania populacji zbyt dużej, by obserwować ją bezpośrednio. Dzięki starannemu doborowi losowemu otrzymujemy grupę respondentów, co do których możemy zakładać, że ich cechy odzwierciedlają cechy szerszej populacji, natomiast starannie skonstruowane, standaryzowane kwestionariusze dostarczają danych w tej samej formie od wszystkich respondentów.

Sondaże są też doskonałym narzędziem pomiaru postaw i poglądów w dużej popula-

---

[1] *Biblia Tysiąclecia*, wyd. 3 popr., Wydawnictwo Pallotinum, Warszawa 1990 (przyp. tłum.).

[2] Karol Marks, *Kwestionariusz dla robotników*, w: Karol Marks, Fryderyk Engels, *Dzieła*, t. 19, Książka i Wiedza, Warszawa 1972, s. 254 (przyp. tłum.).

cji. Sondaże opinii publicznej – na przykład przeprowadzane przez Gallupa, Harrisa, Roperа czy Yankelovicha – to dobrze znane przykłady tego typu zastosowań. Faktycznie badania opinii są tak rozpowszechnione, że czasami społeczeństwo ma wątpliwości, co o nich sądzić. Badacze opinii są krytykowani przez tych, którzy nie sądzą (lub nie chcą uwierzyć), by sondaże mogły być trafne (kandydaci, którzy „przegrywają" w sondażach, często przekonują wyborców, by nie wierzyli sondażom). Lecz sondaże są również krytykowane za to, że są *zbyt* trafne – na przykład gdy używa się badań wyborców, przeprowadzanych w dniu wyborów wśród osób wychodzących z lokali wyborczych (*exit poll*) w celu określenia zwycięzcy jeszcze przed zakończeniem głosowania.

To ogólne nastawienie wobec badań opinii publicznej jest jeszcze bardziej komplikowane przez nierzetelne od strony naukowej „sondaże", które jednak przykuwają uwagę ludzi z powodu tematów, których dotyczą i/lub ich „wyników". Dobrym przykładem mogą tu być „Raporty Hite" o ludzkiej seksualności. Prasa popularna poświęcała Shere Hite znaczną uwagę, jednak autorka ta została doszczętnie skrytykowana przez społeczność badawczą za swoje metody zbierania danych. Dla przykładu, raport Hite z 1987 roku powstał na podstawie kwestionariuszy wypełnianych przez kobiety w całym kraju – ale jakie kobiety? Hite donosiła, że rozdała około 100 000 kwestionariuszy poprzez różne organizacje, a otrzymała z powrotem około 4500.

No tak, 4500 i 100 000 to duże liczby w badaniach sondażowych. A jednak przy zastosowaniu metod badawczych Hite owe 4500 respondentek nie musi wcale reprezentować kobiet w USA lepiej, niż olbrzymia próba zastosowana przez „Literary Digest" w 1936 roku reprezentowała elektorat USA, kiedy to dwa miliony głosów w próbie badawczej wykazały, że Alf Landon pogrąży Roosevelta ogromną przewagą głosów.

Ludzie niekiedy pod przykrywką badań sondażowych usiłują realizować zupełnie inne cele. Mógłbyś na przykład odebrać telefon, w którym zasugerowanoby ci, że zostałeś wybrany do badań sondażowych, a już pierwsze pytanie brzmiałoby: „Czy w ciągu najbliższego tygodnia chciałbyś zarobić we własnym domu milion dolarów?". Albo też powiedzianoby ci, że możesz wygrać nagrodę, jeśli potrafisz powiedzieć, jaki prezydent jest na monecie jednocentowej. (Powiedz, że to Elvis). Niestety, wielu pozbawionych skrupułów telemarketerów próbuje żerować na ludzkiej chęci do współpracy z badaczami przeprowadzającymi sondaże.

Również partie polityczne i organizacje dobroczynne zaczęły przeprowadzać „sondaże" telefoniczne. Często, pod pozorem badania opinii publicznej w związku z jakimś tematem, osoby dzwoniące proszą w końcu respondentów o wsparcie pieniężne.

Wytworem ostatnich kampanii politycznych jest inna forma pozorowanych badań sondażowych, określana jako „sondażowe przepychanie" (*push poll*). Oto, co miało do powiedzenia Amerykańskie Stowarzyszenie Badania Opinii Publicznej (American Association for Public Opinion Polling), krytykując te praktyki:

> „«Sondażowe przepychanie» jest jedną z technik telemarketingu, w której rozmowy telefoniczne są wykorzystywane do zdobywania wyborców poprzez dostarczanie im fałszywych «informacji» o jakimś kandydacie pod pretekstem przeprowadzania badań nad wpływem takich «informacji» na preferencje wyborcze. Prawdziwą intencją nie jest zbadanie opinii publicznej, ale manipulowanie nią – «przepychanie» wyborców od jakiegoś kandydata ku jego przeciwnikowi. Sondaże takie szkalują wybranych kandydatów, rozsiewając na ich temat fałszywe lub mylące informacje. Intencją jest szerzenie propagandy wyborczej pod przykrywką prowadzenia uprawnionych badań opinii publicznej".
>
> (Bednarz 1996)

Krótko mówiąc, określenia „sondaż" czy „badania opinii" są niekiedy nadużywane. Jednak właściwie przeprowadzony sondaż może być przydatnym narzędziem badawczym w naukach społecznych. Zaprojektowanie dobrych (i wiarygodnych) badań sondażowych zaczyna się od sformułowania dobrych pytań. Zajmijmy się teraz tą kwestią.

Wskazówki do zadawania pytań

W badaniach społecznych zmienne często są operacjonalizowane poprzez zadawanie ludziom pytań w celu uzyskania danych do analizy i interpretacji. Czasami pytania te zadaje ankieter, czasami zaś spisuje się je i wręcza respondentom z prośbą o udzielenie odpowiedzi. Badacz może wykorzystać kilka ogólnych wskazówek, które pomogą mu w ułożeniu i zadaniu pytań, służących jako dobre operacjonalizacje zmiennych, a zarazem w uniknięciu pułapek, prowadzących do uzyskiwania niepotrzebnych czy nawet mylących informacji.

Sondaż zakłada użycie **kwestionariusza** – narzędzia zaprojektowanego specjalnie do zbierania informacji, które mogą przydać się do analiz. Logika tkwiąca u podstaw poniższych wskazówek zachowuje swą ważność zawsze, gdy zadajemy ludziom pytania w celu uzyskania danych, chociaż niektóre konkretne wskazówki mogą niekiedy odnosić się bardziej do kwestionariuszy ustrukturyzowanych niż do kwestionariuszy stosowanych w jakościowych wywiadach pogłębionych, które zawierają raczej pytania otwarte.

Wybierz właściwą formę pytania

Zacznijmy od istnienia kilku opcji dostępnych przy układaniu kwestionariusza. Dotyczą one stosowania pytań lub stwierdzeń oraz wyboru pytań otwartych lub zamkniętych.

Pytania i stwierdzenia

Mimo że pojęcie *kwestionariusza* (*questionnaire*) przywodzi na myśl zbiór pytań (*questions*), po zbadaniu typowego kwestionariusza okazałoby się prawdopodobnie, że zawiera on równie dużo stwierdzeń, jak pytań. Dzieje się tak nie bez powodu. Badacz często chce się dowiedzieć, w jakim stopniu respondenci podzielają jakąś konkretną postawę czy punkt widzenia. Jeśli postawę tę da się streścić całkiem krótkim stwierdzeniem, można przedstawić to stwierdzenie respondentom i zapytać, czy się z nim zgadzają, czy nie. Jak być może pamiętasz, Rensis Likert znacznie sformalizował tę procedurę poprzez stworzenie skali Likerta – takiego formatu pytania, w którym respondenci proszeni są o określenie, czy się z czymś zdecydowanie zgadzają, zgadzają, nie zgadzają, czy też zdecydowanie nie zgadzają, albo też zdecydowanie popierają, popierają itd.

Można z pożytkiem stosować zarówno pytania, jak i stwierdzenia. Użycie obu form w danym kwestionariuszu zwiększa elastyczność układania pytań, a także może sprawić, że kwestionariusz będzie ciekawszy.

Pytania otwarte i zamknięte

Jeśli chodzi o zadawanie pytań, badacze mają dwie możliwości. Mogą zadawać **pytania otwarte**, w których prosi się respondenta o sformułowanie własnej odpowiedzi. Można go na przykład zapytać: „Co uważa Pan(i) za najważniejszą obecnie kwestię dla Stanów Zjednoczonych?" i zostawić miejsce na wpisanie odpowiedzi (lub poprosić o ustną odpowiedź na pytanie zadane przez ankietera). Jak przekonamy się w rozdziale 10, pogłębione wywiady jakościowe posługują się niemal wyłącznie pytaniami otwartymi. Pytania takie są jednak również stosowane w badaniach sondażowych.

W przypadku **pytań zamkniętych** prosi się respondenta o wybranie odpowiedzi z listy przedstawionej przez badacza. Pytania zamknięte zawdzięczają swoją wielką popularność w badaniach sondażowych temu, że odpowiedzi na nie cechuje większa jednolitość oraz temu, że są łatwiejsze w opracowaniu.

Odpowiedzi na pytania otwarte trzeba zakodować przed poddaniem ich analizie komputerowej, co zostanie omówione w rozdziale 14. Ów proces kodowania wymaga często, żeby badacz interpretował znaczenie odpowiedzi, co otwiera drogę do nieporozumień i obciążeń związanych z wpływem badacza. Istnieje też niebezpieczeństwo, że niektórzy respondenci będą udzielać odpowiedzi zasadniczo nieadekwatnych w stosunku do intencji badacza. Natomiast odpowiedzi na pytania zamknięte mogą być często przekształcone bezpośrednio na format komputerowy.

Główny problem związany ze stosowaniem pytań zamkniętych polega na ustrukturyzowa-

niu odpowiedzi przez badacza. Jeśli adekwatne odpowiedzi na dane pytanie są względnie jasne, nie ma problemu. Gdy jednak tak nie jest, odpowiedzi ułożone przez badacza mogą pomijać jakieś istotne możliwości odpowiedzi. Na przykład lista odpowiedzi zaproponowanych na pytanie o „najważniejszą obecnie kwestię dla Stanów Zjednoczonych" mogłaby pomijać jakieś kwestie, które respondenci uznaliby za ważne.

Układanie pytań zamkniętych powinno spełniać dwa wymogi strukturalne. Po pierwsze, proponowane kategorie odpowiedzi powinny być wyczerpujące: powinny one zawierać wszystkie możliwe oczekiwane odpowiedzi. Badacze często usiłują spełnić ten warunek przez dodanie takich kategorii jak „Inne (Jakie? .....)". Po drugie, kategorie odpowiedzi muszą być wzajemnie rozłączne: respondent nie powinien czuć się zmuszony do wskazania więcej niż jednej odpowiedzi (w pewnych przypadkach może być pożądane dopuszczenie więcej niż jednej odpowiedzi, ale może to potem stworzyć problemy przy przetwarzaniu danych i ich analizie). Aby przekonać się, czy zaproponowane kategorie są wzajemnie rozłączne, należy przemyśleć starannie wszystkie kombinacje kategorii i zastanowić się, czy ktoś mógłby wskazać więcej niż jedną odpowiedź. Przydatne jest ponadto dodanie instrukcji do pytania, z prośbą do respondenta o wskazanie jednej, najlepszej odpowiedzi, jednak technika ta nie może zastąpić starannie skonstruowanego zestawu odpowiedzi.

Spraw, by pytania były jasne

Powinno być oczywiste, że poszczególne pytania kwestionariusza muszą być jasne i pozbawione dwuznaczności, jednak z powodu szerokiego rozpowszechnienia w sondażach niejasnych i wieloznacznych pytań warto tę kwestię podkreślić. Możemy często tak głęboko zaangażować się w badany temat, że pewne opinie czy punkty widzenia będą jasne dla nas, ale nie dla naszych respondentów, z których wielu poświęcało dotychczas temu tematowi niewiele uwagi albo nie poświęcało jej wcale. Może się też zdarzyć, że w sytuacji, gdy jesteśmy tylko powierzchownie obeznani z danym tematem, nie wyjaśnimy w wystarczający sposób intencji pytania. Pytanie: „Co Pan sądzi o proponowanym planie pokojowym?" może wywołać u respondenta pytanie zwrotne: „Jakim proponowanym planie pokojowym?". Pytania kwestionariusza powinny być precyzyjne, tak aby respondent wiedział dokładnie, o co pyta badacz.

Możliwości nieporozumień są niezliczone i żaden badacz nie jest na nie uodporniony (Polivka, Rothgeb 1993). Jednym z najlepiej ugruntowanych projektów badawczych w Stanach Zjednoczonych jest przeprowadzane przez Biuro Spisowe „Stałe Badanie Ludności" (*Current Population Survey, CPS*). Mierzy ono, poza innymi istotnymi danymi, także krajową stopę bezrobocia. Wchodząca w jego skład część dotycząca wzorców zatrudnienia skupia się na działaniach respondenta w ciągu „ostatniego tygodnia", pod którym to pojęciem Biuro Spisowe rozumie okres od niedzieli do soboty. Badania podjęte w celu sprawdzenia trafności sondażu wykazały, że dla ponad połowy respondentów „ostatni tydzień" dotyczył tylko okresu od poniedziałku do piątku. Podobnie działo się z pojęciem „pracujący w pełnym wymiarze godzin", które było definiowane przez Biuro Spisowe jako 35 godzin w tygodniu lub więcej, a w tych samych badaniach ewaluacyjnych okazało się, że część respondentów stosowała w tym przypadku bardziej tradycyjną definicję 40 godzin tygodniowo. Konsekwencją tego była modyfikacja brzmienia tych pytań w sondażu z 1994 roku w taki sposób, by odpowiadały definicjom Biura Spisowego.

Podobnie używanie terminu „rdzenny Amerykanin" (*Native American*) dla określenia *Amerykanina indiańskiego pochodzenia* prowadzi często do nadreprezentacji tej grupy etnicznej w sondażach. Najwidoczniej wielu respondentów rozumie ten termin jako oznaczający osobę „urodzoną w Stanach Zjednoczonych".

Unikaj podwójnych pytań

Badacze często proszą respondentów o jedną odpowiedź na pytanie, które w rzeczywistości składa się z wielu części. Wydaje się, że najczęściej dzieje się tak wtedy, gdy badacz osobiście identyfikuje się z takim złożonym pytaniem.

można na przykład zapytać respondentów, czy się zgadzają, czy nie zgadzają ze stwierdzeniem: „Stany Zjednoczone powinny zarzucić swój program badań kosmicznych i wydawać pieniądze na inicjatywy dotyczące spraw wewnętrznych". Mimo że wiele osób jednoznacznie zgodzi się z tym stwierdzeniem, a inni jednoznacznie się z nim nie zgodzą, to jeszcze inni nie będą w stanie na nie odpowiedzieć. Niektórzy będą chcieli zaprzestania programu badań kosmicznych i oddania pieniędzy podatnikom. Inni zechcą kontynuacji programu badań kosmicznych, ale jednocześnie chcieliby, by przeznaczano więcej pieniędzy na programy wewnętrzne. Respondenci nie mogą zgodzić się lub nie zgodzić, nie wprowadzając badacza w błąd.

Podwójne pytania i nie tylko

Nawet renomowanym profesjonalnym badaczom zdarzają się niekiedy pytania podwójne i gorsze rzeczy. Przyjrzyjmy się takiemu pytaniu, zadanemu obywatelom USA w kwietniu 1986 roku, w okresie gdy stosunki z Libią były szczególnie trudne. Niektórzy obserwatorzy sugerowali, że Stany Zjednoczone mogą w końcu znaleźć się w stanie prawdziwej wojny z tym małym narodem północnoafrykańskim. Sondaż Harrisa był próbą określenia opinii publicznej w USA na ten temat.

„Jeśli Libia zwiększy teraz liczbę ataków terrorystycznych przeciw USA, a my będziemy wyrządzać dalsze szkody Libii, wówczas w nieunikniony sposób doprowadzi to do rozpoczęcia wojny przez Stany Zjednoczone i najechania tego kraju, co będzie błędem".

Respondenci mieli następujące możliwości odpowiedzi: „zgadzam się", „nie zgadzam się" lub „nie jestem pewien". Zwróć uwagę na elementy zwarte w tym złożonym stwierdzeniu:

1. Czy Libia zwiększy liczbę ataków terrorystycznych przeciw USA?
2. Czy USA będzie wyrządzać dalsze szkody Libii?
3. Czy USA na pewno lub jakoś inaczej rozpocznie wojnę przeciw Libii?
4. Czy USA najedzie na Libię?
5. Czy to będzie słuszne, czy nie?

Tych parę elementów otwiera wiele możliwych punktów widzenia – o wiele więcej niż trzy alternatywy, zaproponowane respondentom w sondażu. Nawet gdybyśmy mieli hipotetycznie założyć, że Libia „zwiększyłaby liczbę ataków terrorystycznych", a Stany Zjednoczone odpłacałyby „wyrządzaniem dalszych szkód", można by oczekiwać co najmniej siedmiu różnych wyników:

	USA nie zacznie wojny	Wojna jest prawdopodobna, ale nie jest nieunikniona	Wojna jest nieunikniona
USA nie uderzy na Libię	1	2	3
USA uderzy na Libię, ale będzie to błędem		4	5
USA uderzy na Libię i będzie to słuszne		6	7

Analiza prognoz dotyczących sytuacji z Libią nie jest jedynym przykładem podwójnych pytań wkradających się do badań opinii publicznej. Oto parę pytań, które w sondażu Harrisa zadano w celu zbadania opinii publicznej na temat ówczesnego radzieckiego generalnego sekretarza Gorbaczowa:

„Sprawia wrażenie takiego typu radzieckiego przywódcy, który rozumie, że i Sowieci, i Amerykanie mogą zniszczyć się nawzajem za pomocą pocisków nuklearnych, zatem lepiej jest dojść do sprawdzalnych porozumień na temat kontroli zbrojeń".

„Wydaje się, że jest bardziej nowoczesny, oświecony i atrakcyjny, co stanowi dobry znak dla światowego pokoju".

„Nawet mimo że wygląda znacznie bardziej nowocześnie i atrakcyjnie, błędem byłoby sądzić, że będzie się bardzo różnił od innych rosyjskich przywódców".

Ile elementów możesz wyróżnić w każdym z tych pytań? Ile opinii mogą mieć ludzie w każdym z tych przypadków? Co tak naprawdę znaczy zwykłe „zgadzam się" lub „nie zgadzam się" w takich przypadkach?

Źródło: Przytoczone w „World Opinion Update", odpowiednio październik 1985 i maj 1996.

Można przyjąć ogólną zasadę, że zawsze gdy w pytaniu pojawia się spójnik *i*, należy sprawdzić, czy nie zadaje się podwójnego pytania. Przyjrzyj się ramce zatytułowanej „Podwójne pytania i nie tylko", w której zawarto parę możliwych wariacji na ten temat.

Respondenci muszą być w stanie udzielić odpowiedzi

Prosząc respondentów o to, by udzielili informacji, należy ciągle zadawać sobie pytanie, czy mogą oni uczynić to wiarygodnie. W badaniach dotyczących wychowania dzieci można by zapytać respondentów, w jakim wieku byli, kiedy zaczęli mówić. Pomijając problem zdefiniowania *mówienia*, jest rzeczą wątpliwą, by większość respondentów pamiętała ten okres w jakimkolwiek stopniu.

Jako inny przykład może posłużyć sytuacja, w której władze samorządu studenckiego od czasu do czasu pytają swoich wyborców, na co powinny być ich zdaniem przeznaczane opłaty za studia. Z reguły prosi się respondentów o wskazanie procentu dostępnych środków, jaki powinien zostać przeznaczony na długą listę działań. Nie dysponując odpowiednią wiedzą na temat charakteru tych działań i związanych z nimi kosztów, respondenci nie mogą dać miarodajnych odpowiedzi. Na przykład koszty administracyjne uzyskałyby małe poparcie, choć mogą one mieć podstawowe znaczenie dla realizacji całego programu.

Pewna grupa badaczy, zajmująca się doświadczeniem nastolatków w prowadzeniu samochodu, obstawała przy zadaniu otwartego pytania o liczbę mil przejechanych od czasu uzyskania prawa jazdy. Chociaż konsultanci przekonywali, że tylko nieliczni będą w stanie określić tę wartość w miarę trafnie, pytanie zostało zadane. W odpowiedzi na nie niektóre nastolatki pisały o przejechaniu setek tysięcy mil.

Respondenci muszą chcieć odpowiadać

Często chcemy się czegoś dowiedzieć od osób, które nie są chętne, by się tym z nami dzielić. Na przykład Yanjie Bian zwraca uwagę na to, że często trudno jest uzyskać szczere odpowiedzi od ludzi w Chinach.

> „Ludzie są (tutaj) generalnie ostrożni, wypowiadając się przy okazjach innych niż prywatne, gdyż chodzi o przeżycie w warunkach autorytarnych. Na przykład w czasie Rewolucji Kulturalnej między rokiem 1966 a 1976 stosowanie technik sondażowych do zbierania prawdziwych i wiarygodnych danych w Chinach na temat doświadczeń życiowych Chińczyków, ich cech i postaw wobec reżimu komunistycznego było niemal niemożliwe z powodu radykalnego ustroju i napięć politycznych w całym kraju".
>
> (1994, s. 19–20)

Niekiedy zdarza się, że respondenci w Stanach Zjednoczonych mówią, że są niezdecydowani, podczas gdy faktycznie mają jakąś opinię, ale sądzą, że są w mniejszości. Mogą wobec tego odczuwać opór przed wyjawieniem tej opinii obcej osobie (ankieterowi). Biorąc ten problem pod uwagę, Organizacja Gallupa zastosowała specjalną formę „tajnego głosowania", która polega na udawaniu rzeczywistych warunków wyborczych, w których „wyborcy" gwarantuje się całkowitą anonimowość. Analizując dane wyborcze z Sondażu Gallupa od 1944 do 1988 roku, Andrew Smith i G. F. Bishop (1992) wykazali, że technika ta w istotny sposób zmniejszyła odsetek respondentów, którzy mówili, że jeszcze się nie zdecydowali, jak będą głosować.

Jednak problem ten nie ogranicza się tylko do badań sondażowych. Richard Mitchell (1991, s. 100) napotkał podobny problem w prowadzonych przez siebie badaniach terenowych wśród amerykańskich surwiwalistów[3]:

> „Na przykład surwiwaliści przyjmują postawy ambiwalentne w odniesieniu do ujawnienia swojej tożsamości czy swoich skłonności. Są przekonani, że tajemnica chroni ich przed kpinami ze strony inaczej myślącej większości, ale

[3] Surwiwaliści – ruch społeczny, stawiający na umiejętność przetrwania w najtrudniejszych warunkach (np. po ewentualnej katastrofie cywilizacyjnej) (przyp. tłum.).

ten wymuszony separatyzm zmniejsza możliwości rekrutacji i wymiany informacji (...).

«Zakonspirowani» surwiwaliści rezygnują z telefonów, przepuszczają swoją pocztę przez skrytki pocztowe, używają pseudonimów i starannie strzegą przed obcymi swoich adresów. Jednak od kiedy zostałem zaproszony na spotkania grupy, stwierdziłem, że są chętnymi do współpracy respondentami".

Pytania powinny być adekwatne

Pytania umieszczane w kwestionariuszu powinny też odnosić się do większości respondentów. Jeśli dotyczą one postaw wobec spraw, nad którymi zastanawiało się lub którymi naprawdę przejmuje się niewielu respondentów, to jest mało prawdopodobne, że wyniki będą przydatne do czegokolwiek. Naturalnie, istnieje ryzyko, że zostanie się wprowadzonym w błąd, gdyż respondenci mogą wyrażać jakieś postawy, nawet jeśli nigdy się nie zastanawiali nad daną kwestią.

Odzwierciedleniem tego są pojawiające się od czasu do czasu sytuacje, gdy badacze zadają pytania odnoszące się do fikcyjnych osób i spraw. W pewnym prowadzonym przeze mnie sondażu politycznym zapytałem respondentów o znajomość każdego z 15 polityków działających w tej miejscowości. Jedno z nazwisk – Tom Sakumoto – stworzyłem sam w ramach eksperymentu metodologicznego. W odpowiedzi 9% respondentów powiedziało, że jest on im znany. Z tych, którzy go znali, mniej więcej połowa twierdziła, że widziała go w telewizji i czytała o nim w gazetach.

Gdy otrzymuje się odpowiedzi dotyczące fikcyjnych kwestii, można je pominąć. Ale gdy kwestia jest rzeczywista, nie sposób powiedzieć, które odpowiedzi faktycznie oddają postawy, a które odzwierciedlają bezsensowne odpowiedzi na nieadekwatne pytanie.

W sytuacji idealnej chcielibyśmy, żeby respondenci po prostu mówili, że nie wiedzą, nie mają zdania, czy też że są niezdecydowani w przypadku, gdy tak jest. Jednak niestety często wymyślają oni jakieś odpowiedzi.

Najlepsze są krótkie pytania

By osiągnąć jednoznaczność i precyzję oraz b odnieść się do sedna sprawy, badacze maj skłonność do budowania długich i skomplik wanych pytań. Należy tego unikać. Responden tom często nie chce się studiować długich py tań, aby je zrozumieć. Respondent powinien być w stanie szybko przeczytać pytanie, zrozumieć jego intencję i bez trudności wybrać lub podać odpowiedź. Generalnie, najlepiej wychodź z założenia, że respondenci będą szybko czytać pytania i szybko na nie odpowiadać. Zatem zadawaj jasne, proste pytania, które będzie łatwo zinterpretować.

Unikaj form przeczących

Pojawienie się przeczenia w pytaniu kwestionariusza otwiera drogę do błędnych interpretacji. Znacząca część respondentów, zapytanych czy się zgadzają, czy nie zgadzają ze stwierdzeniem „Stany Zjednoczone nie powinny uznawać Kuby", przeoczy słówko *nie* i na tej podstawie udzieli odpowiedzi. Czyli niektórzy zgodzą się z tym stwierdzeniem, gdyż popierają uznanie, a inni się zgodzą, bo mu się sprzeciwiają. I można nie dowiedzieć się nigdy, którzy są którzy.

Podobne spostrzeżenia odnoszą się do innych „negatywnych" słów. Na przykład w badaniach dotyczących poparcia dla swobód obywatelskich zapytano respondentów, czy ich zdaniem „następującym kategoriom ludzi powinno się *zabronić* nauczania w szkołach publicznych" i przedstawiono listę, zawierającą takie pozycje, jak komunista, członek Ku Klux Klanu itd. Odpowiedzi „tak" i „nie" były podane przy każdym pytaniu. Porównanie odpowiedzi na to pytanie z innymi pytaniami, odzwierciedlającymi poparcie dla swobód obywatelskich, wskazywało wyraźnie na to, że wielu respondentów dało odpowiedź „tak" raczej, aby wyrazić zgodę na nauczanie przez taką osobę, niż aby wskazać, że powinno się jej tego zabronić. (Późniejsze badania z tej serii podające jako kategorie odpowiedzi „pozwolić" i „zabronić" dały o wiele wyraźniejsze wyniki).

Unikaj obciążonych pytań i pojęć

W rozważaniach dotyczących konceptualizacji i operacjonalizacji w rozdziale 5 mówiliśmy, że nie istnieją żadne ostatecznie prawdziwe znaczenia pojęć, którymi z reguły zajmują się nauki społeczne. *Uprzedzenia* nie mają żadnej absolutnie poprawnej definicji. To, czy dana osoba jest uprzedzona, zależy od naszej definicji tego terminu. Ta sama ogólna zasada odnosi się do odpowiedzi otrzymywanych od osób, które wypełniają kwestionariusz.

Sens czyjejś odpowiedzi na dane pytanie zależy w znacznej mierze od jego sformułowania. Jest to prawdą w odniesieniu do każdego pytania i odpowiedzi. Wydaje się, że pewne pytania bardziej niż inne zachęcają do udzielania określonych odpowiedzi. W kontekście układania kwestionariusza pojęcie **obciążenia**[4] (*bias*) odnosi się do każdej własności pytań, która nakłania respondentów do odpowiadania w określony sposób.

Większość badaczy zgodnie uznaje, że taki efekt wystąpi prawdopodobnie w pytaniu, które zaczyna się od: „Czy nie zgodzi się Pan(i) z prezydentem USA, że...", i żaden renomowany badacz nie posłużyłby się takim pytaniem. Niestety, efekt obciążający występujący w pytaniach i terminach jest znacznie subtelniejszy, niż wynikałoby to z tego przykładu.

Już samo skojarzenie jakiejś postawy lub poglądu z cieszącą się prestiżem osobą lub instytucją może działać sugerująco na odpowiedzi. Pytanie: „Czy zgadza się Pan(i), czy też nie zgadza, z ostatnią decyzją Sądu Najwyższego, mówiącą..." miałoby taki właśnie skutek. Sformułowanie takie nie musi prowadzić do konsensusu ani nawet większości popierającej stanowisko kojarzone z cieszącą się prestiżem osobą lub instytucją, ale zwiększy ono poparcie dla danego sformułowania ponad poziom, jaki by osiągnęło bez takiego skojarzenia.

[4] W języku polskim przyjęło się mówić o „pytaniach sugerujących". Jednak ze względu na nieco szerszy zakres znaczeniowy angielskiego określenia „biased" (obejmuje ono m.in. terminy obciążone emocjonalnie, np. wartościowane negatywnie, co nie musi oznaczać sformułowania „sugerującego") zdecydowano się na tłumaczenie go jako „obciążone" (przyp. tłum.).

Niekiedy wpływ różnego sformułowania pytań jest dość trudno uchwytny. Gdy na przykład Kenneth Rasinski (1989) analizował wyniki kilku edycji Generalnego Sondażu Społecznego dotyczące postaw wobec wydatków rządowych, odkrył on, że sposób określania programów wpływa na stopień osiąganego przez nie poparcia społecznego. Oto kilka porównań:

Więcej poparcia	Mniej poparcia
„Pomoc dla biednych"	„Pomoc społeczna"
„Powstrzymywanie wzrostu przestępczości"	„Egzekwowanie prawa"
„Zajmowanie się uzależnieniem od narkotyków"	„Leczenie narkomanów"
„Rozwiązywanie problemów wielkich miast"	„Pomoc wielkim miastom"
„Poprawa warunków życia czarnych"	„Pomoc czarnym"
„Ochrona bezpieczeństwa socjalnego"	„Bezpieczeństwo socjalne"

Przykładowo, w 1986 roku 62,8% respondentów było zdania, że wydaje się zbyt mało pieniędzy na „pomoc dla biednych", podczas gdy w podobnym badaniu w tym samym roku tylko 23,1% było zdania, że wydajemy zbyt mało na „pomoc społeczną".

W tym kontekście zwracaj szczególną uwagę na to, co badacze nazywają *społeczną akceptowalnością* (*social desirability*) pytań i odpowiedzi. Zawsze, gdy prosimy ludzi o informacje, odpowiadają oni poprzez filtr tego, co pozwoli im dobrze wypaść. Szczególnie odnosi się to do wywiadów bezpośrednich. Przykładowo więc, jakiś konkretny mężczyzna może uważać, że byłoby dużo lepiej, gdyby kobiety trzymano w kuchni, nie dopuszczano do głosowania, zmuszano do milczenia w miejscach publicznych itd. Jednak pytany o to, czy popiera równe prawa dla kobiet, chciałby być może uniknąć etykiety męskiego szowinisty. Zdając sobie sprawę, że jego poglądy nie dotrzymują kroku współczesnemu myśleniu, wybrałby być może odpowiedź „tak".

Najlepszym sposobem ustrzeżenia się tego problemu jest wyobrażenie sobie, jak ty sam czułbyś się, udzielając każdej z odpowiedzi, jakie zamierzasz zaproponować respondentom. Jeśli czułbyś się zakłopotany, nienormalny, nieludzki, głupi, nieodpowiedzialny lub napiętno-

wany społecznie w inny sposób z powodu którejś odpowiedzi, zastanów się poważnie, czy inni będą skłonni udzielać takich odpowiedzi.

Sugerujące działanie niektórych sformułowań często trudno przewidzieć. Zarówno w sondażach, jak i w eksperymentach warto niekiedy poprosić respondentów, by rozważyli jakieś hipotetyczne sytuacje i powiedzieli, jak sądzą, jak by się zachowali. Jako że sytuacje te dotyczą często innych ludzi, użyte do ich określenia nazwy mogą wpływać na odpowiedzi. Na przykład badacze od dawna wiedzieli, że imiona męskie hipotetycznych postaci mogą prowadzić do innych odpowiedzi niż imiona żeńskie. Badania przeprowadzone przez Josepha Kasofa (1993) wskazują na wagę konkretnych imion: czy wywołują one ogólnie pozytywne, czy negatywne obrazy w sensie atrakcyjności, wieku, inteligencji itd. Przegląd przeszłych badań, dokonany przez Kasofa, sugeruje, że istniała tendencja do używania bardziej pozytywnie wartościowanych imion dla mężczyzn niż dla kobiet.

Tak jak zawsze w badaniach, przemyśl starannie ich cel i ułóż pytania, które będą dla tego celu najbardziej odpowiednie. Nigdy nie powinieneś ulec złudnemu myśleniu, że istnieją absolutnie „dobre" i „złe" sposoby zadawania pytań. Co więcej, jeśli masz wątpliwości, jak zadać najlepsze pytanie, pamiętaj, że powinieneś zadać więcej niż jedno.

Otrzymałeś zatem kilka ogólnych wskazówek dotyczących formułowania pytań w celu zebrania danych do analizy i interpretacji. Zajmiemy się teraz tym, jak zbudować kwestionariusz.

Budowa kwestionariusza

Kwestionariusze są używane w związku z wieloma metodami obserwacyjnymi w naukach społecznych. Mimo że standaryzowane kwestionariusze są podstawowym narzędziem w badaniach sondażowych i są z nimi najbardziej bezpośrednio powiązane, to jednak są one również szeroko stosowane w eksperymentach, badaniach terenowych i innych rodzajach działalności związanej ze zbieraniem danych. Dlatego praktyczna umiejętność stworzenia kwestionariusza jest tak ważna dla badaczy. Ponieważ omawiamy rutynowe techniki tworzenia kwestionariusza, zacznijmy od niektórych zagadnień dotyczących formatu.

Ogólny format kwestionariusza

Format kwestionariusza jest równie ważny, jak charakter i słownictwo zadawanych w nim pytań. Niewłaściwie rozplanowany kwestionariusz może prowadzić do przeskakiwania pytań przez respondentów, niepewności z ich strony co do rodzaju potrzebnych danych, a nawet może doprowadzić do wyrzucenia kwestionariusza.

Ogólna zasada to swobodny i nieprzeładowany układ kwestionariusza. Niedoświadczeni badacze często obawiają się, że kwestionariusz będzie wyglądał na zbyt długi, i dlatego wciskają po kilka pytań w jednej linijce, skracają pytania i starają się zmieścić na możliwie niewielu stronach. Takie wysiłki są bezsensowne, a nawet niebezpieczne. Umieszczanie w jednej linijce więcej niż jednego pytania spowoduje, że część respondentów w ogóle przeoczy to drugie pytanie. Inni mogą źle zrozumieć skrócone wersje pytań. Ogólnie rzecz biorąc, respondenci, którzy spędzili sporo czasu na wypełnianiu pierwszej, wyglądającej na krótką, strony kwestionariusza, będą bardziej zniechęceni niż ci, którzy szybko wypełnili kilka stron, które wydawały im się początkowo raczej długie. Co więcej, ci ostatni popełnią mniej błędów i nie będą zmuszeni do wielokrotnego czytania niejasnych skróconych pytań ani do wpisywania długich odpowiedzi na stronę, na której jest mało miejsca.

Nigdy nie dość podkreślania potrzeby swobodnego układu pytań w kwestionariuszu. Zbyt ciasno zapisane kwestionariusze są fatalne, niezależnie od tego, czy są wypełniane przez samych respondentów, czy też przez przeszkolonych ankieterów. Takie kwestionariusze są też zmorą na etapie przetwarzania danych. Powiemy o tym więcej w rozdziale 14.

Formaty odpowiedzi

W jednym z najbardziej rozpowszechnionych typów pytań respondent ma wybrać jedną odpo-

wiedź z podanego zestawu. Z moich doświadczeń wynika, że najlepszym formatem dla tego typu pytań są odpowiednio od siebie oddalone kratki. Dzięki nowoczesnym edytorom tekstu zastosowanie kratek jest w dzisiejszych czasach praktyczną techniką, da się je również łatwo i poręcznie wpisać z klawiatury. Można wpisać niby-kratki używając nawiasów kwadratowych: [ ], ale jeśli kwestionariusz jest tworzony za pomocą komputera, powinno się poświęcić kilka minut na zastosowanie oryginalnych kratek, które nadadzą kwestionariuszowi bardziej profesjonalny wygląd. Oto kilka prostych przykładów:

□ ○ ❑

Zamiast wstawiania kratek do zakreślenia można wydrukować numer kodu obok każdej odpowiedzi i poprosić respondenta o zakreślenie odpowiedniego numeru odpowiedzi (patrz rysunek 9.1). Metoda ta ma dodatkową zaletę – określa numer kodu, jaki ma być wprowadzony do komputera na etapie wprowadzania danych (patrz rozdział 14). Jednak jeśli mają być zakreślane numery, trzeba udzielić respondentowi jasnych i konkretnych instrukcji, gdyż wiele osób może przekreślać numery, co utrudnia dodatkowo wprowadzanie danych. (Zauważ, że bezpieczniej jest stosować tę technikę, gdy kwestionariusze wypełniane są przez ankieterów, gdyż to oni wówczas **rejestrują** odpowiedzi).

Rysunek 9.1

Zakreślanie odpowiedzi

1. Tak
2. Nie
3. Nie wiem

Pytania warunkowe

Dość często niektóre pytania kwestionariusza odnoszą się do jednych respondentów, a nie odnoszą się do innych. Na przykład prowadząc badania dotyczące metod kontroli urodzeń, nie zechcesz zapewne pytać mężczyzn o stosowanie pigułek antykoncepcyjnych.

Taka sytuacja powstaje często, gdy badacze chcą zadać serię pytań na dany temat. Można zapytać, czy respondent należy do jakiejś konkretnej organizacji, a jeśli tak, jak często uczestniczy w jej spotkaniach, czy pełnił w niej jakieś funkcje itd. Albo też można zapytać, czy respondent słyszał coś na temat jakiejś kwestii politycznej, a później badać postawy tych respondentów, którzy o niej słyszeli.

Każde kolejne pytanie w serii takiej, jak przytoczone powyżej, nazywa się **pytaniem warunkowym**. To, czy będzie ono zadane i zostanie na nie udzielona odpowiedź, jest uwarunkowane odpowiedzią na pierwsze pytanie w tej serii[5]. Właściwe stosowanie pytań warunkowych może ułatwić respondentom wypełnianie kwestionariusza, gdyż nie muszą oni próbować odpowiadać na pytania, które się do nich nie odnoszą.

Istnieje kilka formatów pytań warunkowych. Format zaprezentowany na rysunku 9.2 jest chyba najbardziej klarowny i najefektywniejszy. Zwróć uwagę na dwa kluczowe elementy tego formatu. Po pierwsze, pytanie warunkowe jest wydzielone spośród innych pytań przez umieszczenie go z boku i ujęcie w ramkę. Po drugie, jest ono połączone strzałką z właściwą odpowiedzią na pytanie filtrujące. Jak widać na rysunku, tylko respondenci, którzy odpowiedzieli twierdząco, mają odpowiadać na pytanie warunkowe. Pozostali powinni je po prostu przeskoczyć.

Zwróć uwagę, że pytania pokazane na rysunku 9.2 można było sformułować za pomocą jednego pytania. Pytanie to mogłoby brzmieć: „Ile razy, jeśli w ogóle kiedykolwiek, paliłeś marihuanę?". Wówczas kategorie odpowiedzi mogłyby brzmieć: „nigdy", „raz", „2–5 razy" itd. To jedno pytanie odnosiłoby się do wszystkich respondentów i każdy znalazłby dla siebie właściwą kategorię odpowiedzi. Jednak pytanie takie mogłoby delikatnie sugerować respondentom, by odpowiadali, że palili marihuanę, gdyż w pytaniu głównym pytamy o to, ile razy palili, nawet jeśli dopuszcza ono owe *wyjątkowe* przypadki *takich osób, które nie paliły*

[5] Pierwsze pytanie w serii takich pytań jest po polsku określane jako „pytanie filtrujące" (przyp. tłum.).

marihuany ani razu. (Podkreślenie zastosowane w poprzednim zdaniu stanowi jasną wskazówkę co do tego, jak respondenci mogliby odczytać to pytanie). Format pytania warunkowego, zaprezentowany na rysunku 9.2, powinien zminimalizować ową delikatną sugestię, że respondenci palili marihuanę.

Poprawnie zastosowane pytania warunkowe mogą być używane bez obawy, że zdezorientują respondenta, nawet jeśli występują w postaci dość złożonych bloków. Na rysunku 9.3 został przedstawiony bardziej skomplikowany przykład.

Zdarza się, że blok pytań warunkowych jest tak duży, że zajmuje kilka stron. Załóżmy, że chciałbyś zbadać działalność polityczną studentów college'u i chciałbyś zadać więcej pytań tym studentom, którzy brali udział w wyborach federalnych, stanowych lub lokalnych. Mógłbyś oddzielić studentów, których dotyczą te pytania, za pomocą początkowego pytania w rodzaju: „Czy kiedykolwiek głosowałeś(-aś) w wyborach federalnych, stanowych lub lokalnych?", ale byłoby dezorientujące, gdybyś umieszczał pytania warunkowe w ramce ciągnącej się przez wiele stron. Lepiej byłoby

Rysunek 9.2

Format pytania warunkowego

```
23. Czy kiedykolwiek paliłeś(-aś) marihuanę?
    [ ] Tak ──┐
    [ ] Nie   │
              ▼
   Jeśli tak: Ile mniej więcej razy paliłeś(-aś) marihuanę?
              [ ] raz
              [ ] 2-5 razy
              [ ] 6-10 razy
              [ ] 11-20 razy
              [ ] więcej niż 20 razy
```

Rysunek 9.3

Blok pytań warunkowych

```
Czy był(a) Pan(i) kiedykolwiek porwany(a) przez kosmitów?
 [ ] Tak ──┐
 [ ] Nie   │
           ▼
  Jeśli tak: Czy pozwolili Panu(i) kierować statkiem kosmicznym?
             [ ] Tak ──┐
             [ ] Nie   │
                       ▼
            Jeśli tak: Jak prędko Pan(i) leciał(a)?
                       [ ] Z zawrotną prędkością
                       [ ] Z malutką prędkością
```

Rysunek 9.4

Reguły przejścia

```
13. Czy kiedykolwiek brał(a) Pan(i) udział wyborach krajowych, stanowych lub lokalnych?
 [ ] Tak (proszę przejść do pytań 14-25).
 [ ] Nie (proszę opuścić pytania 14-25 i przejść od razu do pytania 26 na stronie 8).
```

umieścić po każdej odpowiedzi instrukcję w nawiasach, mówiącą respondentom, czy mają odpowiadać na pytania warunkowe, czy też mają je przeskoczyć. Metoda ta przedstawiona jest na rysunku 9.4.

Jako uzupełnienie tych instrukcji warto umieścić instrukcję na górze każdej strony, zawierającej tylko pytania poprzedzone filtrem. Może to być na przykład sformułowanie: „Ta strona przeznaczona jest tylko dla respondentów, którzy głosowali w wyborach federalnych, stanowych lub lokalnych". Tego rodzaju jasne instrukcje oszczędzają respondentom frustracji, związanej z czytaniem pytań, które się do nich nie odnoszą, i łamaniem sobie nad nimi głowy, oraz zwiększają prawdopodobieństwo uzyskania odpowiedzi od tych, do których te pytania się odnoszą.

Pytania tabelaryczne

Dość często zdarza się, że chcemy zadać kilka pytań, mających taki sam zestaw kategorii odpowiedzi. Typowym przypadkiem jest stosowanie likertowskich kategorii odpowiedzi. W takich sytuacjach możliwe jest często stworzenie macierzy, czyli tabeli zawierającej pytania i odpowiedzi, na wzór tej pokazanej na rysunku 9.5.

Format taki pod kilkoma względami ma przewagę nad innymi. Po pierwsze, efektywnie wykorzystuje przestrzeń. Po drugie, prawdopodobnie respondentom szybciej pójdzie wypełnienie zestawu pytań przedstawionego w ten sposób. Ponadto zastosowanie tego formatu może przyczynić się do zwiększenia porównywalności odpowiedzi udzielanych na różne pytania, zarówno dla respondentów, jak i dla badaczy. Respondenci, mając możliwość dokonania szybkiego przeglądu swoich wcześniejszych odpowiedzi, mogą wybrać odpowiedź „zdecydowanie się zgadzam" lub „zgadzam się" dla danego stwierdzenia, porównując stopień tej zgody z wcześniej udzielonymi odpowiedziami w danym zestawie.

Stosowanie tego formatu wiąże się jednak również z pewnymi niebezpieczeństwami. Jego zalety mogą być zachętą do takiego ułożenia pytania, żeby pasowało do formatu tabelarycznego, nawet gdy właściwszy byłby inny, bardziej specyficzny zestaw odpowiedzi. Tabelaryczny format pytań może też skłaniać niektórych respondentów do preferowania pewnego zestawu odpowiedzi: mogą oni na przykład przejawiać wzorzec, powiedzmy, zgadzania się ze wszystkimi stwierdzeniami. Jest to szczególnie prawdopodobne, gdy zestaw stwierdzeń zaczyna się od takich, które wskazują na jakieś konkretne poglądy (na przykład liberalne poglądy polityczne), a tylko kilka ostatnich stwierdzeń prezentuje poglądy odmienne. Respondenci mogą zakładać, że wszystkie stwierdzenia przedstawiają podobne poglądy, i czytając szybko mogą źle odczytać niektóre z nich, a w konsekwencji udzielać błędnych odpowiedzi. Można nieco złagodzić ten problem, różnicując stwierdzenia prezentujące różne poglądy oraz skracając i upraszczając wszystkie stwierdzenia.

Kolejność pytań w kwestionariuszu

Na odpowiedzi respondentów może wpływać także kolejność, w jakiej przedstawiane są pytania kwestionariusza. Po pierwsze, pojawienie się jakiegoś pytania może wpływać na odpowie-

Rysunek 9.5

Tabelaryczny format pytania

17. Przy *każdym* z poniższych stwierdzeń proszę zaznaczyć, czy Pan(i) się z nim zdecydowanie zgadza (ZZ), zgadza (Z), nie zgadza (N), zdecydowanie nie zgadza (ZN), czy też nie ma Pan(i) zdania (NW).

	ZZ	Z	N	ZN	NW
a. Nasz kraj potrzebuje więcej prawa i porządku ...	[ ]	[ ]	[ ]	[ ]	[ ]
b. Policja amerykańska powinna zostać rozbrojona ..	[ ]	[ ]	[ ]	[ ]	[ ]
c. W czasie zamieszek powinno się natychmiast strzelać do rabusiów	[ ]	[ ]	[ ]	[ ]	[ ]

itd.

dzi udzielane na następne. Na przykład jeśli zada się kilka pytań na temat niebezpieczeństwa, jakie dla Stanów Zjednoczonych przedstawia terroryzm, a następnie zapyta o to (pytanie otwarte), co zdaniem respondentów stanowi niebezpieczeństwo dla Stanów Zjednoczonych, to terroryzm uzyska więcej wskazań, niż by uzyskał w innym kontekście. W tej sytuacji lepiej byłoby zadać najpierw pytanie otwarte.

Podobnie też gdy prosi się respondentów o określenie ich ogólnej religijności („Jak ważna jest dla Pana(i) ogólnie Pana(i) religia?"), ich odpowiedzi na kolejne pytania, dotyczące poszczególnych aspektów religijności, będą ukierunkowane na spójność z poprzednim stwierdzeniem. Prawdą jest również odwrotne zjawisko. Jeśli najpierw zapyta się respondentów o poszczególne aspekty ich religijności, ich późniejsze ogólne określenie będzie odzwierciedleniem wcześniejszych odpowiedzi.

Wpływ kolejności pytań nie jest jednolity. Gdy J. Edwin Benton i John Daly (1991) przeprowadzali badania sondażowe władz lokalnych, zauważyli, że mniej wykształceni respondenci ulegali wpływowi kolejności pytań w kwestionariuszu w większym stopniu niż ci, którzy mieli lepsze wykształcenie.

Niektórzy badacze próbują przezwyciężyć to oddziaływanie poprzez losowe ułożenie kolejności pytań. Z reguły jest to próżny wysiłek. Przede wszystkim ułożony losowo zestaw pytań wydaje się respondentom uderzająco chaotyczny i bez sensu. Losowa kolejność utrudnia też udzielanie odpowiedzi, gdyż respondenci muszą nieustannie przeskakiwać z tematu na temat, co rozprasza uwagę. W końcu wreszcie, nawet losowa kolejność pytań będzie miała omawiany wyżej wpływ – tyle że nie ma się nad tym efektem żadnej kontroli.

Najbezpieczniej jest zachować odpowiednią wrażliwość na ten problem. Mimo że nie da się całkowicie wyeliminować wpływu kolejności pytań, można spróbować oszacować ten wpływ, tak aby móc sensownie interpretować wyniki. Jeśli w danych badaniach kolejność pytań wydaje się szczególnie ważna, można stworzyć więcej niż jedną wersję kwestionariusza z różną kolejnością pytań. Można wówczas określić wpływ kolejności pytań poprzez porównanie odpowiedzi dla różnych wersji. Wreszcie, można przetestować wstępnie kwestionariusz w różnych formach. (Zagadnienia związane z pilotażem omówimy wkrótce).

Poprawna kolejność pytań różni się w zależności od tego, czy wywiad jest prowadzony przez ankietera, czy też kwestionariusz jest wypełniany samodzielnie przez respondenta. Z reguły w tym drugim przypadku najlepiej jest, gdy kwestionariusz zaczyna się od najciekawszego zestawu pytań. Ewentualni respondenci, którzy od niechcenia rzucają okiem na owe pierwsze pytania, powinni nabrać chęci, by na nie odpowiedzieć. Mogą one dotyczyć postaw, które bardzo pragną wyrazić. Zarazem jednak pierwsze pytania nie powinny stwarzać poczucia zagrożenia (złym pomysłem byłoby zaczynanie wywiadu od pytań o zachowania seksualne czy zażywanie narkotyków). Pytania o nudniejsze dane demograficzne (wiek, płeć itp.) powinny z zasady być umieszczane na końcu kwestionariusza wypełnianego samodzielnie przez respondentów. Umieszczenie ich na początku, co jest skłonnych czynić wielu niedoświadczonych badaczy, nadaje kwestionariuszowi wygląd formularza i osoba otrzymująca go może nie mieć motywacji do wypełniania go.

W przypadku wywiadów przeprowadzanych przez ankietera stosuje się na ogół odwrotne zasady. Gdy drzwi ewentualnego respondenta zostaną uchylone, ankieter musi szybko nawiązać kontakt. Najlepiej jeśli po krótkim przedstawieniu celów badania ankieter zacznie od wyliczenia członków gospodarstwa domowego, uzyskując dane demograficzne każdego z nich. Na takie pytania z reguły szybko się odpowiada i nie są one odczuwane jako zagrażające. Po nawiązaniu wstępnego kontaktu ankieter może przejść do dziedziny postaw i bardziej delikatnych kwestii. Wywiad rozpoczynający się od pytania: „Czy wierzy Pan(i) w czary?" zapewne zakończyłby się dość szybko.

Instrukcje do kwestionariusza

Każdy kwestionariusz, niezależnie od tego, czy ma być wypełniany samodzielnie przez respondenta, czy też służy do wywiadu przeprowadzanego przez ankietera, powinien zawierać

jasne instrukcje i komentarze wszędzie tam, gdzie jest to potrzebne.

Dobrze jest umieścić na początku każdego kwestionariusza, przeznaczonego do samodzielnego wypełniania przez respondentów, podstawowe instrukcje dotyczące wypełniania go. Mimo że wielu ludzi ma w dzisiejszych czasach doświadczenie w wypełnianiu różnych formularzy i kwestionariuszy, zacznij swój kwestionariusz od wyjaśnienia dokładnie respondentom, czego od nich chcesz: że mają wskazać odpowiedzi na pewne pytania poprzez wstawienie znaku wyboru albo znaku X w kratce obok właściwej odpowiedzi lub poprzez wpisanie odpowiedzi tam, gdzie są o to poproszeni. Jeśli stosuje się wiele pytań otwartych, należy dać respondentom jakieś wskazówki co do tego, czy oczekuje się od nich odpowiedzi zwięzłych, czy raczej dłuższych. Jeśli chcesz zachęcić respondentów do rozwinięcia odpowiedzi na pytania zamknięte, powinieneś to zaznaczyć.

Jeśli kwestionariusz zawiera sekcje tematyczne – obejmujące postawy polityczne, religijne, dane społeczno-demograficzne – daj do każdej z nich krótkie wprowadzenie, dotyczące jej zawartości i celu. Na przykład takie: „W tej części chcemy się dowiedzieć, co ludzie uważają za najważniejsze problemy swojej miejscowości". Pytania metryczkowe na końcu kwestionariusza przeznaczonego do samodzielnego wypełniania mogą mieć taki wstęp: „Na koniec chcielibyśmy uzyskać na Pana(i) temat kilka informacji. Umożliwią nam one porównanie poglądów różnych typów ludzi na badane przez nas kwestie".

Takie krótkie wprowadzenia pomagają respondentom zrozumieć kwestionariusz. Dzięki nim może on wyglądać mniej chaotycznie, zwłaszcza jeśli służy pozyskaniu wielu różnych danych. Pomagają także wprowadzić respondenta we właściwy stan umysłu podczas udzielania odpowiedzi.

Niektóre pytania mogą wymagać specjalnych instrukcji, ułatwiających poprawne udzielanie odpowiedzi. Dotyczy to zwłaszcza sytuacji, gdy dane pytanie odbiega od ogólnej instrukcji dla całego kwestionariusza. Kilka konkretnych przykładów zilustruje tę sytuację.

Mimo prób układania wzajemnie rozłącznych odpowiedzi na pytania zamknięte, często możliwa jest więcej niż jedna odpowiedź. Jeśli chcesz uzyskać tylko jedną odpowiedź, musi to być absolutnie jasne. Na przykład: „Wybierz z poniższej listy najważniejszy powód, dla którego zdecydowałeś(-aś) się iść do college'u". Często po zasadniczym pytaniu umieszcza się w nawiasie adnotację: „Proszę wskazać jedną najważniejszą odpowiedź". Z kolei, jeśli chcesz, by respondenci wybierali tyle odpowiedzi, ile jest dla nich właściwych, powinieneś również jasno to określić.

Jeśli zbiór odpowiedzi ma być ułożony przez respondenta według kolejności, instrukcja powinna o tym mówić. Powinien być też zastosowany inny format odpowiedzi (na przykład spacje zamiast kratek). Instrukcja powinna mówić, ile odpowiedzi należy uszeregować (na przykład: wszystkie, tylko pierwszą i drugą, tylko pierwszą i ostatnią, najważniejszą i najmniej ważną). W instrukcji powinien też być określony sposób szeregowania (na przykład: „Wstaw 1 obok najważniejszego pytania, 2 obok następnego według ważności itd."). Szeregowanie odpowiedzi jest często trudne dla respondentów, gdyż może od nich wymagać kilkakrotnego czytania całej listy, zatem technika ta powinna być stosowana w sytuacjach, gdy żadna inna metoda nie przyniosłaby pożądanych rezultatów.

W wieloczęściowych pytaniach tabelarycznych wskazane jest umieszczanie specjalnych instrukcji, chyba że format ten stosowany jest w całym kwestionariuszu. Niekiedy prosi się respondentów o wybranie jednej odpowiedzi w każdej kolumnie pytania tabelarycznego, w innych kwestionariuszach będzie to jedna odpowiedź w rzędzie. Jeśli kwestionariusz zawiera oba typy pytań tabelarycznych, wskazane jest dodanie instrukcji, wyjaśniającej, jaki sposób jest właściwy w danym przypadku.

Pilotaż kwestionariusza

Niezależnie od tego, jak starannie badacze zaprojektują instrument zbierania danych, jakim jest kwestionariusz, zawsze istnieje możliwość – a w zasadzie pewność – pomyłki. Zawsze popełnia się jakieś błędy: pytanie wieloznaczne, pytanie, na które nie da się odpowiedzieć, bądź inne naruszenie omówionych reguł.

Najpewniejszym zabezpieczeniem przed takimi błędami jest przetestowanie całości lub części kwestionariusza, czyli pilotaż. Na przykład rozdaj kwestionariusz dziesięciu osobom w swojej lidze kręglowej. Zazwyczaj nie jest niezbędne, by osoby testujące stanowiły reprezentatywną próbę, ale powinieneś wybrać ludzi, do których pytania kwestionariusza przynajmniej mają zastosowanie.

Najogólniej rzecz biorąc, lepiej jest poprosić respondentów o wypełnienie kwestionariusza, niż o przeczytanie go i wychwycenie błędów. Aż nazbyt często zdarza się, że jakieś pytanie wygląda zrozumiale przy pierwszym czytaniu, ale okazuje się, że nie da się na nie odpowiedzieć.

Stanley Presser i Johnny Blair (1994) opisują kilka różnych strategii pilotażu i określają skuteczność każdej z nich. Dostarczają też danych o kosztach różnych metod.

Jest jeszcze wiele innych wskazówek i zaleceń, dotyczących budowy kwestionariusza, ale uwzględnienie ich wszystkich złożyłoby się na odrębną książkę. Zakończę teraz to omówienie przykładem prawdziwego kwestionariusza, pokazując, jak niektóre z moich uwag znajdują potwierdzenie w praktyce.

Zanim jednak przejdziemy do przykładu, chciałbym napomknąć o kluczowym aspekcie kwestionariusza, o którym mowa jest w rozdziale 14: prekodowaniu. Informacje zbierane za pomocą kwestionariusza z reguły są przekształcane na jakiś format odczytywany przez komputer, zatem najczęściej jest wskazane umieszczenie w samym kwestionariuszu instrukcji związanych z przetwarzaniem danych. Instrukcje te wskazują miejsce zapisywania poszczególnych informacji w plikach komputerowych. W rozdziale 15 omówiony zostanie charakter takiego zapisu i wskazane właściwe notacje w kwestionariuszu. Zauważ jednak w ramach wstępu, że nasza ilustracja jest prekodowana za pomocą tajemniczych liczb pojawiających się obok pytań i kategorii odpowiedzi.

Całościowy przykład

Rysunek 9.6 jest częścią kwestionariusza stosowanego w Generalnym Sondażu Społecznym (*General Social Survey, GSS*) przez Krajowe Centrum Badania Opinii przy Uniwersytecie Chicago. Kwestionariusz ten dotyczy postaw ludzi wobec rządu i jest przeznaczony do samodzielnego wypełniania przez respondentów.

Kwestionariusz do samodzielnego wypełniania

Zajmowaliśmy się dotychczas tym, jak formułować pytania i jak właściwie ułożyć kwestionariusz. Jednak wysiłek związany z wykonaniem tych ważnych zadań byłby zmarnowany, gdyby kwestionariusz nie dostarczał potrzebnych danych – co oznacza, że respondenci muszą go faktycznie wypełnić. Przejdziemy więc teraz do głównej metody uzyskiwania odpowiedzi na pytania kwestionariusza.

Kilkakrotnie w tym rozdziale odnosiłem się do podziału na kwestionariusze wypełniane w trakcie wywiadu przez ankietera oraz kwestionariusze wypełniane samodzielnie przez respondentów. Tak naprawdę można mówić o trzech głównych metodach wypełniania kwestionariusza przez próbę respondentów: samodzielne wypełnianie kwestionariusza przez respondentów, wypełnianie kwestionariusza przez ankietera podczas bezpośredniego wywiadu i sondaże telefoniczne. W tym podrozdziale i dwóch następnych omówimy kolejno każdą z tych metod.

Najczęstszą formą badań za pomocą kwestionariusza wypełnianego samodzielnie przez respondentów jest ankieta pocztowa. Istnieje jednak także kilka innych często stosowanych technik. Niekiedy właściwe jest rozdanie kwestionariuszy grupie respondentów, zebranych w tym samym czasie i w tym samym miejscu. W ten sposób można przeprowadzić sondaż wśród studentów psychologii podczas wykładu. Uczniowie mogą być badani podczas spotkań z wychowawcą.

Ostatnio przeprowadzono kilka eksperymentów dotyczących dostarczania kwestionariuszy do domów. Pracownik badawczy przynosi kwestionariusz do domu osobom dobranym do próby i wyjaśnia cele badań. Następnie kwestio-

10. Oto pewne działania, które rząd może podjąć na rzecz gospodarki. Proszę zaznaczyć poprzez zakreślenie odpowiedniego numeru, czy jest Pan(i) za, czy przeciw.

1. Jestem zdecydowanie za
2. Jestem za
3. Nie jestem ani za, ani przeciw
4. Jestem przeciw
5. Jestem zdecydowanie przeciw

	PROSZĘ ZAKREŚLIĆ NUMER					
a. Prawna regulacja płac	1	2	3	4	5	28/
b. Prawna regulacja cen	1	2	3	4	5	29/
c. Cięcia w wydatkach rządowych	1	2	3	4	5	30/
d. Finansowanie przez rząd projektów wspomagających tworzenie nowych miejsc pracy	1	2	3	4	5	31/
e. Mniej rządowych regulacji w dziedzinie biznesu	1	2	3	4	5	32/
f. Wspieranie przemysłu w rozwijaniu nowych produktów i technologii	1	2	3	4	5	33/
g. Wspieranie upadających gałęzi przemysłu w celu ratowania miejsc pracy	1	2	3	4	5	34/
h. Ograniczenie tygodnia pracy w celu stworzenia dodatkowych miejsc pracy	1	2	3	4	5	35/

11. Poniżej zestawiono różne dziedziny wydatków rządu. Proszę zaznaczyć poprzez zakreślenie odpowiedniego numeru, czy jest Pan(i) za zwiększeniem, czy za zmniejszeniem wydatków rządowych w każdej z tych dziedzin. Proszę pamiętać, że wybór określenia „powinno się wydawać znacznie więcej" może wymagać zwiększenia podatków, by sfinansować takie wydatki.

1. Powinno się wydawać znacznie więcej
2. Powinno się wydawać więcej
3. Powinno się wydawać tyle, co obecnie
4. Powinno się wydawać mniej
5. Powinno się wydawać znacznie mniej
8. Nie mam zdania

	PROSZĘ ZAKREŚLIĆ NUMER						
a. Ochrona środowiska	1	2	3	4	5	8	36/
b. Zdrowie	1	2	3	4	5	8	37/
c. Policja i przestrzeganie prawa	1	2	3	4	5	8	38/
d. Edukacja	1	2	3	4	5	8	39/
e. Wojsko i obronność	1	2	3	4	5	8	40/
f. Świadczenia emerytalne	1	2	3	4	5	8	41/
g. Świadczenia dla bezrobotnych	1	2	3	4	5	8	42/
h. Kultura i sztuka	1	2	3	4	5	8	43/

12. Gdyby rząd *musiał* dokonać wyboru *pomiędzy zmniejszaniem* inflacji a zmniejszaniem bezrobocia, co Pana(i) zdaniem powinno mieć bezwzględne pierwszeństwo?

Zmniejszanie inflacji	1	44/
Zmniejszanie bezrobocia	2	
Nie mam zdania	3	

13. Czy Pana(i) zdaniem związki zawodowe mają w naszym kraju za dużo czy za mało władzy?

O wiele za dużo władzy	1	45/
Za dużo władzy	2	
Mniej więcej tyle co trzeba władzy	3	
Za mało władzy	4	
O wiele za mało władzy	5	
Nie mam zdania	8	

Rysunek 9.6

Próbka kwestionariusza (ciąg dalszy)

14. Czy środowiska gospodarcze mają za dużo, czy za mało władzy?

O wiele za dużo władzy	1	46/
Za dużo władzy	2	
Mniej więcej tyle co trzeba władzy	3	
Za mało władzy	4	
O wiele za mało władzy	5	
Nie mam zdania	8	

15. Czy rząd federalny ma za dużo, czy za mało władzy?

O wiele za dużo władzy	1	47/
Za dużo władzy	2	
Mniej więcej tyle co trzeba władzy	3	
Za mało władzy	4	
O wiele za mało władzy	5	
Nie mam zdania	8	

16. Proszę określić, czy generalnie rzecz biorąc związki zawodowe są dobre dla kraju?

Znakomite	1	48/
Bardzo dobre	2	
Dość dobre	3	
Niezbyt dobre	4	
Wcale nie są dobre	5	
Nie mam zdania	8	

17. Jaka powinna być Pana(i) zdaniem rola rządu w każdej z poniższych dziedzin gospodarki?

1. Rząd powinien być właścicielem
2. Rząd powinien kontrolować ceny i zyski, ale nie powinien być właścicielem
3. Rząd nie powinien być właścicielem ani nie powinien kontrolować cen i zysków
8. Nie mam zdania

	PROSZĘ ZAKREŚLIĆ NUMER				
a. Energetyka	1	2	3	8	49/
b. Hutnictwo	1	2	3	8	50/
c. Bankowość i ubezpieczenia	1	2	3	8	51/

18. Czy Pana(i) zdaniem do obowiązków rządu powinno należeć...

1. Zdecydowanie powinno
2. Raczej powinno
3. Raczej nie powinno
4. Zdecydowanie nie powinno
8. Nie mam zdania

	PROSZĘ ZAKREŚLIĆ NUMER					
a. Zapewnienie pracy każdemu, kto chce pracować	1	2	3	4	8	52/
b. Utrzymywanie kontroli cen	1	2	3	4	8	53/
c. Zapewnienie chorym opieki zdrowotnej	1	2	3	4	8	54/
d. Zapewnienie osobom starszym odpowiedniego standardu życia	1	2	3	4	8	55/

nariusz zostaje u respondenta do wypełnienia, i zostaje odebrany później.

Dostarczanie do domu i wysyłka pocztowa mogą być również stosowane łącznie. Kwestionariusze wysyła się do rodzin, a potem pracownik badawczy chodzi po domach, zbierając je i sprawdzając kompletność ich wypełnienia. Odwrotnością tej techniki jest sytuacja, w której kwestionariusze dostarcza się osobiście respondentom z prośbą o odesłanie ich pocztą do agencji badawczej po wypełnieniu.

Ogólnie rzecz biorąc, jeśli pracownik badawczy sam przynosi kwestionariusz, odbiera go, albo zarówno przynosi, jak i odbiera, zwrot-

ność kwestionariuszy jest większa niż w przypadku zwykłych ankiet pocztowych. Prawdopodobnie dodatkowe eksperymenty z tą techniką wskażą inne sposoby zwiększenia zwrotności przy zmniejszeniu kosztów. Jednak reszta tego podrozdziału jest poświęcona ankiecie pocztowej, gdyż pozostaje ona nadal najbardziej typową formą kwestionariusza wypełnianego przez respondentów.

Dystrybucja pocztowa i zwroty

Zasadniczą metodą zbierania danych w ankiecie pocztowej jest wysłanie kwestionariusza wraz z listem wyjaśniającym i kopertą zwrotną z naklejonym znaczkiem. Respondent ma wypełnić kwestionariusz, włożyć go do koperty i wysłać z powrotem. Jeśli akurat przypadkowo otrzymałeś kiedyś kwestionariusz i zapomniałeś go oddać, warto byłoby przypomnieć sobie przyczyny, dla których go nie oddałeś, i przypominać je sobie za każdym razem, gdy chcesz wysłać kwestionariusze do innych.

Wspólną przyczyną niezwracania kwestionariuszy jest to, że to za duży kłopot. Aby przezwyciężyć ten problem, badacze opracowali parę sposobów ułatwiających oddanie kwestionariusza. Na przykład samowysyłający się kwestionariusz nie wymaga koperty zwrotnej: gdy kwestionariusz zostanie złożony w odpowiedni sposób, adres zwrotny ukaże się na odwrocie. Respondent nie musi zatem martwić się, że zgubił kopertę.

Dostępne są również bardziej wyrafinowane metody. Opisany dalej w tym rozdziale kwestionariusz dla studentów był zszyty w formie broszurki ze specjalną dwuskrzydłową tylną okładką. Po wypełnieniu kwestionariusza respondent musiał tylko rozwinąć dodatkową tylną okładkę, owinąć nią broszurkę i zakleić całość za pomocą samoprzylepnego paska umocowanego wzdłuż rogów okładki. Na dodatkowej tylnej okładce umieszczony był mój adres. Gdy parę lat później powtarzałem te badania, ulepszyłem swoją metodę. Dodatkowa okładka była dołączona zarówno do okładki tylnej, jak i przedniej: jedna do wysłania kwestionariusza do respondentów, a druga do odesłania go z powrotem – w ten sposób uniknąłem stosowania kopert.

Najważniejsze, że wszystko, co zrobisz w celu ułatwienia wypełniania i zwracania kwestionariuszy, ulepszy twoje badania. Wyobraź sobie, że otrzymujesz kwestionariusz bez żadnych wymogów dotyczących sposobu oddania go badaczowi. Przyjmijmy, że musisz: (1) znaleźć kopertę, (2) napisać na niej adres, (3) dowiedzieć się, ile kosztuje przesyłka oraz (4) nakleić odpowiednie znaczki. Jakie jest prawdopodobieństwo, że wyślesz kwestionariusz?

Warto dodać kilka krótkich uwag na temat opcji wysyłki. Dotyczą one wysyłania kwestionariuszy i otrzymywania ich z powrotem[6]. Jeśli chodzi o wysyłkę kwestionariuszy, można zasadniczo wybierać między wysyłką zwykłą a hurtową. Wysyłka pocztą zwykłą jest pewniejsza, a wysyłka hurtowa znacznie tańsza. (Sprawdź w swoim urzędzie pocztowym ceny i procedury). Natomiast co do sposobu otrzymywania kwestionariuszy z powrotem, możesz wybierać między wcześniejszym naklejeniem znaczków a zryczałtowaną opłatą za odpowiedź zwrotną (czyli opłatą uiszczaną przez odbiorcę przesyłki po jej odebraniu – dop. tłum.). Zróżnicowanie kosztów jest w tym przypadku bardziej złożone. Gdy używa się znaczków, płaci się za nie niezależnie od tego, czy ludzie wyślą kwestionariusze z powrotem, czy nie. Przy opłacie zryczałtowanej płaci się tylko za faktycznie wysłane kwestionariusze, ale uiszcza się dodatkową opłatę w wysokości około 5 centów. Oznacza to, że znaczki są tańsze, gdy zwraca się dużo kwestionariuszy, opłata zryczałtowana zaś jest tańsza, gdy zwraca się ich niewiele (a nie wie się z wyprzedzeniem, ile kwestionariuszy zostanie odesłanych z powrotem).

Jest jeszcze wiele czynników, które można brać pod uwagę wybierając wariant ankiety pocztowej. Na przykład niektórzy badacze uważają, że używanie znaczków pocztowych komunikuje więcej „czynnika ludzkiego" i szczerości, niż ma to miejsce w przypadku korzystania z wysyłki hurtowej i zryczałtowanej opłaty za odpowiedź zwrotną. Inni natomiast martwią

[6] Te rozważania autora odnoszą się do warunków amerykańskich (przyp. tłum.).

się, że respondenci mogą odklejać znaczki nad parą, by użyć ich do celów innych niż odesłanie kwestionariusza. Ponieważ zarówno wysyłka hurtowa, jak i zryczałtowana opłata za przesyłki zwrotne wymagają zawarcia umów z pocztą, zapewne uznasz, że znaczki pocztowe są znacznie wygodniejsze przy niedużych sondażach.

Monitorowanie zwrotów

Wysyłka kwestionariuszy pocztą wiąże się z nowym pytaniem badawczym, które może okazać się istotne dla prowadzonych badań. Badacze nie powinni siedzieć bezczynnie, podczas gdy kwestionariusze napływają z powrotem. W tym czasie powinni oni rozpocząć dokładne odnotowywanie zmiennego poziomu zwrotów wśród respondentów.

Nieocenionym narzędziem w tej działalności jest wykres poziomu zwrotów. Dzień, w którym zostały wysłane kwestionariusze, zostaje nazwany na wykresie dniem 1. Każdego następnego dnia nanosi się na wykres liczbę przysłanych z powrotem kwestionariuszy. Najlepiej jest połączyć dwa wykresy. Jeden pokazuje liczbę kwestionariuszy przysłanych każdego dnia – rosnącą, a potem malejącą. Drugi ukazuje ich liczbę lub procent narastająco. Po części działalność ta stanowi nagrodę dla badaczy – bądź co bądź rysują oni obraz swojego udanego gromadzenia danych. Co ważniejsze jednak, jest ona wskazówką co do tego, jak przebiega owo gromadzenie danych. Jeśli planuje się listy przypominające, wykres jest wskazówką co do tego, kiedy należy rozpocząć wysyłanie takich monitów. (Daty kolejnych monitów powinny być zaznaczane na wykresie).

W miarę przychodzenia wypełnionych kwestionariuszy, powinny one być kolejno otwierane, starannie sprawdzane i opatrywane numerami identyfikacyjnymi. Numery te powinny być przypisywane w kolejności nadchodzenia kwestionariuszy, nawet jeśli nadano im już jakieś inne numery identyfikacyjne. Dwa przykłady powinny zilustrować istotne pożytki z takiej procedury.

Załóżmy, że badasz postawy wobec jakiejś postaci z życia politycznego. W trakcie zbierania danych w mediach wybucha afera związana z pozamałżeńskimi przygodami tego polityka. Znając datę publicznego ujawnienia tych spraw i daty przysyłania kwestionariuszy, będziesz w stanie zbadać efekt tego ujawnienia. (Przypomnij sobie rozważania na temat historii w związku z eksperymentami w rozdziale 8).

Mniej sensacyjny przykład zastosowania seryjnych numerów identyfikacyjnych stanowi ich wartość w ocenianiu wpływu braków odpowiedzi w przeprowadzonym sondażu. Unikanie stosowania bardziej bezpośrednich testów tego uwarunkowania może być zarazem wyrazem założenia, że ci, którzy nie wypełnili kwestionariusza, będą bardziej podobni do tych, którzy późno przysłali swoje odpowiedzi, niż do tych, którzy od razu odesłali wypełniony kwestionariusz. Analiza kwestionariuszy otrzymanych w różnych terminach może być w takiej sytuacji wykorzystana do oszacowania stopnia obciążenia próby. Na przykład jeśli średnia ocen badanych studentów wykazuje stały spadek w miarę zbierania danych, przyjmując wyższe wartości wśród studentów, którzy odpowiedzieli od razu, niższe zaś wśród tych, którzy zwlekali z odpowiedzią, możesz wstępnie wnioskować, że ci, którzy w ogóle nie odpowiedzieli, mieli jeszcze niższe średnie ocen. Choć niewskazane byłoby statystyczne szacowanie wielkości tego odchylenia, możesz pozwolić sobie na przybliżone szacunki na podstawie obserwowanych wzorców.

Po wytypowaniu respondentów do wysłania im listów przypominających, przygotowania do wysłania takich listów należy poczynić, gdy zostały zwrócone kwestionariusze. Bardziej szczegółowo zajmiemy się tym procesem w opisie przypadku, przedstawionym w dalszej części tego podrozdziału.

Listy przypominające

Istnieje kilka sposobów przypominania respondentom prośby o odesłanie wypełnionego kwestionariusza. Najprostszy polega na wysłaniu listu z dodatkową zachętą do uczestnictwa w badaniach. Jednak lepszą metodą jest wysłanie z listem przypominającym nowego egzem-

plarza kwestionariusza. Jeśli potencjalni respondenci nie odesłali kwestionariuszy po dwóch czy trzech tygodniach, to prawdopodobnie zgubili je. Nadejście samego listu przypominającego może zachęcić ich do poszukiwania kwestionariuszy, ale jeśli znalezienie ich nie będzie łatwe, samo wysłanie listu pójdzie na marne.

Literatura metodologiczna wskazuje dobitnie, że listy przypominające są skuteczną metodą zwiększania odsetka zwrotów w ankiecie pocztowej. Najogólniej rzecz biorąc, im dłużej potencjalny respondent zwleka z odpowiedzią, tym mniej jest prawdopodobne, że w ogóle odpowie. Tak więc listy przypominające, wysłane w odpowiednim momencie, dostarczają dodatkowego bodźca do udzielenia odpowiedzi.

Efekty wysłania listów przypominających będą widoczne na krzywych poziomu zwrotów, rejestrujących proces zbierania danych. W następstwie pierwotnej wysyłki zwroty będą rosły, a następnie malały, powtórna wysyłka spowoduje wzrost zwrotów, podobnie jak następne monity. Wydaje się, że praktycznie najskuteczniejsze są trzy wysyłki (pierwotna i dwie przypominające).

Ważny jest też moment wysłania listów przypominających. Literatura metodologiczna jest w tym punkcie mniej precyzyjna, ale z moich doświadczeń wynika, że dwa czy trzy tygodnie to rozsądny odstęp między wysyłkami. (Ten okres może być zwiększony o parę dni, jeśli czas doręczania przesyłki – tam czy z powrotem – jest dłuższy niż dwa, trzy dni).

Badacze prowadzący kilka sondaży na tej samej populacji w danym czasie mogą wypracować bardziej zaawansowane reguły. Biuro Badań Sondażowych przy Uniwersytecie Hawajskim często prowadzi badania sondażowe wśród studentów i było w stanie znacznie udoskonalić swoje procedury wysyłkowe. Pracownicy tego biura odkryli stały wzorzec zwrotów, który prawdopodobnie nie zależy od tematu sondażu, jakości instrumentu badawczego itd. W ciągu pierwszych dwóch tygodni od pierwszej wysyłki zwracanych jest około 40% kwestionariuszy, w ciągu dwóch tygodni od pierwszego przypomnienia – dodatkowe 20%, a w ciągu dwóch tygodni od ostatniego przypomnienia – dodatkowe 10%. (Te wielkości zwrotów zostały uzyskane przy procedurze wysyłania dodatkowych kwestionariuszy, a nie tylko samych listów). Możesz uzyskać inne wyniki, ale ten przykład powinien uzmysłowić ci wartość dokładnego zestawiania wielkości zwrotów dla każdego prowadzonego sondażu.

Jeśli uczestników sondażu nie można zidentyfikować na podstawie kwestionariusza, wysłanie przypomnienia tylko do tych, którzy nie wzięli udziału w badaniach, może okazać się niemożliwe. W takiej sytuacji wyślij przypomnienia do wszystkich członków próby, dziękując tym, którzy wzięli już udział w badaniach, i zachęcając do udziału tych, którzy jeszcze tego nie uczynili. (Zamieszczony dalej opis przypadku pokazuje jeszcze inną metodę, jakiej możesz użyć w anonimowej ankiecie pocztowej).

Wystarczający poziom zwrotów

Osoby stawiające pierwsze kroki w badaniach sondażowych często zadają pytanie o procentowy **poziom zwrotów**, który powinien zostać osiągnięty w ankiecie pocztowej. Teoria wnioskowania statystycznego zakłada w odniesieniu do analiz stosowanych w badaniach sondażowych, że wszyscy członkowie dobranej początkowo próby wypełniają i zwracają kwestionariusze. Jako że prawie nigdy się to nie zdarza, obciążenie związane z poziomem zwrotów staje się problemem. Badacze testują możliwość (mając najczęściej nadzieję na jej potwierdzenie), że respondenci są zasadniczo tacy sami jak próba losowa z pierwotnie dobranej próby, a zatem że są nieco mniejszą próbą z badanej populacji. (Jeśli chcesz zapoznać się z bardziej szczegółowym omówieniem obciążenia związanego z poziomem zwrotów, możesz przeczytać publikacje Donalda [1960] oraz Brownlee [1975]).

Tak czy inaczej, całkowity poziom zwrotów jest jakimś wskaźnikiem reprezentatywności respondentów w próbie. Gdy poziom ten jest wysoki, prawdopodobieństwo znaczących obciążeń z nim związanych jest mniejsze niż wtedy, gdy jest on niski. I odwrotnie, niski poziom zwrotów jest sygnałem ostrzegawczym,

gdyż osoby nieuczestniczące w badaniach mogą różnić się od respondentów nie tylko pod tym względem, że nie chcą brać w nich udziału. Na przykład Richard Bolstein (1991) wykazał, że ci, którzy nie wzięli udziału w sondażu przedwyborczym, również z mniejszym prawdopodobieństwem brali udział w wyborach niż uczestnicy sondażu. Zatem szacowanie frekwencji na podstawie wyników sondażu prowadziłoby do zawyżania liczby osób naprawdę później głosujących.

Ale co kryje się pod określeniem wysoki lub niski poziom zwrotów? Szybki przegląd literatury wykaże szeroki zakres poziomu zwrotów. Każdy z nich może być opatrzony stwierdzeniem w rodzaju: „Uzyskany poziom zwrotów można uznać za relatywnie wysoki dla tego rodzaju badań". (Pewien amerykański senator stwierdził to w odniesieniu do sondażu przeprowadzonego wśród wyborców w jego okręgu, w którym osiągnięto 4-procentowy poziom zwrotów). Mimo to można podać kilka opartych na intuicji reguł, dotyczących poziomu zwrotów. Jestem przekonany, że zwrotność na poziomie 50% jest wystarczająca do opracowywania analiz i raportów. Za dobrą należy uznać zwrotność 60-procentową, za bardzo dobrą – 70-procentową. Należy jednak mieć na względzie, że są to tylko luźne wskazówki; nie mają one podstaw statystycznych i znacznie ważniejsze jest wykazanie braku obciążeń, związanych z poziomem zwrotów, niż sam wysoki poziom zwrotów. Jeśli chcesz dalej zgłębiać te zagadnienia, Delbert Miller (1991, s. 145–155) dokonał przeglądu kilku konkretnych sondaży, pozwalającego na wyrobienie sobie lepszego poglądu na zmienność poziomu zwrotów.

Jak łatwo się domyślić, jeden z nieustających tematów debat wśród badaczy sondażowych dotyczy sposobów zwiększenia zwrotności ankiet. Jak pamiętamy, było to głównym problemem we wcześniejszym omówieniu różnych opcji wysyłki i zwrotnego otrzymywania kwestionariuszy. Badacze sondażowi opracowali wiele pomysłowych technik mających na celu zaradzenie temu problemowi. Niektórzy eksperymentowali z nowatorskimi formatami. Inni próbowali płacić respondentom za udział w badaniach. Oczywiście problem z płaceniem polega na tym, że dość drog wychodzi zapłacenie odpowiednio wysokie kwoty setkom czy tysiącom respondentów, a stosowano pewne pomysłowe alternatywne spo soby. Niektórzy badacze powiadali: „Chcemy Pana(i) uzyskać opinie o pewnych sprawach, wa te dwa centy, i chcemy za to zapłacić", załączali dwie monety jednocentowe. Inni załączali ćwier dolara, sugerując, by respondent uszczęśliwił je kieś dziecko. Jeszcze inni dołączali banknoty.

Don Dillman (1978) dokonał znakomitego przeglądu różnych technik stosowanych przez badaczy w celu zwiększenia zwrotności ankiet pocztowych, zarazem oceniając skuteczność każdej z nich. Co więcej, Dillman podkreśla konieczność zwracania uwagi raczej na wszystkie aspekty badań niż jeden czy dwa specjalne triki – podejście takie określa on jako „metodę całościowego projektowania" (*Total Design Method*).

W nowszym badaniu Francis Yammarino, Steven Skinner i Terry Childers (1991) dokonali pogłębionej analizy poziomu zwrotów uzyskanego w wielu badaniach, prowadzonych z zastosowaniem różnych technik. Wyniki są zbyt złożone, aby je krótko streścić, ale nieco wskazówek co do tego, jak zaplanować efektywny sondaż, możesz znaleźć w publikacji tych autorów.

Opis przypadku

Aby zrealizować sondaż, wysyłając ankiety pocztą, trzeba podjąć wiele kroków, które najłatwiej zrozumieć dzięki prześledzeniu rzeczywistych badań. Dlatego w zakończeniu niniejszego podrozdziału szczegółowo opisuję realizację sondażu wśród studentów, który omawialiśmy w rozdziale 7 jako przykład systematycznego doboru próby. Badania te nie stanowią teoretycznego ideału badań tego typu, ale najlepiej nadają się do naszych obecnych celów. Zostały one przeprowadzone przez studentów uczęszczających na moje seminarium z metod badań sondażowych.

Jak być może pamiętasz, z komputerowej bazy danych dobrano próbę 1100 studentów, posługując się procedurą warstwowego dobo-

u systematycznego. Dla każdego z wylosowanych studentów wydrukowano sześć samoprzylepnych naklejek adresowych.

W chwili, gdy byliśmy gotowi do dystrybucji kwestionariuszy, okazało się, że nasze mizerne fundusze badawcze nie wystarczą na pokrycie kilku wysyłek do całej próby 1100 studentów (koszty druku kwestionariuszy okazały się wyższe niż przewidywane początkowo). Dlatego dobraliśmy systematycznie dwie trzecie próby naklejek adresowych, uzyskując podpróbę 773 studentów.

Wcześniej zadecydowaliśmy, że nasze badania będą anonimowe – mieliśmy nadzieję, że w ten sposób nakłonimy respondentów do szczerszych odpowiedzi na niektóre delikatne pytania. (Późniejsze sondaże dotyczące tych samych kwestii przeprowadzone wśród tej samej populacji wykazały, że anonimowość była niepotrzebna). Niemożliwe było zatem określenie tożsamości studenta na podstawie kwestionariusza. Mieliśmy zarazem nadzieję, że uda nam się ograniczyć koszty listów przypominających poprzez wysłanie ich tylko do tych osób, które nie odpowiedziały.

Aby osiągnąć oba te cele, opracowano metodę specjalnych kart pocztowych. Do każdego studenta wysłano kwestionariusz bez żadnych znaków identyfikacyjnych oraz kartę pocztową zaadresowaną do agencji badawczej. Na odwrocie karty umieszczono jedną z naklejek adresowych z danymi studenta. W liście wyjaśniającym proszono studenta o wypełnienie i odesłanie kwestionariusza – z zapewnieniem anonimowości – oraz o jednoczesne odesłanie karty. Otrzymanie jej dawałoby nam informację o odesłaniu kwestionariusza przez danego studenta, bez określania, który to kwestionariusz. Procedura ta ułatwiałaby w ten sposób wysyłkę listów przypominających.

Kwestionariusz o objętości 32 stron został wydrukowany w formie broszurki. Trzyskrzydłowa okładka, opisana wcześniej w tym rozdziale, umożliwiała odesłanie kwestionariusza bez używania dodatkowej koperty.

Na przedniej okładce broszurki wydrukowany był list wyjaśniający cele badań. Tłumaczył on, dlaczego badania są przeprowadzane (aby poznać poglądy studentów związane z różnymi sprawami), w jaki sposób dobrano studentów do badań, dlaczego ważne jest uzyskanie odpowiedzi od każdego z wybranych studentów oraz w jaki sposób należy odesłać kwestionariusz.

Zapewniono studentów, że ich odpowiedzi pozostaną anonimowe, i wyjaśniono metodę kart pocztowych. Następnie określono, pod czyimi auspicjami prowadzone są badania i podano numer telefonu, pod którym można uzyskać więcej informacji. (Około pięciu studentów zwróciło się po dodatkowe informacje).

Drukując list na kwestionariuszu, uniknęliśmy konieczności załączania osobnego listu w tej samej kopercie, upraszczając w ten sposób zadanie złożenia całości przeznaczonej do wysyłki.

Materiały do pierwotnej wysyłki były skompletowane w następujący sposób: (1) na kartach pocztowych naklejono po jednej naklejce dla każdego studenta, (2) drugą naklejkę umieszczano na wysyłanej brązowej kopercie, (3) do każdej koperty zapakowano kartę pocztową i kwestionariusz, kontrolując zgodność nazwiska na karcie z nazwiskiem na kopercie.

Wysyłki pocztowej dokonano hurtowo. Po zapakowaniu kwestionariuszy w koperty, ułożono je według numerów kodowych, powiązano w paczki i zaniesiono na pocztę.

Wkrótce po pierwotnej wysyłce do agencji badawczej zaczęły nadchodzić kwestionariusze i karty pocztowe. Kwestionariusze otwierano, przeglądano i opatrywano numerem identyfikacyjnym zgodnie z opisanym wcześniej w tym rozdziale sposobem. Dla każdej otrzymanej karty pocztowej wyszukiwano pozostałe dla danego nazwiska naklejki i niszczono je.

Po upływie dwóch czy trzech tygodni wykorzystano pozostałe naklejki adresowe, organizując wysyłkę przypominającą. Tym razem przygotowano specjalne, osobne listy zachęcające. Nowy list informował, że wielu studentów zwróciło już kwestionariusze i że jest bardzo ważne, by pozostali też to uczynili.

Zgodnie z przewidywaniami, wysyłka przypominająca spowodowała wzrost zwrotów. Zastosowano te same procedury rejestracyjne. Odesłane karty pocztowe informowały nas o tym, które pozostałe jeszcze naklejki adreso-

we możemy zniszczyć. Niestety, ograniczenia czasowe i finansowe uniemożliwiły nam dokonanie planowanej początkowo trzeciej wysyłki, ale przy zastosowaniu dwóch wysyłek uzyskaliśmy 62% zwrotów.

Ten przykład powinien uzmysłowić ci, z czym w praktyce wiąże się zastosowanie wysyłanych pocztą kwestionariuszy wypełnianych przez respondentów. Przejdźmy teraz do drugiej głównej metody prowadzenia sondaży, jaką są wywiady osobiste.

Kwestionariusz wypełniany przez ankietera

Alternatywną wobec kwestionariuszy wypełnianych samodzielnie przez respondenta metodą zbierania danych w badaniach sondażowych jest **wywiad**. Badacze, zamiast prosić respondentów o zapoznanie się z kwestionariuszem i wpisanie swoich odpowiedzi, wysyłają ankieterów, którzy zadają pytania ustnie, a następnie zapisują odpowiedzi respondentów. Wywiady najczęściej realizuje się w postaci osobistych spotkań, ale większość reguł stosuje się także do wywiadów telefonicznych, omówionych w następnym podrozdziale.

Większość sondaży, przeprowadzanych za pomocą wywiadów, wymaga większej liczby ankieterów, jednak możesz też podjąć tego typu badania na małą skalę o własnych siłach. Poszczególne fragmenty tego podrozdziału dotyczą metod szkolenia i nadzorowania ankieterów, towarzyszących ci w realizacji sondażu.

Niniejszy podrozdział dotyczy konkretnie wywiadów w badaniach sondażowych. W rozdziale 10 omówione są słabiej ustrukturyzowane wywiady pogłębione, prowadzone często w jakościowych badaniach terenowych.

Rola ankietera w badaniach sondażowych

Wypełnianie kwestionariusza **przez ankietera**, a nie przez respondenta, ma **kilka zalet. Za**cznijmy od tego, że sondaże realizowane przez ankieterów z reguły mają wyższy odsetek odpowiedzi niż ankiety pocztowe. Właściwie zaplanowane i przeprowadzone badanie powinno uzyskać efektywność na poziomie co najmniej 80–85% pierwotnie zakładanej próby. (W sondażach finansowanych z funduszy federalnych często wymaga się osiągnięcia jednej z tych dwóch wartości). Najwidoczniej respondenci mają większe opory przed odtrąceniem ankietera stojącego na ich progu, niż przed wyrzuceniem kwestionariusza ankiety pocztowej.

Obecność ankietera na ogół zmniejsza również liczbę wskazań typu „nie wiem" i „brak odpowiedzi". Jeśli dla danych badań ważne jest zminimalizowanie liczby takich odpowiedzi, można przeszkolić ankietera tak, by dopytywał o odpowiedzi („Gdyby musiał(a) Pan(i) wskazać jedną z odpowiedzi, to jak Pan(i) myśli, która z nich byłaby najbliższa Pana(i) odczuciom?").

Ankieterzy mogą też służyć jako zabezpieczenie przed mylnym rozumieniem pytań. Jeśli respondent wyraźnie nie rozumie intencji pytania, albo rozumie ją źle, ankieter może udzielić stosownych wyjaśnień, które umożliwią uzyskanie właściwej odpowiedzi na dane pytanie. (Jak się wkrótce przekonamy, takie wyjaśnienia muszą być ściśle kontrolowane poprzez wymogi formalne).

Ankieter może też w końcu nie tylko zadawać pytania, ale również obserwować respondentów. Może na przykład odnotować, jakiej rasy jest respondent, jeśli ta kwestia zostanie uznana za zbyt delikatną, by o nią zapytać. Podobne obserwacje można zgromadzić na temat charakteru mieszkania, posiadania różnych dóbr, umiejętności posługiwania się przez respondenta językiem angielskim, jego ogólnej reakcji na badania itd. W jednym z badań prowadzonych wśród studentów, w trakcie wywiadu wręczono respondentom do samodzielnego wypełnienia krótki kwestionariusz, dotyczący postaw i zachowań seksualnych. W czasie gdy student wypełniał kwestionariusz, ankieter sporządzał szczegółowe notatki dotyczące ubrania i wyglądu respondenta.

Ta procedura wiąże się z pewną kwestią etyczną. Niektórzy badacze mieli obiekcje co do takich praktyk. Uważali, iż stanowią one po-

gwałcenie istoty zgody na udział w badaniach, jaką wyraził respondent. Chociaż kwestie etyczne przeważnie nie są jednoznacznie rozstrzygalne w badaniach społecznych, trzeba być na nie wrażliwym. Bardziej szczegółowo zajmiemy się nimi w rozdziale 18.

Badania sondażowe są z konieczności oparte na nierealistycznej teorii bodźca i reakcji w odniesieniu do poznania i zachowania. Badacze muszą zakładać, że pytanie kwestionariusza będzie oznaczać to samo dla każdego respondenta i że dana odpowiedź udzielana przez różnych respondentów będzie każdorazowo oznaczać to samo. Pytania w sondażu są zaprojektowane tak, by zbliżyć się do tego ideału tak bardzo, jak tylko jest możliwe, mimo że cel ten jest nieosiągalny.

Również ankieter musi dostosować się do tej idealnej sytuacji. Jego obecność nie powinna wpływać na to, jak respondent postrzega pytania czy udzielane odpowiedzi. Innymi słowy, ankieter powinien być neutralnym medium, poprzez które przekazywane są pytania i odpowiedzi.

Będąc takimi neutralnymi mediami, różni ankieterzy powinni otrzymywać dokładnie takie same odpowiedzi od danego respondenta. (Przypomnij sobie nasze wcześniejsze rozważania na temat rzetelności). Ta neutralność ma szczególne znaczenie w próbach opartych na jednostkach geograficznych. Otóż ze względu na oszczędność czasu i środków danemu ankieterowi powierza się najczęściej do przeprowadzenia wszystkie wywiady w jakimś określonym obszarze geograficznym – w kwartale miejskim czy w kilku sąsiadujących ze sobą kwartałach. Jeśli ankieter uczyni cokolwiek, co będzie wpływało na uzyskiwane przez niego odpowiedzi, wprowadzone w ten sposób obciążenie może być później interpretowane jako cecha charakterystyczna tego obszaru.

Załóżmy, że przeprowadzany jest sondaż badający postawy wobec taniego budownictwa mieszkaniowego w celu umożliwienia dokonania wyboru miejsca na nowe osiedle, finansowane ze środków publicznych. Ankieter wyznaczony dla danego obszaru geograficznego może – poprzez swoje słowa czy gesty – komunikować swój niesmak dla idei budowania tanich osiedli mieszkaniowych. Respondenci mogą mieć więc skłonność do udzielania odpowiedzi zgodnych z poglądami ankietera. Wyniki tego sondażu wskazywałyby, że w danym obszarze geograficznym istnieje silny opór przeciw budowie osiedla na tym terenie, podczas gdy tak naprawdę ten pozorny opór jest po prostu odbiciem postawy ankietera.

Ogólne zasady pracy ankieterskiej

Sposób, w jaki należy przeprowadzać wywiady, będzie różnił się nieco w zależności od badanej populacji i do pewnego stopnia będzie uwarunkowany tematem badań. Niemniej jednak pewne ogólne zasady dają się zastosować do większości sytuacji wywiadu.

Wygląd i zachowanie

Można przyjąć jako zasadę, że ankieterzy powinni być ubrani podobnie jak ludzie, z którymi będą przeprowadzać wywiady. Zamożnie ubrany ankieter może mieć trudności w uzyskaniu dobrej współpracy i odpowiedzi ze strony uboższych respondentów; niezamożnie ubrany ankieter może napotkać podobne trudności ze strony zamożniejszych respondentów. Jeśli już odzież i wygląd ankietera muszą się różnić od odzieży i wyglądu respondentów, ważne jest, aby skromne ubranie ankietera było czyste i schludne. Czystość jest chyba najbliższa neutralności, jeśli nie pobożności. Mimo że schludność i czystość w stylu klasy średniej mogą nie być akceptowane przez wszystkie sektory społeczeństwa amerykańskiego, pozostają one jednak podstawową normą i prawdopodobnie będą akceptowalne dla największej liczby respondentów.

Ubranie i wygląd są z reguły uważane za oznaki postaw i poglądów danej osoby. W czasie, gdy pisane są te słowa, podarte dżinsy, zielone włosy i kolczyki z żyletek mogą komunikować – słusznie czy niesłusznie – że ankieter jest radykalny politycznie, permisywny w sferze seksualnej, przychylny używaniu narkotyków itd. Każde z tych wrażeń może uwarunkować odpowiedzi albo wywrzeć wpływ na samą

chęć uczestniczenia w badaniach przez osoby, z którymi ma być przeprowadzony wywiad.

Co do zachowania, ankieterzy powinni być co najmniej mili. Ponieważ będą oni wtrącać się w prywatne życie i poglądy respondenta, muszą komunikować szczere zainteresowanie poznaniem respondenta, bez sprawiania wrażenia szpiegowania. Muszą być odprężeni i przyjaźnie nastawieni, bez zbytniej swobody ani namolności. Dobrzy ankieterzy mają także zdolność do bardzo szybkiego określania typu osoby, z jaką respondent będzie się czuł najswobodniej – typu osoby, z jaką byłoby mu najprzyjemniej rozmawiać. Oczywiście wywiad będzie skuteczniejszy wtedy, gdy ankieter może stać się takim typem osoby. Poza tym, ponieważ respondenci proszeni są o dobrowolne poświęcenie czasu i o ujawnienie informacji na tematy osobiste, zasługują oni na jak najprzyjemniejsze przeżycia dostarczane przez badacza i ankietera.

Znajomość kwestionariusza

Gdy ankieter nie zna kwestionariusza, cierpią na tym badania, a respondent niepotrzebnie musi znosić dodatkową uciążliwość. Wywiad prawdopodobnie zabierze więcej czasu, niż potrzeba i będzie nieprzyjemny. Co więcej, ankieter nie nabędzie znajomości kwestionariusza przez przejrzenie go dwa czy trzy razy. Musi on dokładnie go przestudiować, pytanie za pytaniem, musi też przećwiczyć czytanie go na głos.

Ostatecznie ankieter musi być w stanie odczytywać respondentowi pytania kwestionariusza bez błędów, bez potykania się o słowa czy wyrażenia. Dobrym wzorem jest aktor czytający kwestie w sztuce czy filmie. Muszą być one odczytywane w taki sposób, jakby stanowiły naturalną rozmowę, ale rozmowa ta musi posługiwać się dokładnie tym językiem, jaki jest zapisany w kwestionariuszu.

Na tej samej zasadzie ankieter musi zapoznać się z uwagami do kwestionariusza. Nie da się uniknąć sytuacji, w których pewne pytania nie będą pasowały idealnie do sytuacji danego respondenta i ankieter musi wówczas określić sposób interpretacji pytania. Uwagi dla ankietera powinny dawać odpowiednie wskazówki w takich przypadkach, ale ankieter musi znać układ i zawartość tych wskazówek wystarczająco dobrze, by móc z nich skutecznie korzystać. Lepiej, by ankieter opuścił dane pytanie, niż by spędzał pięć minut, szukając w uwagach wyjaśnienia lub próbując zinterpretować odnośne instrukcje.

Trzymanie się dokładnego brzmienia pytań

Pierwsza część tego rozdziału zajmowała się znaczeniem brzmienia pytania dla uzyskiwanych na nie odpowiedzi. Nieznaczna zmiana sformułowania danego pytania może prowadzić do wybierania przez respondenta odpowiedzi „tak" zamiast „nie". Dlatego ankieterzy powinni być poinstruowani, że mają trzymać się ściśle brzmienia pytań. Jeśli tego nie robią, zmarnowany zostaje wysiłek, jaki autor kwestionariusza włożył w staranne formułowanie pytań, w celu uzyskania potrzebnych informacji oraz zapewnienia, że respondenci będą interpretować pytania zgodnie z jego intencjami.

Dokładne zapisywanie odpowiedzi

Zawsze gdy kwestionariusz zawiera pytania otwarte, a więc takie, które wymagają od respondenta sformułowania własnych odpowiedzi, ankieter musi zapisywać te odpowiedzi dokładnie tak, jak je sformułował respondent. Nie wolno ich streszczać, parafrazować ani poprawiać błędów gramatycznych.

Ta dokładność jest szczególnie ważna dlatego, że ankieter nie wie, jak będą kodowane odpowiedzi. Tak naprawdę sami badacze najczęściej tego nie wiedzą, dopóki nie przeczytają jakichś stu odpowiedzi. Na przykład kwestionariusz może zawierać pytanie o to, co respondenci sądzą o sytuacji komunikacyjnej w swojej miejscowości. Jakiś respondent może odpowiedzieć, że na drogach jest za dużo samochodów i że powinno się coś zrobić, by ograniczyć ich liczbę. Inny może być zdania, że potrzeba więcej dróg. Jeśli ankieter zapisze te obie odpowiedzi, streszczając je w podobny sposób – jako „zatłoczone drogi" – badacze nie będą mogli wykorzystać w interpretacji istotnych różnic między oryginalnymi odpowiedziami.

Wypowiedzi słowne są niekiedy zbyt niewyraźne lub wieloznaczne, by można je było zinterpretować. Jednak ankieter może zrozumieć intencję odpowiedzi poprzez gesty czy ton głosu respondenta. W takiej sytuacji powinien on mimo to zapisać dokładną wypowiedź słowną, ale powinien także dodać na marginesie zarówno uwagi na temat jego własnej interpretacji odpowiedzi, jak i uzasadnienie tej interpretacji.

Ujmując to ogólniej, ankieterzy mogą stosować uwagi na marginesie dla wyjaśnienia różnych aspektów odpowiedzi, które nie są widoczne w samym zapisie słownym, takich jak widoczna złość respondenta, zakłopotanie, niepewność przy udzielaniu odpowiedzi itd. Jednak zawsze należy zapisywać również odpowiedź słowną w jej dokładnym brzmieniu.

Dopytywanie

Zdarza się, że respondenci w trakcie wywiadu udzielają odpowiedzi nieadekwatnych lub niekompletnych. W sytuacjach takich przydatne może być **dopytywanie**, czyli prośba o dokładniejsze wyjaśnienie. Na przykład pytanie zamknięte może zawierać stwierdzenie odnoszące się do jakiejś postawy, wraz z propozycją odpowiedzi „zdecydowanie się zgadam", „raczej się zgadzam", „raczej się nie zgadzam" i „zdecydowanie się nie zgadzam". Jednak respondent może odpowiedzieć: „Myślę, że to prawda". Ankieter powinien na taką odpowiedź zapytać: „Powiedziałby pan, że się pan zdecydowanie zgadza, czy też raczej zgadza?". W razie potrzeby ankieter może wyjaśnić, że musi zakreślić którąś z podanych w kwestionariuszu odpowiedzi. Jeśli respondent uparcie odmawia dokonania wyboru, ankieter powinien zapisać dosłownie odpowiedź udzieloną przez respondenta.

Dopytywanie częściej stosuje się do uzyskiwania odpowiedzi na pytania otwarte. Na przykład na pytanie o warunki komunikacyjne respondent mógłby po prostu odpowiedzieć „fatalne". Ankieter powinien uzyskać dokładniejsze wyjaśnienie tej odpowiedzi poprzez różne dodatkowe pytania. Niekiedy najlepszym dopytaniem jest cisza; gdy ankieter siedzi cicho z uniesionym ołówkiem, respondent prawdopodobnie wypełni tę pauzę dodatkowymi komentarzami. (Technika ta jest z powodzeniem stosowana przez reporterów prasowych). Odpowiednie zachęty słowne mogą brzmieć: „Jak to się dzieje?" lub „W jaki sposób?". Chyba najbardziej uniwersalne dopytanie to: „Czy coś jeszcze?".

Ankieterzy często muszą dopytywać aż do uzyskania odpowiedzi zawierających wystarczającą do celów analitycznych ilość informacji. Jednak takie dopytywanie *musi* być zawsze całkowicie neutralne; w żaden sposób nie wolno dopuścić do tego, by wpływało ono na charakter następującej po nim odpowiedzi. Zawsze gdy można przewidzieć, że dane pytanie może wymagać dopytywania w celu uzyskania adekwatnej odpowiedzi, powinno się umieścić w kwestionariuszu po tym pytaniu jedno lub więcej praktycznych sformułowań, służących do dopytywania. Działanie takie ma dwie ważne zalety. Po pierwsze, ma się więcej czasu na opracowanie najlepszych, najbardziej neutralnych sformułowań dopytujących. Po drugie, wszyscy ankieterzy zastosują w razie potrzeby takie same sformułowania. Tak więc nawet jeśli dopytanie nie jest absolutnie neutralne, wszyscy respondenci zostaną poddani działaniu tego samego bodźca. Jest to ta sama wskazówka logiczna, o której była mowa przy formułowaniu brzmienia pytania. Chociaż pytanie nie powinno być naładowane wartościująco albo obciążone, najważniejsze jest, by każdemu respondentowi zadano je w identycznym brzmieniu – nawet jeśli nie jest ono wolne od obciążeń.

Koordynacja i kontrola

Większość badań sondażowych wymaga współpracy wielu ankieterów. W sondażach prowadzonych na dużą skalę wynajmuje się płatnych ankieterów. Badacze będący studentami mogą sami zrekrutować kolegów do pomocy przy wywiadach. Zawsze gdy angażuje się do wywiadów więcej niż jednego ankietera, praca ankieterów musi być starannie kontrolowana. Ta kontrola ma dwa aspekty: szkolenie ankieterów oraz nadzór nad nimi po rozpoczęciu przez nich pracy.

Sesja szkoleniowa dla ankieterów powinna zacząć się od opisu tematu badań. Mimo że ankieterzy mogą być zaangażowani jedynie w fazie zbierania danych, warto, by rozumieli, co stanie się dalej z przeprowadzonymi przez nich wywiadami i jakim celom to służy. Gdy ankieterzy nie wiedzą, o co chodzi, na ogół niższe jest ich morale i motywacja.

Część szkolenia dotycząca samego ankietowania powinna rozpocząć się od omówienia ogólnych wskazówek i procedur, takich jak te, którymi zajmowaliśmy się wcześniej w tym podrozdziale. Następnie cała grupa powinna razem przejść przez kwestionariusz – pytanie po pytaniu. Nie pytaj po prostu, czy ktoś ma jakieś pytania do pierwszej strony kwestionariusza. Przeczytaj na głos pierwsze pytanie, wyjaśnij jego cel, a następnie poświęć uwagę pytaniom i komentarzom, które ankieterzy być może zgłoszą. Po wyjaśnieniu wszystkich zgłoszonych pytań i komentarzy, przejdź do następnego pytania kwestionariusza.

Dobrym pomysłem jest przygotowanie uwag towarzyszących kwestionariuszowi. *Uwagi* mają charakter komentarzy wyjaśniających, jak radzić sobie w trudnych lub niejednoznacznych sytuacjach, jakie mogą wystąpić w związku z konkretnymi pytaniami kwestionariusza. Układając kwestionariusz, spróbuj wyobrazić sobie wszystkie problematyczne przypadki, jakie mogą się pojawić – nietypowe okoliczności, które mogą spowodować, że trudno będzie udzielić odpowiedzi na pytanie. Uwagi do sondażu powinny dawać szczegółowe wskazówki, jak radzić sobie z takimi sytuacjami. Na przykład problemów może przysporzyć nawet tak prosta sprawa jak wiek. Przypuśćmy, że respondent mówi, że w przyszłym tygodniu skończy 25 lat. Ankieter może nie być pewien, czy przyjąć obecny wiek respondenta, czy wiek, jaki uzyska w najbliższym czasie. Uwagi do tego pytania powinny określać, co należy zrobić. (Zapewne uściśliłbyś, że należy odnotować wiek osiągnięty przy okazji ostatnich urodzin).

Po przygotowaniu zestawu uwag przejrzyj go z ankieterami, gdy analizujecie pojedyncze pytania kwestionariusza. Upewnij się, że ankieterzy w pełni rozumieją uwagi i przyczyny, dla których one powstały, jak również same pytania.

Jest bardzo możliwe, że ta część szkolenia wywoła wiele kłopotliwych pytań ze strony ankieterów. Będą oni pytać: „Co powinienem zrobić, gdy...?". W takich sytuacjach unikaj szybkich, nieprzemyślanych odpowiedzi. Jeśli masz przygotowane uwagi do kwestionariusza, pokaż, w jaki sposób można na ich podstawie określić rozwiązanie tego problemu. Jeśli nie masz uwag, pokaż, jak preferowane rozwiązanie tego problemu pasuje do ogólnej logiki pytania i do celów badań. Udzielanie na takie pytania ze strony ankieterów odpowiedzi pozbawionych elementu wyjaśnienia spowoduje tylko zakłopotanie wśród ankieterów i sprawi, że będą oni traktować swoją pracę mniej poważnie. Jeśli nie potrafisz odpowiedzieć na takie pytanie, gdy zostanie ono zadane, przyjmij je i poproś o nieco czasu na podjęcie decyzji. Potem przemyśl starannie sprawę i upewnij się, że twoja odpowiedź wraz z uzasadnieniem dotarła do wszystkich ankieterów.

Gdy już przeszliście przez cały kwestionariusz, przeprowadź jeden czy dwa wywiady pokazowe przed wszystkimi. Najlepiej, gdybyś przeprowadził wywiad z kimś spoza grona ankieterów. Bądź świadom tego, że przeprowadzany przez ciebie wywiad będzie służył jako model dla szkolonych osób, więc zrób to dobrze. Ponadto najlepiej będzie, jeśli wywiad pokazowy będzie przeprowadzony w możliwie realistyczny sposób. Nie zatrzymuj się podczas pokazu, żeby wskazać, jak poradziłeś sobie ze skomplikowaną sytuacją: poradź sobie z nią, a wyjaśnij potem. Nie ma znaczenia, czy osoba, z którą przeprowadzasz wywiad, udziela prawdziwych odpowiedzi, czy też przybiera jakąś hipotetyczną tożsamość dla tego celu, jeśli tylko jej odpowiedzi są spójne.

Po wywiadzie pokazowym poleć ankieterom, by dobrali się w pary i przeprowadzili wywiady ze sobą nawzajem. Gdy skończą, zaproponuj, by odwrócili role i przećwiczyli to jeszcze raz. Przeprowadzanie wywiadów jest najlepszą szkołą przeprowadzania wywiadów. W czasie gdy ankieterzy ćwiczą, przechadzaj się wśród nich, wsłuchując się, jak to robią, aby przekonać się, jak im to wychodzi. Gdy ćwiczenia się zakończą, cała grupa powinna wspólnie przeanalizować swoje doświadczenia i zadać ewentualne pytania.

Końcowy etap szkolenia ankieterów powinien obejmować jakieś „prawdziwe" wywiady. Niech ankieterzy przeprowadzą parę wywiadów w rzeczywistych warunkach, jakie będą obowiązywać przy ostatecznym sondażu. Możesz wyznaczyć im respondentów lub pozwolić im na ich samodzielne dobranie. Nie dopuszczaj jednak, by ćwiczono na osobach, które dobrałeś do próby. Gdy każdy z ankieterów zrealizuje trzy do pięciu wywiadów, spotkaj się z nimi ponownie. Sprawdź, czy w wypełnionych kwestionariuszach nie występują ślady błędnego zrozumienia pytań. Jeszcze raz odpowiedz na wszelkie ewentualne pytania ankieterów. Gdy już nabierzesz pewności, że dany ankieter wie, co ma robić, przydziel mu kilka prawdziwych wywiadów, korzystając z próby dobranej do badań.

Jest sprawą zasadniczą, by praca ankieterów była nadzorowana przez cały czas trwania badań. Powinieneś odebrać od nich ich dotychczasową pracę po przeprowadzeniu przez nich nie więcej niż 20 czy 30 wywiadów. Możesz przydzielić 20 wywiadów, poczekać aż ankieter odda z powrotem wypełnione kwestionariusze, przejrzeć je i przydzielić na przykład następne 20. Choć może to wyglądać na nadmierną ostrożność, musisz stale zabezpieczać się przed nieporozumieniami, które mogą nie być oczywiste na wczesnym etapie badań.

Jeśli jesteś jedynym ankieterem, może wydawać się, że te uwagi cię nie dotyczą. Jednak byłoby mądrze przygotować na przykład uwagi do potencjalnie kłopotliwych pytań w twoim kwestionariuszu. Jeśli tego nie zrobisz, narażasz się na ryzyko podejmowania w czasie badań decyzji *ad hoc*, których potem możesz żałować lub nie pamiętać. Również nacisk na ćwiczenie dotyczy w takim samym stopniu zarówno projektów jednoosobowych, jak i złożonych sondaży na zlecenie, realizowanych przez znaczną liczbę ankieterów.

Sondaże telefoniczne

Przez lata sondaże telefoniczne cieszyły się raczej złą reputacją wśród profesjonalnych badaczy. Sondaże telefoniczne z definicji ograniczone są do osób, które mają telefony. Wiele lat temu metoda ta prowadziła do zasadniczych obciążeń związanych z przynależnością do klasy społecznej poprzez wyłączenie z takich sondaży ludzi ubogich. Jaskrawą tego demonstracją było fiasko pisma „Literary Digest" w 1936 roku. Jak sobie przypominasz, nawet mimo że kontaktowano się z wyborcami drogą pocztową, próba była częściowo dobrana spośród abonentów telefonicznych, którzy nie byli typowi dla społeczeństwa podnoszącego się właśnie z Wielkiego Kryzysu. Jednak Biuro Spisowe (1996a, tabela 1224) oceniło, że do 1993 roku 93,4% wszystkich gospodarstw domowych posiadało telefony, tak więc istniejąca wcześniej forma obciążenia klasowego została istotnie zmniejszona.

Problem dotyczący próby stanowią też telefony nieuwzględnione w spisach. Próba do badań sondażowych dobrana ze stronic lokalnej książki telefonicznej pomijałaby całkowicie wszystkich takich ludzi – z reguły zamożniejszych – którzy zażądali, by nie publikowano ich numerów telefonów. To ewentualne obciążenie próby zostało zniwelowane dzięki technice, która znacząco ulepszyła dobór próby w sondażach telefonicznych – losowemu wybieraniu numerów.

Sondaże telefoniczne mają wiele zalet, leżących u podłoża rosnącej popularności tej metody. Zapewne największe z nich to pieniądze i czas, w takiej właśnie kolejności. W wywiadach bezpośrednich, prowadzonych w gospodarstwach domowych, trzeba jechać czasami wiele kilometrów do domu respondenta, aby następnie nikogo tam nie zastać, wrócić do agencji badawczej i następnego dnia pojechać tam znowu – być może znowu nikogo nie zastając. Jest taniej i szybciej, gdy wycieczki te wykonywane są za pomocą palców.

Prowadząc wywiad przez telefon, można ubrać się w dowolny sposób, nie wpływając przez to na odpowiedzi udzielane przez respondenta. A respondenci będą niekiedy uczciwiej podawali odpowiedzi nieakceptowane społecznie, jeśli nie będą musieli patrzeć ankieterowi w oczy. Na tej samej zasadzie może okazać się możliwe dopytywanie w bardziej drażliwych kwestiach, choć tak być nie musi. W pewnej mierze ludzie są bardziej nieufni, kiedy nie widzą osoby zadającej pytania – jest to być może konsekwencja „son-

daży" nastawionych na sprzedaż subskrypcji czasopism i apartamentów typu „time-share".

Mimo wszystko ankieterzy mogą zakomunikować przez telefon całkiem sporo na swój temat, mimo że są niewidoczni. Na przykład badacze przywiązują wagę do wpływu, jaki może wywrzeć nazwisko ankietera (zwłaszcza jeśli temat badań związany jest z kwestiami etnicznymi) i zastanawiają się, na ile etyczne byłoby, gdyby wszyscy ankieterzy używali nijakich „pseudonimów", takich jak Smith czy Jones. (Niekiedy ankieterki proszą o zezwolenie na takie praktyki, chcąc uniknąć późniejszego napastowania przez mężczyzn, z którymi przeprowadzają wywiady).

Wywiady telefoniczne mogą dawać większą kontrolę nad zbieraniem danych, gdy zatrudnia się wielu ankieterów. Jeśli wszyscy oni dzwonią z siedziby agencji badawczej, mogą uzyskać wyjaśnienia od osoby nadzorującej, gdy wynikną jakieś problemy, a to zazwyczaj jest nieuniknione. Ankieter pracujący sam na zabitej deskami wsi może być zmuszony do odkładania tego na cotygodniowe spotkania z koordynatorem.

W końcu jeszcze jeden ważny czynnik, związany z rosnącym stosowaniem sondaży telefonicznych, wiąże się z bezpieczeństwem osobistym. Don Dillman (1978, s. 4) opisuje tę sytuację w następujący sposób:

> „Ankieterzy muszą być w stanie swobodnie działać w atmosferze, w której na obcych patrzy się nieufnie i muszą skutecznie odpierać obiekcje respondentów co do udziału w wywiadzie. Coraz częściej ankieterzy muszą być gotowi pracować nocą, by móc skontaktować się z mieszkańcami wielu gospodarstw domowych. W niektórych przypadkach powoduje to konieczność zapewnienia ochrony ankieterom, pracującym w takich częściach miasta, w których istnieje wyraźne zagrożenie bezpieczeństwa jednostki".

Troska o bezpieczeństwo działa zatem hamująco na wywiady bezpośrednie na dwa sposoby. Ewentualni respondenci mogą odmówić udziału w badaniach, obawiając się obcego ankietera. Z kolei sami ankieterzy też mogą być narażeni na pewne niebezpieczeństwa. To wszystko jest jeszcze pogarszane przez możliwość narażenia badacza na poniesienie wielkich kosztów, jeśli pojawią się jakieś problemy.

Jednak badania telefoniczne są również związane z problemami. Jak już wspominałem, rozwój tej metody jest krępowany mnożącymi się przypadkami pozornych „sondaży", które w rzeczywistości są akcjami handlowymi prowadzonymi pod pozorem badań. Nawiasem mówiąc, jeśli masz jakieś pytania związane z otrzymywanymi przez ciebie tego rodzaju telefonami, zapytaj wprost ankietera, czy zostałeś wybrany tylko do udziału w sondażu, czy też w grę wchodzi jakaś „okazja" zakupu. W razie jakichkolwiek wątpliwości dobrym pomysłem jest też uzyskanie od ankietera jego nazwiska, numeru telefonu i nazwy firmy. Odłóż słuchawkę, jeśli osoba dzwoniąca odmawia podania którejkolwiek z tych informacji.

Dla badacza owa łatwość, z jaką ludzie mogą odłożyć słuchawkę, jest kolejną wadą sondaży telefonicznych. Jest mało prawdopodobne, by respondent nakazał opuścić dom w połowie wywiadu ankieterowi, który już został wpuszczony do środka. Znacznie łatwiej jest przerwać nagle wywiad telefoniczny, mówiąc coś w rodzaju: „Ojej! Ktoś puka. Muszę iść" albo: „O MÓJ BOŻE! Świnie żżerają moje Volvo!". (Tego rodzaju rzecz jest znacznie trudniejsza do upozorowania, gdy ankieter siedzi u kogoś w salonie).

Inny potencjalny problem z wywiadami telefonicznymi to rozpowszechnione stosowanie automatycznych sekretarek.

Badania przeprowadzone przez Walker Research (1988) wykazały, że połowa właścicieli automatycznych sekretarek przyznawała się do używania ich do „przesiewania" przychodzących rozmów przynajmniej czasami. Z kolei badania Tuckela i Feinberga (1991) wykazały, że automatyczne sekretarki nie mają jeszcze istotnego wpływu na możność skontaktowania się ankieterów z potencjalnymi respondentami. Niemniej jednak badacze ci zakończyli konkluzją, że w miarę jak automatyczne sekretarki rozpowszechniają się coraz bardziej, „cechy społeczno-demograficzne ich właścicieli będą się zmieniać". Ten fakt może prawdopodobnie doprowadzić do powstania „odmiennych wzorców zachowań, związanych z korzystaniem z automatycznych sekretarek" (1991, s. 216).

Voice Capture™

James E. Dannemiller

SMS Research, Honolulu

Rozwój różnych technik CATI jest dobrodziejstwem dla badań sondażowych i marketingowych, choć służy on głównie wspomaganiu zbierania, kodowania i analizy „zwykłych danych". Technika Voice Capture™ opracowana przez firmę Syrvey Systems oferuje jednak zupełnie niezwykłe możliwości, które dopiero zaczynamy poznawać.

W trakcie wywiadu CATI ankieter może włączyć w komputerze opcję cyfrowego nagrywania rozmowy z respondentem. Na przykład po ustaleniu faktu, że respondent zmienił ostatnio swój ulubiony program informacyjny w telewizji, ankieter może zapytać: „Dlaczego pan go zmienił?" i może rozpocząć nagrywanie dosłownej wypowiedzi. (Na początku wywiadu ankieter prosi o pozwolenie na nagrywanie fragmentów wywiadu).

Później koderzy mogą odtworzyć odpowiedzi i zakodować je – tak jak zrobiliby to z odpowiedziami zapisanymi przez ankietera. Jest to łatwiejszy i dokładniejszy sposób wykonania tradycyjnego zadania. Ale to jest banalne zastosowanie tej nowej techniki.

Jest też możliwe włączenie takich ustnych danych podczas analizy jako części tabel. Możemy stworzyć tabelę ilustrującą rozkład płci przez wiek przez powody przełączania się na inne telewizyjne programy informacyjne. Czyli możemy wysłuchać po kolei odpowiedzi młodych mężczyzn, młodych kobiet, mężczyzn w średnim wieku itd. W pewnych badaniach tego rodzaju stwierdziliśmy, że młodsi i starsi mężczyźni na ogół oglądają jeden program informacyjny, natomiast mężczyźni w średnim wieku na ogół oglądają coś jeszcze. Wysłuchując odpowiedzi mężczyzn w średnim wieku, jednej za drugą, słyszeliśmy często uwagę: „Cóż, teraz, gdy jestem starszy...". Ten rodzaj dygresji mógłby zostać pominięty w notatkach sporządzanych w pośpiechu przez ankieterów, ale komentarze te wyróżniały się dobitnie w ustnych danych. Wydawało się, że mężczyźni w średnim wieku mówią nam, że uważają, iż „dojrzałość" wymaga, by oglądali oni jakiś określony program, bardziej zaawansowany wiek zaś pozwala im cofnąć się do tego, co najbardziej lubią.

Tego rodzaju dane są szczególnie atrakcyjne dla zleceniodawców, zwłaszcza w badaniach satysfakcji klienta. Zamiast podsumowywać, co naszym zdaniem podoba się, a co nie podoba klientom naszego zleceniodawcy, możemy pozwolić, by przemówili oni do niego bezpośrednio swoimi własnymi słowami. To jest jak zogniskowany wywiad grupowy na żądanie. Idąc krok do przodu, stwierdziliśmy, że wysłuchanie takich odpowiedzi przez pracowników liniowych (na przykład kasjerów bankowych) miało na nich większy wpływ, niż gdyby ich przełożeni mówili im, co robią dobrze, a co źle.

Choć te doświadczenia są same w sobie ekscytujące, jestem przekonany, że dopiero zaczęliśmy wypróbowywać możliwości stosowania takich niekonwencjonalnych form danych.

Wywiad telefoniczny wspomagany komputerowo (CATI)

W rozdziale 14 przyjrzymy się niektórym aspektom wpływu, jaki wywarły komputery na prowadzenie badań społecznych – zwłaszcza przetwarzanie danych i analizę. Komputery zmieniają też charakter wywiadów telefonicznych. Jedną z innowacji jest wywiad telefoniczny wspomagany komputerowo (CATI, *computer-assisted telephone interviewing*). Metoda ta jest w coraz większym stopniu wykorzystywana w badaniach sondażowych o charakterze akademickim, rządowym czy komercyjnym. Chociaż w praktyce istnieją jej różne odmiany, pokażemy tutaj, jak CATI może wyglądać.

Wyobraź sobie ankietera ze słuchawkami od telefonu na uszach, siedzącego przed terminalem komputera i jego ekranem. Komputer centralny został tak zaprogramowany, by losował numery telefonów i dzwonił pod nie. (Losowe generowanie numerów telefonicznych pozwala uniknąć problemów związanych z numerami telefonów nie umieszczonymi w książ-

kach telefonicznych). Na ekranie komputera widnieje wprowadzenie („Dzień dobry, nazywam się...") i pierwsze pytanie, które ma być zadane („Czy może mi Pan(i) powiedzieć, ile osób mieszka pod tym adresem?").

Gdy respondent odbierze telefon, ankieter mówi „dzień dobry", przedstawia krótko cel badań i zadaje pierwsze pytanie wyświetlone na ekranie. Po udzieleniu przez respondenta odpowiedzi, wpisuje ją do terminala komputera – albo dosłowną odpowiedź na pytanie otwarte, albo też kod kategorii dla właściwej odpowiedzi na pytanie zamknięte. Odpowiedź jest natychmiast zapisywana w pamięci komputera. Na ekranie pojawia się drugie pytanie, ankieter zadaje je i wprowadza odpowiedź do komputera. W ten sposób wywiad posuwa się naprzód.

Oprócz oczywistych korzyści w zakresie gromadzenia danych, CATI automatycznie przygotowuje dane do analizy; tak naprawdę badacz może zacząć analizę danych jeszcze przed ukończeniem całości wywiadów, uzyskując w ten sposób szybszy wgląd w wyniki. Jeszcze inna innowacja, umożliwiona przez technikę komputerową, opisana jest w ramce zatytułowanej „Voice Capture™".

Nowe technologie a badania sondażowe

Jak widzieliśmy na przykładzie wywiadów telefonicznych wspomaganych komputerowo (CATI), wiele spośród nowych technologii, które wpływają na ludzkie życie, otwiera również nowe możliwości w dziedzinie badań sondażowych. Na przykład najnowsze innowacje w ankietach wypełnianych samodzielnie przez respondenta dotyczą wykorzystania komputerów. Testowaniu poddano m.in. poniższe techniki.

CAPI – wywiad osobisty wspomagany komputerowo (*computer assisted personal interviewing*): podobny do CATI, ale stosowany w wywiadach bezpośrednich, a nie telefonicznych.

CASI – ankieta samodzielnie wypełniana przez respondenta, wspomagana komputerowo (*computer assisted self-interviewing*): pracownik badawczy przychodzi do domu respondenta z komputerem, respondent zaś odczytuje pytania na ekranie komputera i wprowadza swoje odpowiedzi.

CSAQ – skomputeryzowany kwestionariusz do samodzielnego wypełniania przez respondenta (*computerized self-administered questionnaire*): respondent otrzymuje kwestionariusz na dyskietce, elektronicznej tablicy ogłoszeń lub w inny sposób i uruchamia program, który zadaje pytania i przyjmuje odpowiedzi. Następnie respondent zwraca plik z danymi.

TDE – wprowadzanie danych za pomocą wybierania tonowego (*touchtone data entry*): respondent inicjuje proces, dzwoniąc do organizacji badawczej. Zadaje mu ona serię skomputeryzowanych pytań, na które respondent odpowiada wciskając klawisze na klawiaturze telefonu.

VR – rozpoznawanie głosu (*voice recognition*): zamiast prosić respondenta o używanie klawiatury telefonu, jak ma to miejsce w *TDE*, system akceptuje odpowiedzi mówione.

Nicholls ze współpracownikami donoszą, że techniki te są bardziej efektywne od tradycyjnych i, jak się wydaje, nie pociągają za sobą strat w jakości danych.

Jeffery Walker (1994) badał możliwość prowadzenia sondaży za pomocą faksu. Kwestionariusze wysyła się faksem do respondentów i prosi się ich o przefaksowanie ich z powrotem po wypełnieniu. Takie sondaże mogą oczywiście być reprezentatywne tylko dla tej części populacji, która dysponuje faksem. Walker donosi, że sondaże przy użyciu faksu nie osiągają tak wysokiego odsetka odpowiedzi jak wywiady bezpośrednie, ale ze względu na to, że są postrzegane jako pilne, poziom odpowiedzi uzyskiwany w nich jest wyższy niż w ankietach pocztowych lub sondażach telefonicznych. W jednym z testowanych przypadków do wszystkich respondentów, którzy zignorowali ankietę pocztową, wysłano ponownie kwestionariusz faksem i 83% z nich przysłało odpowiedzi.

Jak już wspominałem wcześniej, należy wystrzegać się „sondaży", których celem jest najwyraźniej zbieranie pieniędzy. Praktyka ta wtargnęła już na obszar „sondaży faksowych", o czym może świadczyć faks, zatytułowany: „Czy należy zdelegalizować rewolwery?". W celu

umożliwienia wyrażenia opinii „tak" lub „nie" podano dwa numery faksu. Małym drukiem zamieszczono uwagę: „Koszt połączenia wynosi 2,95 USD za minutę, co nie jest wygórowaną ceną za większą demokrację. Połączenie trwa średnio od 1 do 2 minut". Można się domyślać, gdzie trafiało owe 2,95 USD.

Nowe technologie w badaniach sondażowych obejmują również korzystanie z Internetu i sieci www – dwóch wynalazków o największym zasięgu pod koniec XX wieku. Załącznik B zajmuje się wykorzystaniem sieci do poszukiwania literatury i innych związanych z tym aspektów procesu badawczego. Część badaczy jest zdania, że Internet może być również wykorzystywany do prowadzenia poważnych badań sondażowych.

Podstawowe zastrzeżenie, zgłaszane przez wielu badaczy społecznych pod adresem sondaży „*online*", dotyczy ich reprezentatywności: czy ludzie poddani badaniom sondażowym „online" mogą być reprezentacją jakiejś sensownej populacji, takiej jak dorośli Amerykanie, wszyscy wyborcy itd.? Jest to ten sam rodzaj krytyki, jaki zgłaszano pod adresem sondaży z wykorzystaniem faksu, a wcześniej – sondaży telefonicznych.

Camilo Wilson (1999), twórca Cogixa (www.cogix.com) zwraca uwagę na fakt, że niektóre populacje pasują idealnie do sondaży „online": a mianowicie te, które odwiedzają daną stronę www. Wilson podaje przykład badań rynku prowadzonych dla firm operujących w sieci, które *powinny* być przeprowadzane poprzez sieć, jego firma zaś opracowała program *Views-Flash*, służący właśnie temu celowi. Mimo że w sondażach sieciowych łatwo jest zebrać dane od wszystkich odwiedzających daną stronę, techniki doboru próby mogą zdaniem Wilsona zapewnić wystarczająco wiele danych o użytkownikach, bez konieczności irytowania tysięcy czy milionów potencjalnych użytkowników.

A co z sondażami na całej populacji? Jak wspomniałem, w środowisku badawczym zaczęła się dyskusja: Humphrey Taylor i George Terhanian (1999, s. 20) wywołali część tej dyskusji artykułem „Znów mamy oszałamiające czasy" („Heady Days Are Here Again"). Uznając konieczność zachowania ostrożności, nalegali oni zarazem na sprawiedliwe traktowanie sondaży „online":

> „Jeden z testów wiarygodności każdej nowej metody zbierania danych zależy od jej zdolności do rzetelnego i trafnego przewidywania zachowań wyborczych. Było to powodem, dla którego ostatniej jesieni próbowaliśmy oszacować wyniki wyborów na gubernatora i do Senatu USA w 1998 roku w 14 stanach, z wykorzystaniem czterech różnych sondaży przeprowadzanych przez Internet".

Badacze porównali swoje wyniki z 52 sondażami telefonicznymi dotyczącymi tych samych wyborów. W sondażu przeprowadzonym przez Internet trafnie wytypowano 21 z 22 zwycięzców, czyli 95%. Jednak samo wytypowanie zwycięzców nie jest wystarczającym testem na efektywność: na ile wyniki sondażu były zbliżone do faktycznej procentowej liczby głosów, uzyskanej przez poszczególnych kandydatów? Taylor i Terhanian podają, że ich sondaże „online" odbiegały od faktycznych wyników średnio o 6,8%. 52 sondaże telefoniczne odbiegały od tych samych wyników średnio o 6,2%.

Warren Mitofsky (1999) jest krytykiem sondaży „online". Poza faktem, że podważył on sposób, w jaki Taylor i Terhanian obliczyli podany przez siebie zakres błędu, wezwał on również do stworzenia solidniejszych podstaw teoretycznych dla ugruntowania nowej techniki.

Jednym z kluczowych zagadnień związanych z sondażami przez Internet jest właściwe określenie i zastosowanie wag dla różnych rodzajów respondentów, o czym mówiliśmy w kontekście doboru kwotowego w rozdziale 7. Taylor i Terhanian są świadomi ograniczeń próby kwotowej, ale ich pierwsze doświadczenia z sondażami „online" sugerują ich zdaniem, że powinno się stosować tę technikę. W rzeczy samej, kończą swe rozważania konkluzją: „Jest to rozpędzony pociąg, którego nie można zatrzymać, i przyspiesza on coraz bardziej. Ci, którzy nie wsiadają do niego, ryzykują, że pozostaną daleko w tyle" (1999, s. 23).

Uwagi zgłaszane obecnie w związku z sondażami „online" są podobne do uwag w stosunku do sondaży telefonicznych, zamieszczo-

nych w pierwszym wydaniu tej książki w 1975 roku. To, czy sondaże przez Internet będą cieszyć się w przyszłości podobnym uznaniem i szerokim zastosowaniem jak obecnie sondaże telefoniczne, jest kwestią otwartą. Studenci rozważający możliwość zastosowania tej techniki powinni mieć pełną świadomość jej ewentualnych ograniczeń.

Porównanie różnych metod badań sondażowych

Skoro już poznaliśmy kilka sposobów zbierania danych sondażowych, poświęćmy chwilę na ich bezpośrednie porównanie.

Kwestionariusze do samodzielnego wypełniania przez respondentów są generalnie tańsze i szybsze niż wywiady bezpośrednie, przeprowadzane osobiście przez ankietera. Te względy mogą być istotne dla pozbawionego środków finansowych studenta, który chce przeprowadzić badania sondażowe, by napisać na ich podstawie referat semestralny czy rozprawę. Co więcej, przy zastosowaniu formatu kwestionariusza wypełnianego przez respondenta przeprowadzenie sondażu ogólnokrajowego nie kosztuje więcej niż przeprowadzenie sondażu lokalnego na próbie tej samej wielkości. Natomiast sondaż ogólnokrajowy realizowany za pomocą wywiadów (bezpośrednich bądź telefonicznych) kosztowałby znacznie więcej niż sondaż o zasięgu lokalnym. Do sondaży pocztowych potrzeba też z reguły mniej personelu: jedna osoba może samodzielnie przeprowadzić sensowny sondaż pocztowy, choć nie można nie doceniać potrzebnego przy tym nakładu pracy. Poza tym respondenci mają niekiedy opory przed opisywaniem w wywiadach postaw lub zachowań kontrowersyjnych czy odbiegających od normy, ale chętnie udzielą odpowiedzi w anonimowym kwestionariuszu przeznaczonym do samodzielnego wypełnienia.

Również sondaże realizowane dzięki wywiadom mają wiele zalet. Na przykład na ogół otrzymuje się w nich mniej niekompletnych kwestionariuszy. Chociaż respondenci mogą przeskakiwać pytania w samodzielnie wypełnianym kwestionariuszu, to ankieterzy są poinstruowani, by tego nie robić. W wywiadach CATI komputer dodatkowo gwarantuje kompletność wypełnienia kwestionariusza. Ponadto sondaże realizowane poprzez wywiady mają z reguły wyższy odsetek odpowiedzi niż sondaże przeprowadzane przy użyciu kwestionariuszy do samodzielnego wypełniania.

Mimo że sondaże przeprowadzane przy zastosowaniu kwestionariuszy do samodzielnego wypełniania mogą być bardziej efektywne przy drażliwych kwestiach, to sondaże z zastosowaniem wywiadów są zdecydowanie bardziej efektywne przy kwestiach bardziej skomplikowanych. Do najważniejszych przykładów należy wyliczenie członków gospodarstwa domowego i określenie, czy dany adres odpowiada więcej niż jednej jednostce mieszkalnej. Choć pojęcie jednostki mieszkalnej zostało udoskonalone i wystandaryzowane przez Biuro Spisowe i można przeszkolić ankieterów, by je opanowali, to jednak przekazanie go w kwestionariuszu do samodzielnego wypełniania przez respondenta jest rzeczą skrajnie trudną. Ta przewaga sondaży z zastosowaniem wywiadu odnosi się generalnie do wszystkich skomplikowanych pytań warunkowych.

Realizując sondaż poprzez wywiady, można przeprowadzić go na próbie adresów lub numerów telefonów zamiast nazwisk. Ankieter może przyjść pod wskazany adres lub zadzwonić pod wskazany numer, opowiedzieć, o co chodzi w badaniach, a nawet – zgodnie z instrukcjami – wybrać właściwą osobę pod tym adresem do udziału w wywiadzie. W przeciwieństwie do tego, sondaże realizowane za pomocą kwestionariuszy do samodzielnego wypełniania adresowanych ogólnie do „mieszkańców" notorycznie mają bardzo niską zwrotność.

I wreszcie, jak już się przekonaliśmy, ankieterzy zadający respondentom pytania w wywiadach bezpośrednich mogą poczynić istotne obserwacje, niezależnie od odpowiedzi na zadawane pytania. Przebywając w gospodarstwie domowym, którego dotyczy wywiad, mogą odnotować cechy charakterystyczne jego otoczenia, samego mieszkania itd. Mogą też odnotować cechy respondentów albo jakość interakcji z nimi – czy respondent miał kłopoty z komu-

nikacją, czy był wrogo nastawiony, czy wydawało się, że kłamie itd.

Główne zalety sondaży telefonicznych w stosunku do wywiadów bezpośrednich skupiają się przede wszystkim na czasie i pieniądzach. Wywiady telefoniczne są znacznie tańsze[7] i mogą być szybko zorganizowane i przeprowadzone. Większe jest również bezpieczeństwo ankieterów, którzy przeprowadzają wywiady na obszarach o wysokiej przestępczości. Ponadto wpływ ankieterów na odpowiedzi jest nieco zmniejszony, gdy pozostają oni niewidoczni dla respondentów. Za jeden ze wskaźników popularności wywiadów telefonicznych może posłużyć zestawiona przez Johnny'ego Blaira i jego kolegów (1995) bibliografia rodzajów próby stosowanych w wywiadach telefonicznych – uwzględniała ona ponad 200 pozycji.

Sondaże przeprowadzane przy użyciu Internetu mają wiele zalet i wad ankiet pocztowych. Gdy dostępne oprogramowanie zostanie udoskonalone, mogą one stać się znacznie tańsze. Ich poważną wadą jest jednak trudność zapewnienia, że respondenci w sondażu „online" są reprezentatywni dla jakiejś większej populacji.

Oczywiście każda z metod badań sondażowych ma swoje miejsce w badaniach społecznych. Ostatecznie należy wyważyć wady i zalety różnych metod w odniesieniu do własnych potrzeb badawczych i własnych zasobów.

Mocne i słabe strony badań sondażowych

Niezależnie od zastosowanej konkretnej metody sondaże – podobnie jak inne metody obserwacji w badaniach społecznych – mają charakterystyczne dla siebie mocne i słabe strony. Powinno się je brać pod uwagę przy podejmowaniu decyzji, czy sondaż będzie właściwą metodą dla danych celów badawczych.

Sondaże są szczególnie przydatne przy opisywaniu cech większej populacji. Starannie dobrana próba losowa w połączeniu ze standaryzowanym kwestionariuszem umożliwiają formułowanie wyszukanych twierdzeń opisowych o ogóle studentów, o mieście, o społeczności kraju czy też o dowolnej innej dużej populacji. Sondaże służą do określenia stopy bezrobocia, zamiarów wyborczych i tym podobnych zjawisk z zadziwiającą dokładnością. I chociaż analiza oficjalnych dokumentów – takich jak rejestry ślubów, urodzeń czy zgonów – może zapewnić taką samą dokładność w kilku tematach, to jednak żadna inna metoda obserwacji nie wykazuje takiej ogólnej zdolności.

Sondaże, zwłaszcza te przeprowadzane przy użyciu kwestionariuszy do samodzielnego wypełniania, umożliwiają stosowanie dużych prób. Niczym niezwykłym nie są sondaże na próbie 2000 osób. Duża liczba przypadków jest bardzo istotna zarówno w analizach opisowych, jak i wyjaśniających, zwłaszcza gdy poddaje się analizie wiele zmiennych naraz.

Sondaże są w pewnym sensie elastyczne. Można w nich zadać wiele pytań dotyczących danego tematu, co daje istotną elastyczność w analizach. Model eksperymentalny może wymagać przyjęcia z góry jakiejś konkretnej definicji operacyjnej danego pojęcia, natomiast w sondażu możliwe jest opracowanie definicji operacyjnych dopiero na podstawie faktycznie uzyskanych obserwacji.

Wreszcie, standaryzowane kwestionariusze mają istotną zaletę, jeśli chodzi ogólnie o pomiar. We wcześniejszych rozdziałach omawialiśmy wieloznaczny charakter większości pojęć: nie mają one żadnych rzeczywistych znaczeń. Religijność jednej osoby jest zupełnie inna niż religijność innej. Choć powinno się móc zdefiniować pojęcia w sposób najlepiej odpowiadający celom badawczym, to jednak zastosowanie tych samych definicji do wszystkich badanych może okazać się niełatwe. Badacz sondażowy jest zmuszony do spełniania tego wymogu, gdyż musi on zadawać dokładnie te same pytania wszystkim badanym i zakładać te same znaczenia u wszystkich respondentów udzielających jakiejś konkretnej odpowiedzi.

Sondaże mają również słabe strony. Przede wszystkim zdaje się, że wymóg standaryzacji często kończy się upychaniem na siłę materii do zbyt ciasnej formy. Standaryzowane pytania

[7] Ta opinia autora odnosi się do warunków amerykańskich (przyp. tłum.).

kwestionariusza stanowią często najmniejszy wspólny mianownik przy określaniu postaw, poglądów, sytuacji materialnej czy doświadczeń ludzi. Przy układaniu pytań, które będą przynajmniej w minimalnym stopniu odpowiednie dla wszystkich respondentów, można przeoczyć coś, co jest najodpowiedniejsze dla wielu z nich. W tym znaczeniu sondaże często sprawiają wrażenie powierzchownego traktowania złożonych tematów. I chociaż problem ten może zostać częściowo przezwyciężony dzięki zaawansowanym analizom, jest on zawsze obecny w badaniach sondażowych.

Badania sondażowe rzadko mają też do czynienia z kontekstem życia społecznego. Mimo że kwestionariusze mogą dostarczyć informacji na ten temat, rzadko zdarza się, by badacze sondażowi mieli tego rodzaju wyczucie całościowej sytuacji życiowej, w której myślą i działają respondenci, jakie mają, powiedzmy, badacze posługujący się obserwacją uczestniczącą (zobacz rozdział 10).

Sondaże są również pod wieloma względami nieelastyczne. Modyfikacje badań, prowadzonych za pomocą obserwacji uczestniczącej, mogą być uzasadnione zmieniającymi się warunkami zewnętrznymi, natomiast dla badań sondażowych typowy jest wymóg niezmieniania początkowego kształtu badań w jego trakcie. Na przykład jako badacz terenowy możesz uświadomić sobie istnienie ważnej nowej zmiennej oddziałującej na badane przez ciebie zjawisko i możesz rozpocząć staranne zbieranie informacji na jej temat. W sondażu badacz zapewne pozostałby nieświadom wagi takiej nowej zmiennej i nie mógłby niczego zrobić.

Wreszcie, sondaże poprzez sztuczność sytuacji badawczej prowadzą do artefaktów, o których wspominaliśmy wcześniej w związku z eksperymentami. To, że ktoś udziela konserwatywnych odpowiedzi na pytania kwestionariusza, niekoniecznie musi oznaczać, że jest on konserwatywny; to, że jakaś inna osoba udziela nacechowanych uprzedzeniami odpowiedzi na pytania kwestionariusza, niekoniecznie oznacza, że jest ona uprzedzona. Ta słabość sondażu jest szczególnie wyraźna w dziedzinie działań. Sondaże nie mogą mierzyć działań społecznych; mogą one jedynie gromadzić relacje badanych o ich przeszłych działaniach bądź też o działaniach przyszłych lub hipotetycznych.

Problem pewnej sztuczności badań sondażowych, prowadzącej do powstawania artefaktów, ma dwa aspekty. Po pierwsze, przedmiot badań może nie nadawać się do pomiaru za pomocą kwestionariusza. Po drugie, sam akt badania tego przedmiotu – na przykład postaw – może mieć na niego wpływ. Respondent w sondażu być może ani przez chwilę nie zaprzątał sobie głowy tym, czy gubernator powinien być postawiony w stan oskarżenia, dopóki ankieter nie zapytał go o opinię na ten temat. Może on wyrobić sobie taką opinię dopiero w trakcie badania.

Ogólnie można powiedzieć, że słabą stroną sondaży jest trafność, mocną zaś – rzetelność. Na przykład w porównaniu z badaniami terenowymi, sztuczność formy sondażu obciąża jego trafność. Ilustracją tego niech będzie fakt, że opinie ludzi w różnych kwestiach rzadko przybierają formę zdecydowanej zgody, zgody, niezgody lub zdecydowanej niezgody z jakimś konkretnym stwierdzeniem. Ich odpowiedzi na pytania sondażu należy zatem traktować jako przybliżone wskaźniki tego, co badacze mieli na myśli, układając pytania. Uwagi te należy jednak odczytywać w kontekście wcześniejszych rozważań na temat wieloznaczności samego pojęcia trafności. Stwierdzenie, że coś jest trafną bądź nietrafną miarą, zakłada istnienie „prawdziwej" definicji tego, co jest mierzone, obecnie zaś wielu badaczy odrzuca to założenie.

Rzetelność jest prostszą sprawą. Prezentując wszystkim badanym standaryzowany bodziec, badania sondażowe przyczyniają się w znacznym stopniu do eliminowania nierzetelności obserwacji powodowanej przez badacza. Także staranne sformułowanie pytań może znacząco zredukować nierzetelność powodowaną przez samych badanych.

Tak jak się to dzieje z wszystkimi metodami obserwacji, pełna świadomość słabości, właściwych badaniom sondażowym lub tylko możliwych w nich, może częściowo im zapobiegać w niektórych przypadkach. Ostatecznie jednak, prowadząc badania na dany temat, najbezpieczniej jest stosować kilka metod badawczych.

Analiza wtórna

Jeko jeden ze sposobów prowadzenia obserwa- i badania sondażowe obejmują następujące tapy: (1) budowa kwestionariusza, (2) dobór róby, (3) zbieranie danych, czy to poprzez wy-wiady, czy też kwestionariusze do samodzielne-go wypełniania. Jak widać, sondaże są na ogół poważnym przedsięwzięciem. Nie jest niczym niezwykłym w sondażach prowadzonych na dużą skalę, że postęp od konceptualizacji do uzyskania gotowych danych zabiera kilka mie-sięcy czy nawet ponad rok. (Sondaże na mniej-szą skalę mogą być, rzecz jasna, zrealizowane szybciej). Dzięki metodzie nazywanej *analizą wtórną* badacze mogą jednak oddawać się swo-im zainteresowaniom badawczym – poprzez analizę danych sondażowych uzyskanych na przykład na podstawie ogólnokrajowej próby 2000 respondentów – bez spędzania takiej ilości czasu i wydawania takiej ilości pieniędzy, ja-kiej wymaga tego rodzaju sondaż.

Analiza wtórna to taka forma badań, w której dane zebrane i przetworzone przez jednego badacza są ponownie analizowane – często w odmiennym celu – przez innego. Począwszy od lat 60. badacze sondażowi za-częli sobie uświadamiać potencjalną wartość archiwizowania danych sondażowych w celu umożliwienia ich analizy przez uczonych, któ-rzy nie mieli nic wspólnego z planowaniem tych sondaży czy zbieraniem do nich danych. Nawet jeśli jakiś badacz przeprowadził sondaż i zanalizował dane, te same dane mogłyby być dalej analizowane przez innych, mających nie-co odmienne zainteresowania badawcze. Gdy-byś zatem zainteresował się związkiem między poglądami politycznymi a postawami wobec równości płci, mógłbyś zbadać ten problem poprzez analizę dowolnego zbioru danych, który akurat zawiera pytania odnoszące się do tych dwóch zmiennych.

Początkowo archiwa danych były bardzo podobne do bibliotek, z wyjątkiem paru różnic. Po pierwsze, zamiast książek archiwa danych zawierały zbiory danych: najpierw w postaci kart perforowanych, potem taśm magnetycz-nych. W dzisiejszych czasach na ogół mieszczą się one na dyskach komputerowych, CD-ROM--ach lub serwerach z dostępem online. To zu-gie, w bibliotece oczekuje się zwrotu książek, natomiast dane otrzymane z archiwum danych można sobie zatrzymać.

Najbardziej obecnie znanym przykładem analiz wtórnych jest Generalny Sondaż Spo-łeczny (*General Social Survey, GSS*)[8]. Co rok lub dwa lata rząd federalny zleca Narodowemu Centrum Badania Opinii (*National Opinion Re-search Center, NORC*) przy Uniwersytecie Chicago przeprowadzenie wielkiego ogólnokrajowego sondażu gromadzącego dane na temat ogrom-nej liczby zmiennych społecznych. Sondaże te są przeprowadzane właśnie w celu udostęp-nienia danych naukowcom za niewielką opłatą albo nieodpłatnie. Możesz dowiedzieć się wię-cej o tych badaniach pod adresem

http://www.icpsr.umich.edu:8080/ICPSR-STUDY/03728.xml

Dostępne są także liczne inne źródła infor-macji o istnieniu i możliwościach nabycia da-nych sondażowych do wtórnej analizy. Doskonałym źródłem jest Roper Center for Public Opi-nion Research, istniejące przy Uniwersytecie Connecticut (http://www.ropercenter.uconn.edu/). Publikuje ono też czasopismo dotyczące badań opinii publicznej *Public Perspective*. Polling the Nations (http://www.pollingthenations.com/) jest magazynem online, zawierającym tysiące sondaży, przeprowadzonych w USA i 70 innych krajach. Płatna subskrypcja umożliwia użyt-kownikom otrzymywanie określonych wyni-ków konkretnych badań, nie zaś otrzymywanie badań w całości.

Poza granicami USA, Holenderski Instytut Usług Informacji Naukowej (*the Netherlands In-stitute for Scientific Information Services*) pod ad-resem http://www.niwi.knaw.nl/us/homepag.htm umożliwia swoim użytkownikom odnalezienie europejskich badań, zawierających interesujące ich zmienne. Można też wypróbować Centralne Archiwum Badań Społecznych (*Central Archive for Social Science Research*) przy Uniwersytecie Ko-

[8] Polskim odpowiednikiem tych badań jest Pol-ski Generalny Sondaż Społeczny. Więcej informacji na ten temat można znaleźć pod adresem:

http://www.iss.uw.edu.pl/osrodki/obs/pgss/pl (przyp. tłum.).

lońskim w Niemczech (http://www.gesis.org/en/za). (W załączniku B podano adresy wielu stron www, przydatnych badaczom społecznym, i tam właśnie możesz znaleźć sporo innych źródeł danych).

Zalety wtórnej analizy są oczywiste i ogromne: jest tańsza i szybsza od przeprowadzania oryginalnych badań, a w zależności od tego, kto przeprowadził źródłowy sondaż, można korzystać z pracy najlepszych profesjonalistów. Istnieją jednak również wady tej metody. Kluczowy problem to stale powracająca kwestia trafności. Gdy jakiś badacz zbiera dane potrzebne do konkretnych celów badawczych, nie masz żadnej gwarancji, że dane te będą się nadawały także do twoich celów. Z reguły okaże się, że badacz zadawał pytanie, które „w przybliżeniu" mierzyło to, co interesuje ciebie, ale ty zadałbyś to pytanie nieco inaczej – lub zadałbyś także inne pytanie, związane z poprzednim. Twoim problemem jest teraz, czy zadane pytanie stanowi trafny pomiar zmiennej, którą chcesz analizować. Mimo to analiza wtórna jest ogromnie użyteczna. Ilustruje ona ponadto po raz kolejny zakres możliwości dostępnych przy poszukiwaniu odpowiedzi na pytania dotyczące życia społecznego. Choć żadna metoda sama w sobie nie rozwiąże wszystkich zagadek, sposoby zdobywania wiedzy są nieograniczone. A gdy skupisz się na jakiejś sprawie, przyglądając się jej z kilku niezależnych perspektyw, uzyskasz w ten sposób znacznie większą wiedzę na jej temat.

Główne zagadnienia

- Badania sondażowe, popularna metoda badań społecznych, polegają na wypełnieniu kwestionariuszy przez próbę respondentów dobraną z pewnej populacji.
- Badania sondażowe nadają się zwłaszcza do opisowych badań dużych populacji; dane sondażowe mogą zostać również wykorzystane do celów wyjaśniania.
- Kwestionariusze są narzędziem zbierania danych poprzez (1) zadawanie ludziom pytań lub (2) przedstawienie im stwierdzeń reprezentujących różne stanowiska wraz z prośbą o określenie, czy się z nimi zgadzają, czy nie. Pytania mogą być otwarte (respondenci sami formułują odpowiedzi) lub zamknięte (wybierają oni odpowiedź z listy).
- Przy układaniu pytań kwestionariusza należy uwzględniać pewne wskazówki: (1) pytania muszą być jasne i precyzyjne; (2) pytania powinny dotyczyć tylko jednej sprawy (tzn. należy unikać pytań podwójnych); (3) respondenci muszą być w stanie udzielić odpowiedzi; (4) respondenci muszą chcieć udzielić odpowiedzi; (5) pytania powinny dotyczyć respondentów; (6) zasadniczo pytania powinny być krótkie; (7) należy unikać negatywnych określeń, by nie sprawiać kłopotu respondentom; (8) sformułowanie pytań powinno być takie, by nie prowadziło do obciążonych odpowiedzi.
- Format kwestionariusza może mieć wpływ na jakość uzyskanych danych.
- Niezbędny jest przejrzysty format pytań warunkowych, tak aby zapewnić, że respondenci odpowiadają na wszystkie przeznaczone dla nich pytania.
- Pytania tabelaryczne są efektywną formą przedstawiania wielu pytań mających wspólne kategorie odpowiedzi.
- Kolejność pytań w kwestionariuszu może mieć wpływ na udzielane odpowiedzi.
- Jasne instrukcje mają istotne znaczenie dla uzyskania właściwych odpowiedzi.
- Kwestionariusze powinny być przetestowane w badaniu próbnym (pilotażu), zanim zostaną rozdane respondentom wchodzącym w skład próby.
- Kwestionariusze mogą być wypełniane na trzy sposoby: samodzielnie wypełniane przez respondentów, wypełniane przez ankietera w wywiadach bezpośrednich bądź poprzez sondaż telefoniczny.
- Ogólnie godne polecenia w przypadku ankiet pocztowych jest zaplanowanie wysyłki przypominającej, związanej z ponownym wysłaniem kwestionariuszy do tych respondentów, którzy nie odpowiedzieli na pierwszą prośbę. Właściwie monitorowane zwroty kwestionariuszy są dobrą wskazówką co do momentu dokonania ponownej wysyłki.
- Podstawową cechą ankieterów powinna być

neutralność; nie można dopuścić do tego, by ich obecność w procesie zbierania danych miała jakikolwiek wpływ na odpowiedzi udzielane na pytania kwestionariusza. Należy starannie przeszkolić ankieterów, zapoznając ich z kwestionariuszem, tak aby trzymali się ściśle brzmienia i kolejności pytań i dokładnie zapisywali odpowiedzi respondentów.

Ankieterzy mogą stosować dopytywanie w celu uzyskania dokładniejszych wyjaśnień w przypadku niekompletnych lub wieloznacznych odpowiedzi. Dopytywanie powinno być neutralne. W idealnej sytuacji wszyscy ankieterzy powinni używać tych samych sformułowań dopytujących.

- Sondaże telefoniczne mogą być tańsze i bardziej efektywne niż wywiady bezpośrednie i mogą zapewnić większą kontrolę nad procesem zbierania danych. Obiecujący jest zwłaszcza rozwój technik wywiadów telefonicznych wspomaganych komputerowo (CATI).
- Nowe technologie są źródłem dodatkowych możliwości dla badaczy społecznych. Obejmują one różne rodzaje wspomaganego komputerowo zbierania danych i ich analizy, jak również możliwość prowadzenia badań za pomocą faksu lub Internetu. Jednak te dwie ostatnie metody należy stosować ostrożnie, gdyż respondenci mogą nie być reprezentatywni dla docelowej populacji.

Sondaże przeprowadzane przy zastosowaniu kwestionariusza wypełnianego samodzielnie przez respondentów mają przewagę nad wywiadami pod względem oszczędności kosztów, szybkości, braku wpływu ankietera oraz pod względem możliwości zapewnienia anonimowości i dyskrecji w celu zachęcenia respondentów do wypowiedzi na drażliwe kwestie.

Sondaże przeprowadzane za pomocą wywiadów mają przewagę nad sondażami realizowanymi z zastosowaniem kwestionariusza wypełnianego samodzielnie przez respondentów pod względem mniejszej liczby niekompletnych kwestionariuszy i mniejszej liczby źle zrozumianych pytań, ogólnie większego odsetka udzielonych odpowiedzi oraz pod względem większej elastyczności w kwestiach doboru próby i dodatkowych obserwacji.

- Główne zalety sondaży telefonicznych w porównaniu z wywiadami bezpośrednimi to oszczędność kosztów i czasu. Ankieterzy w badaniach telefonicznych są również bezpieczniejsi niż w wywiadach osobistych i mogą mieć mniejszy wpływ na sam wywiad.
- Sondaże przez Internet mają wiele wspólnych wad i zalet z ankietami pocztowymi. Choć są tańsze, trudno jest zagwarantować, że respondenci reprezentują jakąś szerszą populację.
- Ogółem zalety badań sondażowych wiążą się z oszczędnością kosztów, dużą ilością danych, które można zebrać, i możliwością dobierania prób z dużych populacji. Inną specyficzną zaletą badań sondażowych jest standaryzacja zbieranych danych.
- Słabością badań sondażowych jest ich pewna sztuczność, potencjalna powierzchowność i dość mała elastyczność. Trudno jest stosować sondaże dla zrozumienia pełnego znaczenia procesów społecznych w ich naturalnej oprawie. Ogólnie rzecz biorąc, słabą stroną sondaży jest trafność, mocną zaś – rzetelność.
- Analiza wtórna to dla badaczy społecznych opcja taniego i łatwego „zbierania" danych, ale może to się odbywać kosztem trafności.

Kluczowe pojęcia

kwestionariusz
respondent
pytania otwarte
pytania zamknięte
obciążenie
pytania warunkowe
poziom zwrotów
wywiad
dopytywanie
analiza wtórna

Pytania i ćwiczenia

1. Ułóż dla każdego z podanych niżej pytań otwartych pytanie zamknięte, które mogłoby się znaleźć w kwestionariuszu.

a. Jaki był całkowity dochód Pana(i) rodziny w ostatnim roku?
b. Co Pan(i) sądzi o programie kosmicznym wykorzystującym wahadłowce?
c. Jak ważna jest w Pana(i) życiu religia?
d. Jaka jest główna przyczyna, dla której poszedłeś/poszłaś do college'u?
e. Jak sądzisz, co jest głównym problemem twojej miejscowości?

2. Ułóż blok pytań warunkowych do zastosowania w kwestionariuszu przeznaczonym do samodzielnego wypełniania przez respondenta, który dotyczyłby następujących informacji:
   a. Czy respondent jest zatrudniony?
   b. Jeśli nie jest zatrudniony, czy szuka pracy?
   c. Jeśli nie jest zatrudniony i nie szuka pracy, to czy jest rencistą, studentem lub prowadzi gospodarstwo domowe?
   d. Jeśli respondent szuka pracy, to od jak dawna?
3. Odszukaj w prasie jakiś kwestionariusz (na przykład przeznaczony do przeprowadzenia sondażu wśród czytelników). Wybierz z niego przynajmniej pięć pytań i poddaj je krytycznej analizie, oceniając każde z nich pozytywnie lub negatywnie.
4. Przyjrzyj się teraz swojemu aktualnemu wyglądowi. Określ, jakie aspekty twojego wyglądu mogłyby stanowić problem, gdybyś był ankieterem pracującym przy realizacji sondażu na przekrojowej próbie społeczeństwa.
5. Znajdź jakiś sondaż przeprowadzany w sieci. Opisz krótko ten sondaż i zastanów się, jakie są jego mocne i słabe strony.

Dodatkowe lektury

Babbie Earl, *Survey Research Methods*, Wadsworth, Belmont, Cal. 1990. Wszechstronny przegląd metod sondażowych. (Myślałeś, że napiszę, że jest beznadziejny)? Podręcznik ten, chociaż nieco pokrywa się z niniejszym, obejmuje także te aspekty technik sondażowych, które zostały tutaj przeze mnie pominięte.

Bradburn Norman M., Seymour Sudman, *Polls and Surveys: Understanding What They Tell Us*, Jossey-Bass, San Francisco 1988. Weterani badań sondażowych odpowiadają na najczęściej zadawane pytania dotyczące uprawianego przez nich zajęcia.

Dillman Don A., *Mail and Telephone Surveys: The Total Design Method*, Wiley, New York 1978. Znakomity przegląd literatury metodologicznej na temat sondaży pocztowych i telefonicznych. Dillman podpowiada wiele dobrych sposobów na zwiększenie odsetka odpowiedzi.

Elder Glen H. Jr., Pavalko Eliza K., Clipp Elizabeth C., *Working with Archival Data: Studying Lives*, Sage, Newbury Park, Cal. 1993. Książka omawia możliwości i techniki wykorzystania istniejących archiwów danych w USA, zwłaszcza tych, które zawierają dane dynamiczne.

Feick Lawrence F., *Latent Class Analysis of Survey Questions That Include Don't Know Responses*, „Public Opinion Quarterly" 1989, 53, s. 525–547. Jak wykazuje analiza, „nie wiem" może oznaczać wiele różnych rzeczy.

Fowler Floyd J. Jr., *Improving Survey Questions: Design and Evaluation*, Sage, Thousand Oaks, Cal. 1995. Wszechstronne omówienie budowy kwestionariusza, w tym liczne sugestie dotyczące wstępnego testowania pytań. Książka omawia logikę pozyskiwania informacji poprzez pytania sondażu i daje liczne wskazówki dotyczące efektywności. Podano również wiele przykładowych pytań, które można wykorzystać.

Groves Robert M., *Theories and Methods of Telephone Surveys*, „Annual Review of Sociology" 1990, t. 16, s. 221–240, Richard W. Scott, Judith Blake (red.), Palo Alto, Cal. 1990. Próba umieszczenia sondaży telefonicznych w kontekście teorii socjologicznych i psychologicznych i omówienia różnego rodzaju błędów występujących na ogół przy stosowaniu tej metody badawczej.

Miller Delbert, *Handbook of Research Design and Social Measurement*, Sage, Newbury Park, Cal. 1991. Znakomity informator. W książce tej, zwłaszcza w części 6 autorzy przytaczają i opisują szeroki zakres mierników operacyjnych stosowanych we wczesnych badaniach społecznych. W paru przypad-

kach zaprezentowano format kwestionariuszy zastosowanych w tych badaniach. Choć jakość tych przykładów jest nierówna, stanowią one doskonałą ilustrację różnych możliwych odmian.

Sheatsley Paul F., *Questionnaire Construction and Item Writing*, w: *Handbook of Social Research*, Peter H. Rossi, James D. Wright, Andy B. Anderson (red.), Academic Press, New York 1983, s. 195–230. Znakomita analiza tematu dokonana przez eksperta w tej dziedzinie.

Smith Eric R. A. N., Peverill Squire, *The Effects of Prestige Names in Question Wording*, „Public Opinion Quarterly" 1990, 54, s. 97–116. Prestiżowe nazwiska nie tylko mają ogólny wpływ na odpowiedzi udzielane w sondażu, ale wpływają one również na takie rzeczy, jak korelacja między wykształceniem a liczbą odpowiedzi „nie wiem".

Swafford Michael, *Soviet Survey Research: The 1970's vs. the 1990's*, „AAPOR News" 1992, 19 (3), s. 3–4. Autor przeciwstawia generalne hamowanie badań sondażowych, jakie obserwował w czasie swojego pierwszego pobytu w ZSRR w latach 1973–1974, powrotowi do ich stosowania w nowszych czasach. Odnotowuje on na przykład, że rząd sowiecki zlecił ogólnokrajowy sondaż, aby poznać opinię publiczną na temat możliwego zjednoczenia Niemiec.

Tourangeau Roger, Rasinski Kenneth A., Bradburn Norman, D'Andrade Roy, *Carryover Effects in Attitude Surveys*, „Public Opinion Quarterly" 1989, 53, s. 495–524. Autorzy zadali sześć głównych pytań w sondażu telefonicznym, przeprowadzonym na próbie 1100 respondentów, zmieniając pytania bezpośrednio je poprzedzające. Okryli zasadnicze różnice.

Williams Robin M. Jr., *The American Soldier: An Assessment, Several Wars Later*, „Public Opinion Quarterly" 1989, 53, s. 155–174. Jedno z klasycznych studiów z historii badań sondażowych jest ponownie analizowane przez jednego z jego autorów.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

10

Jakościowe badania terenowe

Przegląd holograficzny

Jakościowe badania terenowe umożliwiają badaczom obserwowanie życia społecznego w jego naturalnym środowisku: idź tam, gdzie toczy się proces, i obserwuj go. Jeśli badacz prowadzi obserwacje w sposób przemyślany, dobrze zaplanowany i aktywny, to ten rodzaj badań umożliwia dogłębniejsze zrozumienie wielu zjawisk społecznych, niż byłoby to możliwe przy zastosowaniu innych metod obserwacyjnych.

Wprowadzenie

Tematy odpowiednie dla badań terenowych

Rzeczy, które należy brać pod uwagę w jakościowych badaniach terenowych
Różne role obserwatora
Stosunek do badanych

Niektóre paradygmaty jakościowych badań terenowych
Naturalizm
Etnometodologia
Teoria ugruntowana
Studium przypadku i metoda rozszerzonych przypadków
Etnografia instytucjonalna
Badania uczestniczące

Prowadzenie jakościowych badań terenowych
Przygotowanie do pracy w terenie
Wywiady jakościowe
Grupy fokusowe
Rejestrowanie obserwacji

Mocne i słabe strony jakościowych badań terenowych
Trafność
Rzetelność

Etyka badawcza w jakościowych badaniach terenowych

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Kilka rozdziałów wcześniej sugerowałem, że prowadzimy badania społeczne przez całe życie. Stanie się to jeszcze jaśniejsze, gdy spojrzymy na chyba najbardziej oczywistą metodę obserwacji: jakościowe badania terenowe. W pewnym sensie prowadzimy badania terenowe, ilekroć obserwujemy zachowania społeczne lub uczestniczymy w nich i próbujemy te zachowania zrozumieć, czy to w sali wykładowej, w poczekalni u lekarza czy w samolocie. Ilekroć przedstawiamy innym, co udało się nam dostrzec, zdajemy sprawę z naszych dociekań badawczych.

Badania tego rodzaju są w naukach społecznych jednocześnie i bardzo stare, i zupełnie nowe. Wiele technik omawianych w tym rozdziale jest stosowanych od stuleci. Wśród badaczy zajmujących się naukami społecznymi szczególnie związani z tą metodą są antropologowie, którzy wnieśli znaczący wkład w jej rozwój jako techniki naukowej. Co więcej, narzędzi nieco podobnych do tej metody używa wielu ludzi, którzy, ściśle biorąc, nie powinni być uważani za badaczy społecznych. Jednym z przykładów są reporterzy prasowi, innym pracownicy pomocy społecznej.

Chociaż są to działania „naturalne", stanowią one również umiejętności, których można się nauczyć, i które można doskonalić. Niniejszy rozdział omawia te umiejętności, zwracając uwagę na kilka szczegółów, przedstawia niektóre z głównych paradygmatów badań terenowych oraz opisuje pewne konkretne techniki, które czynią naukowe badania terenowe bardziej użytecznymi od codziennej obserwacji, z którą wszyscy mamy do czynienia.

Używam terminu *jakościowe badania terenowe* dla odróżnienia tego typu metod obserwacyjnych od metod zaprojektowanych do generowania danych dla analiz ilościowych. Tak więc badania sondażowe dostarczają danych, z których można obliczyć odsetek bezrobotnych w populacji, średni dochód itp. Badania terenowe w większości przynoszą dane jakościowe: obserwacje, których nie można łatwo sprowadzić do liczb. Badacz terenowy może na przykład określić postępowanie liderów na zjeździe partyjnym jako „zachowanie paternalistyczne", urzędnikowi zaś podczas przesłuchania publicznego przypisać „stosowanie uników", nie próbując wyjaśnić ani paternalizmu, ani uników za pomocą liczbowych wartości lub stopni. Badania terenowe mogą zostać zastosowane także do zbierania danych ilościowych – na przykład rejestrowania liczby poszczególnych typów interakcji w danych warunkach terenowych – typowe badania terenowe są jednak jakościowe.

Obserwacja terenowa jest też odmienna od innych modeli obserwacji w tym, że nie jest ona tylko gromadzeniem danych. Często, a być może na ogół, jest także działalnością generującą teorię. Jako badacz terenowy rzadko znajdziesz się w sytuacji, gdy twoim zadaniem będzie towarzyszyć precyzyjnie sformułowana hipoteza, którą trzeba poddać testowaniu. Częściej będziesz starał się zrozumieć toczący się proces, którego nie można przewidzieć z góry – prowadząc wstępne obserwacje, opracowując pierwsze ogólne konkluzje wskazujące poszczególne rodzaje dalszej obserwacji, podejmując te obserwacje i poprzez to rewidując swoje konkluzje itd. Mówiąc krótko, przemienność indukcji i dedukcji omówiona w pierwszej części tej książki być może nigdzie nie jest lepiej widoczna i istotna niż w dobrych badaniach terenowych. Jednak w celu lepszego wytłumaczenia istoty badań terenowych rozdział ten koncentruje się przede wszystkim na niektórych teoretycznych podstawach badań terenowych i na technikach zbierania danych. Rozdział 13 omawia zaś sposoby analizy danych jakościowych.

Tematy odpowiednie dla badań terenowych

Jedną z kluczowych zalet badań terenowych jest całościowa perspektywa, jaką mogą dać one badaczowi. Przez podążanie wprost do zjawiska społecznego poddawanego badaniu i obserwowanie go na tyle całościowo, na ile jest to możliwe, badacz ma szansę uzyskać głębsze i pełniejsze jego rozumienie. Jako taki ten rodzaj obserwacji nadaje się szczególnie, choć nie

wyłącznie, do tematów badawczych i studiów społecznych, które nie poddają się prostemu ujęciu ilościowemu. Badacze terenowi mogą rozpoznawać wiele niuansów w postawach i zachowaniach, które mogą umykać badaczom stosującym inne metody.

Badania terenowe są szczególnie odpowiednie do poznawania tych postaw i zachowań, które najlepiej można zrozumieć w ich naturalnym otoczeniu, w przeciwieństwie do nieco sztucznych warunków przeprowadzania eksperymentu i sondażu. Na przykład badania terenowe stanowią pierwszorzędną metodę studiowania dynamiki konwersji religijnej na spotkaniu poświęconemu odnowie, tak jak statystyczne analizy przepływu wiernych byłyby lepszym narzędziem do odkrycia, czy bardziej skłonni do konwersji byli mężczyźni, czy kobiety.

Na koniec, badania terenowe dobrze nadają się do badania procesów społecznych dziejących się w czasie. Badacz terenowy będzie więc w stanie badać przejawy niezadowolenia i w końcu wybuch zamieszek w trakcie zachodzenia tych wydarzeń, a nie na podstawie ich późniejszej rekonstrukcji.

Inne sytuacje, w których właściwe jest zastosowanie metod badań terenowych, to demonstracja w campusie, postępowanie przed sądem, negocjacje pracownicze, przesłuchania publiczne i inne podobne wydarzenia dziejące się w stosunkowo ograniczonym czasie i przestrzeni. Kilka takich obserwacji musi składać się na bardziej całościowe badania w czasie i przestrzeni.

W *Analyzing Social Settings* (1995, s. 101–113) John i Lyn Loflandowie omawiają kilka elementów życia społecznego nadających się do przeprowadzenia badań terenowych.

1. *Działania zwyczajowe*: różne rodzaje zachowań, takie jak rozmowa lub czytanie książki.
2. *Epizody*: różnorodne wydarzenia, takie jak rozwód, przestępstwo lub choroba.
3. *Spotkania*: dwie osoby lub więcej spotyka się i zachodzi między nimi interakcja.
4. *Role*: analiza stanowisk zajmowanych przez ludzi i związanych z nimi zachowań: zawody, role odgrywane w rodzinie, grupy etniczne.
5. *Związki*: zachowanie właściwe dla dwóch lub więcej ról: relacja matka–syn, przyjaźń i podobne.
6. *Grupy*: małe grupy, takie jak grupa przyjaciół, drużyna sportowa, grupa robocza.
7. *Organizacje*: organizacje formalne, jak szpitale lub szkoły.
8. *Miejsca zamieszkania*: „społeczności" o małej skali, jak wsie, getta i społeczności sąsiedzkie, w przeciwieństwie do wielkich, trudnych do badania społeczeństw, takich jak naród.
9. *Światy społeczne*: niejednoznaczne byty społeczne o mglistych granicach i populacjach, takie jak „świat sportu" lub „Wall Street".
10. *Style życia lub subkultury*: w jaki sposób wielu ludzi przystosowało się do życia w grupach, takich jak „klasa rządząca" lub „miejska klasa niższa".

W tych wszystkich układach społecznych badania terenowe mogą ujawnić rzeczy, które w inny sposób nie byłyby widoczne. Oto konkretny przykład.

Jednym z zagadnień, które mnie szczególnie interesują (Babbie 1985), jest charakter odpowiedzialności za sprawy publiczne: kto jest odpowiedzialny za robienie rzeczy, z których wspólnie korzystamy? Kto jest odpowiedzialny za utrzymywanie w czystości miejsc publicznych, takich jak parki, centra handlowe, budynki itp.? Kto jest odpowiedzialny za oznakowanie dziurawych jezdni? A kiedy silny wiatr poprzewraca pojemniki na śmieci i toczy je po ulicach, kto jest odpowiedzialny za usunięcie ich z ulic?

Na pierwszy rzut oka odpowiedź na te pytania jest całkiem jasna. Mamy formalne i nieformalne umowy społeczne, przypisujące odpowiedzialność za te czynności. Pracownicy służb publicznych są odpowiedzialni za utrzymywanie miejsc publicznych w czystości. Pracownicy wydziału transportu są odpowiedzialni za znaki drogowe, a za pojemniki na śmieci toczące się w wietrzny dzień jest prawdopodobnie odpowiedzialna policja. Kiedy zaś te służby nie wypełniają swoich zadań, na ogół poszukujemy winnego.

Tym, co mnie fascynuje, jest stopień, w jakim odpowiedzialność za sprawy publiczne ponoszona przez ściśle określone jednostki nie tyl-

ko zwalnia innych od tej odpowiedzialności, ale faktycznie zakazuje im brania za nie odpowiedzialności. Jestem przekonany, że nie zostałoby zaakceptowane, gdyby ktoś taki jak ty lub ja wziął osobistą odpowiedzialność za sprawy publiczne, które nie zostały nam przypisane.

Pozwolę sobie zilustrować, co mam na myśli. Gdybyś szedł przez park publiczny i wyrzucił jakieś śmieci, zauważyłbyś, że twoje zachowanie było nie do przyjęcia dla otaczających cię ludzi. Zaczęliby gniewnie spoglądać, szeptać między sobą, być może ktoś zwróciłby ci uwagę. Niezależnie od formy stałbyś się obiektem wyraźnych negatywnych sankcji za śmiecenie. Jednak paradoksalnie, gdybyś spacerując przez ten sam park podszedł do kupki śmieci, które wyrzucił ktoś inny, i uprzątnął ją, prawdopodobnie twoje postępowanie również byłoby nie do przyjęcia dla ludzi wokół ciebie. Stanąłbyś w obliczu negatywnych sankcji za posprzątanie tych śmieci.

Kiedy po raz pierwszy omawiałem tę prawidłowość ze studentami, większość z nich uważała, że to przekonanie jest absurdalne. Chociaż spotkałyby nas negatywne sankcje za śmiecenie, sprzątanie miejsc publicznych powinno oczywiście przynieść reakcje pozytywne. Ludzie powinni być zadowoleni, że to robimy. Oczywiście wszyscy moi studenci powiedzieli, że byłoby im miło, jeśli ktoś posprzątałby w miejscu publicznym. Wydawało się prawdopodobne, że wszyscy inni także byliby zadowoleni, gdyby ich zapytano, jak zareagują, gdy ktoś będzie sprzątał jakieś miejsce publiczne lub w inny sposób brał na siebie odpowiedzialność za rozwiązanie jakiegoś problemu społecznego.

W celu rozstrzygnięcia tej sprawy zasugerowałem moim studentom, by zaczęli usuwać problemy społeczne, jakie napotykają w trakcie swych codziennych działań. Kiedy to zrobili, zadałem im dwa pytania i poprosiłem o zanotowanie odpowiedzi:

1. Jak się czuli, kiedy rozwiązywali problem publiczny, za który odpowiedzialność nie była im przypisana?
2. Jak reagowali inni, będący blisko?

Moi studenci sprzątali śmieci, mocowali znaki drogowe, odstawiali na miejsce przewrócone na jezdnię pojemniki na śmieci, sprzątali i dekorowali wspólne pomieszczenia w akademikach, przycinali drzewa ograniczające widoczność na skrzyżowaniach, naprawiali wyposażenie publicznych placów zabaw, sprzątali publiczne toalety i załatwiali wiele innych spraw publicznych, za które „nie byli odpowiedzialni".

Większość opisała, że czuła się niezręcznie, niezależnie od tego, co robiła. Czuli się głupio, jak świętoszki, jakby zwracali na siebie uwagę, a także odczuwali wszystko to, co normalnie powstrzymuje nas przed podejmowaniem takich działań. W prawie wszystkich przypadkach odczuwany przez nich dyskomfort rósł jeszcze w wyniku reakcji otoczenia. Jeden ze studentów usuwał właśnie zniszczony i od dawna nie używany pojemnik na gazety z przystanku autobusowego, gdzie od miesięcy stanowił on problem, gdy przybyła policja wezwana przez sąsiada. Inny student zdecydował się przeczyścić zatkaną kratkę ściekową na swojej ulicy i został skrzyczany przez sąsiada, który uparł się, żeby zostawić to służbom porządkowym. Każdy, kto zbierał śmieci, stawał się obiektem drwin, śmiechu i ogólnego poniżenia. Gdy pewien młody mężczyzna zbierał śmieci rozrzucone wokół pojemnika, jakiś przechodzień zadrwił „niezdara!". Stało się dla nas jasne, że są tylko trzy akceptowane wyjaśnienia dla zbierania śmieci w miejscu publicznym:

1. Naśmieciłeś i zostałeś złapany – ktoś zmusił cię do posprzątania po sobie.
2. Naśmieciłeś i poczułeś się winny.
3. Kradniesz śmieci.

W normalnym stanie rzeczy po prostu nie akceptuje się ludzi biorących odpowiedzialność za sprawy publiczne.

Oczywiście nie moglibyśmy odkryć natury i siły konwencji, dotyczącej brania osobistej odpowiedzialności za rzeczy publiczne, bez badań terenowych. Normy społeczne sugerują, że podejmowanie odpowiedzialności jest rzeczą dobrą, czasami określa się je jako właściwe zachowanie obywatelskie. Gdybyśmy zapytali ludzi, co sądzą o braniu odpowiedzialności, uzyskalibyśmy powszechną zgodę, że jest to właściwe. Trafny obraz mogliśmy uzys-

kać jedynie wychodząc w teren, podejmując działania i obserwując to, co nastąpi.

Za interesujący przypis do tej historii może posłużyć fakt, że moi studenci i ja zauważyliśmy, iż ilekroć przechodnie byli w stanie porzucić swe początkowe reakcje i zauważyć, że studenci po prostu czują się odpowiedzialni za usprawnienie pewnych rzeczy, byli skłonni przyłączyć się do nich. Chociaż istnieje silnie podzielana i zgodna opinia co do tego, że poczuwanie się do odpowiedzialności za sprawy publiczne jest „niebezpieczne", gotowość wzniesienia się przez kogoś ponad tę opinię sprawiła, jak się wydaje, że stawało się to bezpieczne także dla innych, którzy przyłączali się do tych działań.

Zaletą badań terenowych jest zatem możliwość badania życia społecznego w jego naturalnym środowisku. Chociaż niektóre rzeczy mogą być właściwie zbadane za pomocą kwestionariusza lub w laboratorium, nie dotyczy to wszystkich zjawisk. Natomiast bezpośrednia obserwacja w terenie pozwala badaczowi obserwować subtelności komunikacji i pewne wydarzenia, które w inny sposób nie mogłyby być przewidziane ani zmierzone.

Rzeczy, które należy brać pod uwagę w jakościowych badaniach terenowych

Każda metoda badawcza ma swoją specyfikę, którą należy uwzględniać; jakościowe badania terenowe nie są tu wyjątkiem. Kiedy stosujesz metodę badań terenowych, stajesz przed decyzjami dotyczącymi twojej roli obserwatora i relacji z ludźmi, których obserwujesz. Rozpatrzmy kilka kwestii związanych z tymi decyzjami.

Różne role obserwatora

W badaniach terenowych obserwatorzy mogą odgrywać każdą z kilku ról, w tym także uczestniczyć w tym, co chcą obserwować (taka była sytuacja studentów, którzy naprawiali rzeczy wykorzystywane publicznie). W tym rozdziale używamy raczej określenia *badania terenowe* zamiast często stosowanego określenia *obserwacja uczestnicząca*, ponieważ badacze terenowi nie zawsze uczestniczą w tym, co badają, chociaż zwykle będą prowadzić badania bezpośrednio w miejscu, w którym toczą się wydarzenia. Catherine Marshall i Gretchen Rossman (1995, s. 60) podkreślają:

> „Badacz może zaplanować dla siebie rolę pociągającą za sobą zróżnicowany poziom «uczestniczenia» – stopień faktycznego uczestnictwa w życiu codziennym. Z jednej strony jest pełny uczestnik, który zajmuje się życiem codziennym, przyjmując jakąś rolę lub zestaw ról stworzonych w danym otoczeniu. Z drugiej strony jest całkowity obserwator, który w ogóle nie angażuje się w interakcje społeczne, a nawet może unikać uwikłania w świat, który bada. Oczywiście wszystkie możliwe typy mieszane na tym kontinuum są dostępne badaczowi".

W tym znaczeniu pełnym uczestnikiem może być rzeczywisty uczestnik tego, co jest badane (na przykład uczestnik odbywającej się w campusie demonstracji), albo też ktoś, kto udaje, że jest rzeczywistym uczestnikiem. W każdym przypadku, gdy działasz jako pełny uczestnik, pozwalasz ludziom, by widzieli w tobie jedynie uczestnika, a nie badacza. Na przykład jeżeli badasz grupę złożoną z ludzi niewykształconych i mających trudności z wysławianiem się, nie byłoby właściwe, gdybyś mówił i zachowywał się jak profesor uniwersytetu albo student.

Ten typ badań stawia przed nami kwestię etyczną, w której sami badacze społeczni są podzieleni. Czy jest etyczne oszukiwanie ludzi, których badasz, w nadziei, że będą ci ufać tak, jak nie ufaliby osobie określonej jako badacz? Czy potencjalne korzyści z badań zrównoważą te dylematy? Mimo że wiele stowarzyszeń zawodowych zajmowało się tą kwestią, normy, które zostały wprowadzone, są ciągle raczej niejednoznaczne, gdy odnieść je do konkretnych sytuacji.

Z powyższymi rozważaniami etycznymi związane są także rozważania naukowe. Żaden badacz nie oszukuje swoich badanych wyłącznie w celu oszukiwania. Robi to raczej w przekonaniu, iż dane będą bardziej trafne i rzetelne,

że badani będą bardziej naturalni i szczerzy, kiedy nie będą wiedzieć, że badacz prowadzi projekt badawczy. Gdyby ludzie wiedzieli, że są badani, mogliby zmienić swoje zachowanie na wiele różnych sposobów. Po pierwsze, mogliby przegonić badacza. Po drugie, mogliby zmodyfikować swoje wypowiedzi i zachowania, aby wydawały się bardziej godne szacunku niż byłyby bez tego. Po trzecie, sam proces społeczny mógłby radykalnie się zmienić. Na przykład studenci planujący spalenie budynku administracji uniwersytetu mogliby całkiem zarzucić ten plan, gdyby dowiedzieli się, że jednym z członków ich grupy jest badacz społeczny prowadzący projekt badawczy.

Jednakże, jeśli jesteś pełnym uczestnikiem, możesz wpływać na to, co badasz. Grając rolę uczestnika, musisz uczestniczyć. Jednak twoje uczestnictwo może znacząco wpływać na proces społeczny, który badasz. Załóżmy na przykład, że zapytano cię o zdanie na temat tego, co grupa powinna dalej robić. Niezależnie od tego, co odpowiesz, w pewien sposób będziesz miał wpływ na toczący się proces. Jeśli grupa podąży za twoją sugestią, twój wpływ na proces jest oczywisty. Jeśli tego nie zrobi, proces, w którym odrzucono twoją sugestię, może mieć wpływ na to, co wydarzy się później. Wreszcie, jeśli powiesz, że nie wiesz, co należy zrobić, możesz zwiększyć ogólne poczucie niezdecydowania grupy i jej trudności z podejmowaniem decyzji.

Ostatecznie, cokolwiek uczestnik obserwator zrobi, lub czego nie zrobi, będzie w pewnym stopniu wpływało na to, co będzie obserwowane; jest to po prostu nieuniknione. Mówiąc bardziej poważnie, to co zrobisz lub czego nie zrobisz, może mieć istotny wpływ na to, co się wydarzy. Nie ma przed tym pełnej ochrony, jednak wrażliwość w tej kwestii może zapewnić ochronę częściową. (Wpływ ten nazywany efektem Hawthorne został obszerniej omówiony w rozdziale 8).

Ze względu na tych kilka etycznych i naukowych okoliczności badacze terenowi często wybierają inną rolę niż rola pełnego uczestnika. Możesz w pełni uczestniczyć w działaniach badanej grupy, ale wyjaśnić, że jednocześnie prowadzisz badania. Na przykład jako członek drużyny siatkówki możesz wykorzystać swoją sytuację i rozpocząć badania z socjologii sportu, informując o tym kolegów z drużyny. Jednak i w tej roli kryją się niebezpieczeństwa. Ludzie, których badasz, mogą skierować dużą część swojego zainteresowania w stronę projektu badawczego, zamiast koncentrować się na naturalnych zachowaniach społecznych, sprawiając, że obserwowany proces przestaje być typowy. Lub też odwrotnie, ty sam możesz zacząć za bardzo identyfikować się z interesami i punktem widzenia uczestników. Możesz „zasymilować się" i stracić wiele ze swojego naukowego dystansu.

Na drugim biegunie całkowity obserwator bada proces społeczny, w żaden sposób nie stając się jego częścią. Całkiem możliwe, że dzięki temu, iż badacz nie rzuca się w oczy, badani nie uświadamiają sobie, że są przedmiotem badań. Siedzenie na przystanku i obserwowanie przechodzących nieuważnie na pobliskim skrzyżowaniu jest jednym z przykładów takiego postępowania badawczego. Choć całkowity obserwator rzadziej niż pełny uczestnik wpływa na to, co jest badane i rzadziej się „asymiluje", jest on też mniej zdolny do pełnego zrozumienia tego, co bada. Jego obserwacje mogą być bardziej pobieżne i ulotne.

Fred Davis (1973) opisuje skrajne role, w jakie może wcielić się obserwator, jako „Marsjanina" i „konwertytę". Ta druga pociąga za sobą coraz głębsze wchodzenie w badane zjawisko, co stwarza ryzyko „zasymilowania". Omówimy je w następnym podrozdziale.

Aby zrozumieć podejście „Marsjanina", wyobraź sobie, że zostałeś wysłany, aby obserwować pewne nowo odkryte życie na Marsie. Prawdopodobnie będziesz postrzegał siebie jako całkowicie odrębnego od Marsjan. Część badaczy społecznych przyjmuje ten poziom odrębności, kiedy obserwuje kultury lub klasy społeczne odmienne od ich własnej.

Marshall i Rossman (1995, s. 60–61) piszą również, że badacze mogą różnić się ilością czasu, jaki spędzili w otoczeniu, które obserwowali: możesz być na scenie przez cały czas lub pokazywać się od czasu do czasu. Ponadto możesz koncentrować swoją uwagę na wybranych aspektach otoczenia społecznego lub dą-

żyć do obserwowania wszystkich elementów – przyjmując rolę właściwą ze względu na twoje cele.

Różne sytuacje wymagają od obserwatora wchodzenia w różne role. Niestety, nie ma jasnych wytycznych dla tego wyboru – musisz polegać na tym, jak ty rozumiesz sytuację i jak ją oceniasz. Jednak przy podejmowaniu decyzji musisz kierować się zarówno względami metodologicznymi, jak i etycznymi. Ponieważ często są one sprzeczne, twoje decyzje będą niejednokrotnie trudne, a czasami okaże się, że twoja rola ogranicza prowadzone badania.

Stosunek do badanych

Wprowadziwszy zróżnicowane role, jakie w związku z obserwacją może podejmować badacz terenowy, skupimy się teraz szczególnie na stosunku badacza do badanych i ich poglądów.

Odnotowaliśmy już wcześniej, że mamy możliwość wcielenia się w rolę jednostki o innym statusie społecznym, niż faktycznie posiadamy. Rozważmy teraz, co mógłbyś myśleć i czuć w takiej sytuacji.

Przypuśćmy, że zdecydowałeś się badać grupę religijną, która zwerbowała wielu ludzi z twojego sąsiedztwa. Mógłbyś badać tę grupę, przystępując do niej lub udając, że przystępujesz. Zastanów się przez chwilę, jaka jest różnica między „rzeczywistym" a „udawanym" przystąpieniem. Polega ona głównie na tym, czy faktycznie przyjmujesz wierzenia, postawy i poglądy podzielane przez „rzeczywistych" członków. Jeśli członkowie grupy wierzą, że Jezus przyjdzie w nocy w przyszły czwartek, aby zniszczyć świat i ocalić członków grupy, to czy ty w to wierzysz, czy po prostu udajesz, że wierzysz?

W takich momentach badacze społeczni tradycyjnie dążą do podkreślania wagi „obiektywności". W naszym przykładzie ów nakaz sprowadzałby się do unikania wciągnięcia w wierzenia grupy. Nie zaprzeczając zaletom tego rodzaju obiektywności, obecnie badacze społeczni widzą również korzyści uzyskiwane poprzez przyjęcie punktu widzenia tych, których się bada. John i Lyn Loflandowie (1995, s. 61) określają to „rozumieniem uczestnika". Nie będziesz w stanie w pełni zrozumieć myśli i działań członków grupy, dopóki nie przyjmiesz ich punktu widzenia jako prawdziwego – przynajmniej przejściowo. Aby w pełni docenić zjawisko, które podjąłeś się zbadać, powinieneś uwierzyć, że Jezus przyjdzie w czwartkową noc.

Przyjęcie punktu widzenia obcego jest dla większości ludzi niewygodną perspektywą. Wystarczająco trudne może okazać się samo poznanie punktów widzenia, które wydają ci się dziwne; czasami możesz mieć trudności nawet z tolerowaniem pewnych punktów widzenia. Ale przyjęcie jednego z nich za własny jest dziesięć razy gorsze. Robert Bellah (1970, 1974) zaproponował termin *realizm symboliczny* na podkreślenie potrzeby traktowania badanych wierzeń jako godnych szacunku zamiast jako obiektów drwin. Jeśli weźmiesz na serio tę perspektywę, powinieneś zrozumieć, dlaczego William Shaffir i Robert Stebbins (*Experiencing Fieldwork...* 1991, s. 1) konkludują: „praca terenowa zalicza się na pewno do najbardziej nieprzyjemnych form działalności, jakie wypracowała ludzkość".

W przyjęciu punktu widzenia badanych czai się oczywiście pewne niebezpieczeństwo. Porzucając obiektywność na rzecz tego punktu widzenia, tracisz możliwość postrzegania i rozumienia zjawiska w określonych ramach, niedostępnych badanym przez ciebie ludziom. Z jednej strony, przyjęcie wierzenia, że świat skończy się w czwartek w nocy, pozwala ci ocenić aspekty tego wierzenia dostępne jedynie wierzącym; z drugiej strony, wychodząc poza ten pogląd, masz możliwość zastanowienia się nad niektórymi pezyczynami, dla których ludzie w to wierzą. Możesz odkryć, że niektórzy przyjęli to wierzenie pod wpływem jakiegoś przeżycia traumatycznego (jak utrata pracy lub rozwód), inni zaś zostali przyprowadzeni do owczarni poprzez uczestnictwo w pewnych sieciach społecznych (na przykład cała drużyna kręglowa, do której należą, przystała do tej grupy wyznaniowej). Zauważ, że członkowie grupy mogą nie zgadzać się z takimi „obiektywnymi" wyjaśnieniami i możesz nie zbliżyć się do nich na tyle, by działać w sposób uprawniony w ramach poglądów grupy.

Powstaje tu pozorny dylemat, ponieważ obie te postawy mają istotne zalety, ale wydaje się, że wzajemnie się wykluczają. W rzeczywistości istnieje jednak możliwość ich pogodzenia. Czasami można po prostu dowolnie zmieniać punkt widzenia. Gdy jest to uzasadnione, można w pełni przyjąć wierzenia grupy: później można wyjść poza nie (a ściślej, można przyjąć punkt widzenia nauk społecznych). Kiedy staniesz się bardziej biegły w badaniach tego rodzaju, możesz dojść do utrzymywania sprzecznych punktów widzenia jednocześnie, zamiast przerzucać się z jednego na drugi i na odwrót.

Podczas swoich badań nad mediami – ludźmi, którzy pozwalają duchom zajmować ich ciała i przemawiać przez siebie – zorientowałem się, że mogę w pełni uczestniczyć w seansie spirytystycznym, bez odcinania się od klasycznych nauk społecznych. Zamiast „uwierzyć" w rzeczywistość seansu, dostrzegłem, że możliwe jest raczej zawieszenie własnych przekonań, kiedy znajdowałem się w tej sferze: nie była to ani wiara w jej prawdziwość (podzielana przez większość pozostałych uczestników), ani też niewiara (podzielana przez większość naukowców). Mówiąc inaczej, byłem otwarty na każdą możliwość. Zauważ, jak różni się to od naszej normalnej potrzeby, by „wiedzieć", czy te sprawy są uprawnione, czy też nie.

Badacze społeczni często postrzegają omówione właśnie zagadnienia jako problem *zwrotności* w tym znaczeniu, że rzeczy oddziałują same na siebie. Tak więc twoje własne cechy mogą mieć wpływ na to, co widzisz i jak to interpretujesz. Jednak ta kwestia jest szersza i odnosi się do badanych, podobnie jak do badacza. Wyobraź sobie, że przeprowadzasz wywiad z osobą bezdomną: (1) na ulicy, (2) w schronisku dla bezdomnych, (3) w biurze opieki społecznej. Otoczenie badania może wpływać na odpowiedzi. Innymi słowy, można uzyskać różne wyniki w zależności od tego, gdzie przeprowadza się wywiad. Ponadto jako badacz mógłbyś zachować się różnie w różnym otoczeniu. Jeżeli zastanowisz się nad tą sprawą, będziesz w stanie określić inne aspekty kontaktu badawczego, które komplikują zadanie „prostej obserwacji tego, co się dzieje".

Problem, który omawiamy, mógłby być postrzegany jako psychologiczny, dziejący się głównie w głowach badacza i badanych. Odpowiada mu jednak problem na poziomie społecznym. Kiedy głęboko zaangażujesz się w życie ludzi, których badasz, prawdopodobnie poruszą cię ich osobiste problemy i kryzysy. Wyobraź sobie na przykład, że jeden z członków grupy zachorował i wymaga zawiezienia do szpitala. Czy powinieneś zapewnić transport? Oczywiście, że tak. Załóżmy, że ktoś chce pożyczyć pieniądze, aby kupić sprzęt stereo. Czy powinieneś pożyczyć? Chyba nie. A jeśli na przykład potrzebują pieniędzy na jedzenie?

Nie ma czarno-białych reguł postępowania w takich sytuacjach, musisz jednak uzmysłowić sobie, że będziesz musiał zmierzyć się z nimi bez względu na to, czy ujawnisz, że jesteś badaczem, czy nie. Takie problemy zwykle nie powstają przy badaniach innego typu – na przykład przy sondażach i eksperymentach – lecz są one nieodłączną częścią badań terenowych.

To omówienie stosunku badacza do badanych kłóci się z tradycyjnym postrzeganiem „obiektywności naukowej". Lecz zanim zakończymy ten podrozdział, pójdźmy o krok dalej.

W tradycyjnym spojrzeniu na naukę ukryte są pewne różnice władzy i statusu, oddzielające badacza od badanych. Kiedy na przykład omawialiśmy schematy eksperymentalne w rozdziale 8, było oczywiste, kto kontroluje sytuację: eksperymentator. Eksperymentator organizował badanie i *mówił badanym*, co robić. Często również eksperymentator był jedyną osobą, która wiedziała, czego naprawdę dotyczy eksperyment. Dosyć podobnie można wypowiedzieć się o badaniach sondażowych. Osoba prowadząca sondaż układa pytania, decyduje, kto będzie wybrany do badań, i odpowiada za interpretację zebranych danych.

Socjologowie często postrzegają związek tego rodzaju jako relację władzy czy statusu. W projektach eksperymentalnych i sondażowych badacz po prostu dysponuje większą władzą i wyższym statusem niż badani. Badacz ma specjalną wiedzę, która nie jest dostępna badanym. Badacze nie są tak niegrzeczni, by powiedzieć, że są lepsi od badanych, ale w pew-

nym sensie istnieje takie założenie. (Zauważ, że podobne ukryte założenie istnieje w przypadku autorów i czytelników podręczników).

W badaniach terenowych takie założenie może być problematyczne. Kiedy pierwsi europejscy antropologowie rozpoczynali badania społeczeństw zwanych początkowo „prymitywnymi", nie było wątpliwości, że to antropolog wiedział lepiej. Rdzenni mieszkańcy „wierzyli" w czary, antropolog zaś „wiedział", że nie są one prawdą. Rdzenni mieszkańcy mówili o niektórych swoich rytuałach, że uspokoją one bogów, antropologowie zaś wyjaśniali, że „rzeczywistą" funkcją tych rytuałów było tworzenie tożsamości społecznej, kreowanie solidarności grupy itd.

Jednak im więcej badaczy społecznych wyruszało w teren, by bezpośrednio badać swoich bliźnich, tym bardziej stawali się oni świadomi owego ukrytego założenia o przewadze badacza i tym bardziej poszukiwali oni dla niego alternatywy. Gdy przejdziemy teraz do różnych paradygmatów badań terenowych, zobaczymy niektóre ze sposobów, w jakie na bieżąco radzono sobie z tą kwestią.

Niektóre paradygmaty jakościowych badań terenowych

Opisywałem dotychczas badania terenowe jako takie, które są po prostu prowadzone w miejscu, gdzie toczą się wydarzenia i polegają na obserwowaniu tych wydarzeń. Jednak faktycznie jest wiele różnych podejść do tej metody badawczej. Niniejszy podrozdział analizuje kilka paradygmatów badań terenowych: naturalizm, etnometodologię, teorię ugruntowaną, studium przypadku i metodę rozszerzonych przypadków, etnografię instytucjonalną i badania uczestniczące. Chociaż ten przegląd nie wyczerpuje różnorodności metod, to jednak pozwoli docenić istniejące możliwości.

Naturalizm

Naturalizm to **stara tradycja badań jakościowych**. Pierwsi **badacze terenowi działali zgodnie** z założeniem pozytywistycznym mówiącym, że rzeczywistość społeczna jest „na zewnątrz", że badacz może ją naturalnie obserwować i opisywać „taką, jaką jest naprawdę" (Gubrium, Holstein 1997). Ta tradycja zapoczątkowana została w latach 30. i 40. XX wieku na Wydziale Socjologii Uniwersytetu Chicago, którego profesorowie i studenci snuli się po mieście, obserwując lokalne grupy sąsiedzkie i wspólnoty, a następnie wyciągając wnioski z obserwacji. Badacze z tego okresu i ich podejście badawcze są obecnie często określani jako „szkoła chicagowska".

Jedną z pierwszych i najbardziej znanych prac, ilustrujących tę tradycję badawczą, jest etnografia włosko-amerykańskiego osiedla Cornerville, opisana przez Williama Foote Whyte'a w książce *Street Corner Society* (1943). **Etnografia** to badania skupione bardziej na szczegółowym i dokładnym opisie niż na wyjaśnianiu. Podobnie jak inni naturaliści, Whyte był przekonany, że aby w pełni poznać życie społeczne ulicy, musi starać się stać się jego uczestnikiem. Nawiązał kontakt z „Docem", informatorem, który okazał się jednym z przywódców gangu ulicznego. „Doc" umożliwił Whyte'owi wejście w swój świat i Whyte mógł uczestniczyć w działaniach ludzi z Cornerville. Studium Whyte'a przyniosło coś, czego nie mogły dać sondaże: bogaty w szczegóły obraz życia wśród włoskich imigrantów Cornerville.

Ważną cechą studium Whyte'a jest to, że opisuje on rzeczywistość ludzi z Cornerville w ich kategoriach. Podejście naturalistyczne opiera się na opowiadaniu „ich" historii tak, jak one „naprawdę są", nie zaś tak, jak rozumie „je" etnograf. Narracje zebrane przez Whyte'a zostały na podstawie ich fasadowej wartości przyjęte jako społeczna „prawda" o mieszkańcach Cornerville.

Czterdzieści lat później David A. Snow i Leon Anderson (1987) przeprowadzili eksploracyjne badania terenowe, dotyczące życia bezdomnych w Austin w Teksasie. Głównym zadaniem badaczy było zrozumienie, jak bezdomni konstruują i negocjują swoją tożsamość, wiedząc, że społeczeństwo, w którym żyją, piętnuje bezdomność. Snow i Anderson byli przekonani, że dla tego celu zbiór danych musiał powstać

naturalnie. Podobnie jak Whyte w *Street Corner Society*, odnaleźli kilku kluczowych informatorów, którym towarzyszyli w ich codziennych peregrynacjach, związanych z dniem pracy w miejscach zbierania śmieci lub pod mostami. Snow i Anderson zdecydowali się zapamiętywać rozmowy, w których uczestniczyli, lub „gadki", jakie bezdomni prowadzili między sobą. Na koniec dnia obaj badacze relacjonowali wszystkie zaobserwowane rozmowy i sporządzali szczegółowe notatki. Badacze nagrywali także wywiady pogłębione ze swoimi kluczowymi informatorami.

Snow i Anderson opisali swój pobyt wśród bezdomnych w trakcie 12 miesięcy w sumie przez 405 godzin w 24 różnych miejscach. Na podstawie tego bogatego zbioru danych zidentyfikowali trzy wzorce występujące w konwersacjach bezdomnych. Po pierwsze, bezdomni wykazywali próby zachowania „dystansu" wobec innych bezdomnych, wobec pracy o niskim statusie, jaką aktualnie mieli, czy wobec Armii Zbawienia, od której byli zależni. Po drugie, „przyjmowali" swoją tożsamość jako mieszkańców ulicy, członkostwo w grupie, czy pewne wierzenia na temat powodów, dla których są bezdomni. Po trzecie, opowiadali „fikcyjne historyjki", które zawsze kontrastowały z ich życiem codziennym. Na przykład często mówili, że zarabiają więcej pieniędzy, niż zarabiali faktycznie, lub nawet, że „będą bogaci".

Etnometodologia

Etnometodologia, o której mówiłem już w rozdziale 2 jako o paradygmacie badawczym, jest bardzo odmiennym podejściem do jakościowych badań terenowych. Ma ona swoje korzenie w filozoficznej tradycji fenomenologii, co może tłumaczyć, dlaczego etnometodologowie sceptycznie podchodzą do sposobu, w jaki ludzie opisują swoje doświadczenia na temat rzeczywistości (Gubrium, Holstein 1997). Alfred Schutz (1967, 1970), który zapoczątkował fenomenologię, przekonywał, że rzeczywistość jest społecznie konstruowana, a nie jest „poza" nami, co umożliwiałoby obserwację. Ludzie opisują swój świat nie „takim, jaki on jest", lecz „takim, jakie oni mu nadają znaczenie". Tak więc fenomenolog mógłby argumentować, że ludzie z ulicy Whyte'a opisywali swoje życie w gangu w sposób oddający przypisywane mu przez nich znaczenie. Ich opisy nie powiedziałyby nam jednak, jak i dlaczego nadawali mu oni określone znaczenie. Z tego powodu badacze nie mogą polegać na opowieściach badanych, jeśli chcą odtworzyć rzeczywistość społeczną.

Tradycyjni etnografowie wierzyli, że zanurzanie się w konkretną kulturę oraz opowieści ich informatorów przedstawiają rzeczywistość, fenomenologowie natomiast widzą potrzebę „nadawania sensu" postrzeganiu świata przez informatorów. Idąc za tą tradycją, część badaczy terenowych uznaje za konieczne opracowywanie technik ujawniania tego, w jaki sposób ludzie nadają sens swojemu codziennemu światu. Jak widzieliśmy w rozdziale 2, socjolog Harold Garfinkel sugerował, że badacz *łamie reguły*, dzięki czemu ujawniane są oczekiwania ludzi, traktowane dotąd jako oczywiste. Jest to technika, którą Garfinkel nazwał „etnometodologią".

Garfinkel stał się znany z angażowania swoich studentów do wykonywania serii eksperymentów, zwanych przez niego „przełamującymi", zaplanowanych z myślą o oderwaniu się od zwyczajności (Heritage 1984). Na przykład Garfinkel (1967) poprosił swoich studentów o wykonanie „eksperymentu wyjaśniającego konwersację". Studenci mieli zaaranżować zwyczajną rozmowę ze znajomym lub z przyjacielem i prosić o wyjaśnianie wszelkich stwierdzeń drugiej osoby. Za pomocą tej techniki ujawniono elementy konwersacji, które normalnie przyjmuje się za oczywiste. Oto dwa przykłady relacji studentów Garfinkla (1967, s. 42):

> **Przypadek 1**
>
> „Badana opowiada eksperymentatorowi, członkowi tej samej wspólnoty użytkowników samochodów, o tym, jak złapała gumę jadąc do pracy poprzedniego dnia:
>
> Złapałam gumę.
>
> (E) Co masz na myśli, mówiąc złapałam gumę?
>
> Przez chwilę wydawała się zaskoczona. Następnie odpowiedziała wrogo: «Co masz na myśli, pytając 'Co masz na myśli?'. Guma to guma. Oto, co mam na myśli. Nic specjalnego. Co za głupie pytanie»".

Przypadek 6

„Ofiara przyjaźnie macha ręką.

(S) Co u ciebie słychać?

(E) Co słychać w związku z czym? Z moim zdrowiem, finansami, szkołą, spokojem ducha...?

(S) (Zrobił się czerwony na twarzy i nagle stracił panowanie nad sobą). Słuchaj, ja tylko starałem się być grzeczny. Cholera, szczerze mówiąc wcale mnie nie obchodzi, co u ciebie słychać".

Przez odkładanie lub „branie w nawias" własnych oczekiwań, dotyczących tych codziennych rozmów, eksperymentatorzy uwidaczniają subtelności codziennych interakcji. Na przykład, chociaż „co słychać" ma wiele możliwych znaczeń, nikt z nas nie ma trudności ze zrozumieniem, co to oznacza w okazjonalnych kontaktach, tak jak niczego nie podejrzewający badany ujawnił to w swoim końcowym komentarzu.

Etnometodologowie nie interesują się więc po prostu tym, jak badani postrzegają świat. W przedstawionych przykładach moglibyśmy sobie wyobrazić, że badani uznali eksperymentatorów za niegrzecznych, głupich czy aroganckich. To sama konwersacja, a nie informator, staje się przedmiotem studiów etnometodologicznych. W zasadzie w etnometodologii uwaga zwrócona jest na „ukryte wzorce" interakcji, które regulują nasze życie codzienne.

Etnometodologowie uważają, że badacze stosujący analizę naturalistyczną „(tracą) zdolność do analizowania świata potocznego i jego kultury, jeżeli (używają) narzędzi analitycznych i intuicji, które same są częścią badanego świata lub kultury" (Gubrium, Holstein 1997, s. 43). Lawrence Wieder daje wspaniały przykład, jak różne od podejścia etnometodologicznego jest podejście naturalistyczne (Gubrium, Holstein 1997). W swoim studium *Language and Social Reality: The Case of Telling the Convict Code* (1988) rozpoczął on od tradycyjnego etnograficznego podejścia do skazanych w ośrodku przystosowawczym dla zwalnianych więźniów: zamierzał stać się członkiem grupy, nawiązując przyjazne stosunki ze współwięźniami i prowadząc obserwację uczestniczącą. Pierwszym zadaniem, jakie sobie postawił, było opisanie życia pensjonariuszy takim, jakie ono „naprawdę było" dla nich. Obserwacje Wiedera pozwoliły mu opisać „kodeks skazańca", który był jego zdaniem źródłem zachowań dewiacyjnych, prezentowanych przez współwięźniów wobec personelu. Kodeks ten składał się z szeregu reguł, takich jak „nie całuj w dupę", „nie kapuj", „nie ufaj personelowi". Były one przestrzegane przez współwięźniów, którzy niweczyli usiłowania personelu, dotyczące ułatwienia im powrotu z więzienia do społeczeństwa.

Dla Wiedera stało się oczywiste, że kodeks był czymś więcej niż narzędziem wyjaśnienia dewiacyjnych zachowań skazanych; była to „metoda perswazji moralnej i uzasadniania" (Wieder 1988, s. 175). Od tego momentu zamienił on podejście naturalistyczne na etnometodologiczne. Badacze terenowi, związani z naturalizmem, dążą do zrozumienia życia społecznego tak, jak pojmują je jego uczestnicy, etnometodologowie zaś skupiają się raczej na określaniu metod, za pomocą których dochodzi do zrozumienia. Wieder spostrzegł, że skazani używają kodeksu, by nadać znaczenie swym interakcjom z innymi skazanymi oraz z personelem. Etnografia ośrodka przekształciła się w ten sposób w etnografię kodeksu. Na przykład skazani powiedzieliby „wiesz, że ja nie sypię", odwołując się do kodeksu jako uzasadnienia odmowy odpowiedzi na pytanie Wiedera (s. 168). Według Wiedera kodeks „działa jak narzędzie do przerywania lub zmiany tematu rozmowy" (s. 175). Nawet personel odwołuje się do kodeksu, by uzasadnić swoją niechęć do pomocy skazanym. Choć więc kodeks był czymś, co wymuszało określone zachowania, funkcjonował on także jako narzędzie kontroli interakcji.

Teoria ugruntowana

Teoria ugruntowana (*grounded theory*) powstała w trakcie współpracy Barneya Glasera i Anselma Straussa, socjologów, którzy połączyli ze sobą dwie główne tradycje badawcze: pozytywizm i interakcjonizm. W zasadzie **teoria ugruntowana** jest próbą wyprowadzania teorii z analizy wzorców, tematów i wspólnych kategorii, ujawnionych w danych obserwacyjnych. Pierwsza główna prezentacja tej metody znaj-

tuje się w książce Glasera i Straussa *The Discovery of Grounded Theory* (1967). Teorię ugruntowaną można przedstawić jako próbę łączenia podejścia naturalistycznego z pozytywistyczną troską o „systematyczny zestaw procedur" przy przeprowadzaniu badań jakościowych.

Strauss i Juliet Corbin (1990, s. 44–46) sugerują, że teoria ugruntowana pozwala badaczowi być jednocześnie i naukowym, i twórczym, jeśli tylko badacz stosuje się do trzech wskazówek:

> „**Raz na jakiś czas cofaj się o krok i pytaj:** O co tu chodzi? Czy to, co myślę, że widzę, pasuje do rzeczywistości zapisanej w danych? Dane same z siebie nie kłamią (...).
>
> **Zachowaj sceptyczną postawę.** Wszystkie wyjaśnienia teoretyczne, kategorie, hipotezy, pytania dotyczące danych, niezależnie od tego, czy pochodzą bezpośrednio, czy pośrednio z porównań, z literatury czy z doświadczenia, powinny być traktowane jako tymczasowe. Zawsze wymagają sprawdzenia na podstawie rzeczywistych danych i nigdy nie mogą być przyjmowane jako fakt (...).
>
> **Stosuj procedury badawcze.** Procedury zbierania i analizy danych są zaprojektowane tak, by nadać badaniom walor ścisłości. Jednocześnie pomagają likwidować obciążenia i prowadzą do sprawdzenia przynajmniej niektórych twoich założeń, które mogłyby wpływać na nierealistyczne odczytywanie danych".

Teoria ugruntowana kładzie nacisk na procedury badawcze. Szczególnie ważne dla osiągnięcia trafności i rzetelności w analizie danych jest systematyczne kodowanie. Ze względu na to nieco pozytywistyczne traktowanie danych reprezentanci teorii ugruntowanej są bardzo otwarci na stosowanie badań jakościowych łącznie z badaniami ilościowymi. Oto dwa przykłady zastosowania tego podejścia.

Badanie zmiany akademickiej

Studium zmiany akademickiej na uniwersytetach wykonane przez Cliftona F. Conrada (1978) jest jednym z pierwszych przykładów wykorzystania teorii ugruntowanej. Conrad liczył na odkrycie głównych źródeł zmian w akademickich programach nauczania i zarazem na wyjaśnienie procesu zmiany. Używając pochodzącego z teorii ugruntowanej pomysłu *teoretycznego doboru próby*, zgodnie z którym grupy czy instytucje są dobierane pod kątem ich teoretycznej doniosłości, Conrad dobrał cztery uniwersytety do swoich badań. Na dwóch z nich głównym motorem zmian był formalny komitet ds. programów, na pozostałych dwóch motorem zmiany była nieformalna grupa.

Conrad wyjaśniał krok po kroku zalety zastosowania teorii ugruntowanej do budowania jego teorii zmiany akademickiej. Opisał proces systematycznego kodowania danych w celu stworzenia kategorii, które musiały „wyłonić się" z danych, a następnie określał wzajemne dopasowanie tych kategorii. Stałe przechodzenie od danych do teorii i od teorii do danych pozwoliło mu na ponowną ocenę trafności początkowych wniosków.

Na przykład, pierwotnie wydawało się, że zmiana akademicka jest powodowana głównie przez administratora, który naciska na jej przeprowadzenie. Ponownie analizując dane i poszukując bardziej przekonującego wyjaśnienia, Conrad wskazał na grupy interesu jako bardziej przekonujące źródło zmian. Pojawienie się takich grup interesu pozwoliło faktycznie administratorowi zostać agentem tych zmian.

Ocena, jak dane z każdego z dwóch typów uniwersytetów przystają do drugiego typu, pomogła udoskonalić budowanie teorii. Konkluzja Conrada była taka, że zmiany w programach uniwersyteckich bazują na następującym procesie: konflikt i grupy interesu wyłaniają się pod wpływem nacisku wewnętrznych i zewnętrznych sił społecznych; naciskają one na dokonanie interwencji administracyjnej i zalecenie przez administrację dokonania zmian w obecnych programach akademickich; zmiany te są następnie wprowadzane przez najsilniejsze ciało decyzyjne.

Rumunia na zakupach

Wiele napisano o ogromnych zmianach wywołanych przejściem od komunizmu do kapitalizmu w byłym ZSRR i u jego wschodnioeuropej-

skich sojuszników. Patrick C. Jobes i współpracownicy (1997) chcieli dowiedzieć się, jak wyglądało to przejście w mniejszej skali, wśród przeciętnych Rumunów. Skoncentrowali się na robieniu zakupów.

Zauważając, że zakupy są zwykle postrzegane jako rutynowe, dość racjonalne działanie, badacze sugerowali, że mogły one stać się problemem społecznym w radykalnie zmieniającej się gospodarce. Użyli metody teorii ugruntowanej w celu przebadania zakupów Rumunów jako problemu społecznego, poszukując sposobów, w jakie zwykli ludzie rozwiązują ten problem.

Pierwszym zadaniem badaczy było sprawdzenie, jak Rumuni postrzegają i rozumieją zadanie, jakim jest robienie zakupów. Badacze – uczestnicy zajęć poświęconych problemom społecznym – rozpoczęli od przeprowadzenia wywiadów z 40 osobami robiącymi zakupy. Pytano je, czy doświadczyły problemów związanych z robieniem zakupów i jakie działania podjęły, aby się z nimi uporać.

Gdy zakończono pierwsze wywiady, badacze przejrzeli dane w poszukiwaniu kategorii odpowiedzi – najbardziej powszechnych problemów osób robiących zakupy i ich rozwiązań. Jednym z częstszych problemów był brak pieniędzy. Doprowadziło to badaczy do pierwszej hipotezy roboczej: „pozycja społeczno-ekonomiczna osób robiących zakupy jest powiązana z tym, jak postrzegają one problemy i jak poszukują rozwiązań" (1997, s. 133). Ta i inne hipotezy pomogły badaczom skoncentrować uwagę na bardziej określonych zmiennych w kolejnych wywiadach.

Kontynuując, badacze poszukiwali możliwości badania innych typów osób robiących zakupy. Podczas wywiadów ze studentami zauważono na przykład, że różne typy osób robiących zakupy są zainteresowane różnymi typami dóbr, co z kolei wpływa na problemy, przed którymi stoją, i na rozwiązania, jakie wypróbowują.

Kiedy w wyniku kontynuowania wywiadów wypracowano dodatkowe hipotezy, badacze rozpoczęli tworzenie mniej lub bardziej zestandaryzowanego zestawu pytań zadawanych osobom robiącym zakupy. Początkowo wszystkie pytania były otwarte, lecz w końcu stworzono również pytania zamknięte.

Badania te są przykładem kluczowych, indukcyjnych zasad teorii ugruntowanej: dane zbierane są bez postawienia hipotez. Wstępnych danych używa się do ustalenia kluczowych zmiennych, tak jak postrzegają je badani, podobnie też hipotezy na temat związków między zmiennymi są wyprowadzane z zebranych danych. Dalsze zbieranie danych przynosi lepsze rozumienie i z kolei poprawia się jakość samego zbierania danych.

Studium przypadku i metoda rozszerzonych przypadków

Badacze społeczni często mówią o **studium przypadku**, które skupia się na jednym lub kilku przykładach pewnego zjawiska społecznego, takiego jak wioska, rodzina, gang nastolatków. Charles Ragin i Howard Becker (1992) podkreślają, że panuje bardzo ograniczona zgoda na temat tego, co może tworzyć „przypadek", i termin ten jest stosowany szeroko. Badanym przypadkiem może być na przykład raczej jakiś okres, a nie jakaś określona grupa ludzi. Ograniczenie uwagi do konkretnego przykładu czegoś jest zasadniczą cechą studium przypadku.

Głównym celem studium przypadku może być opis, jak na przykład wówczas, gdy antropolog opisuje kulturę plemienia przedpiśmiennego. Natomiast pogłębione studium konkretnego przypadku może przynieść wyjaśnienia podobne do tych, jakie uzyskali badacze społeczności Helen i Robert Lyndowie (1929, 1937) oraz W. Lloyd Warner (1949), gdy próbowali zrozumieć strukturę i stratyfikację społeczną w małych miastach w Stanach Zjednoczonych.

Badacze stosujący studium przypadku mogą poszukiwać jedynie idiograficznego zrozumienia konkretnego badanego przypadku, lub – jak widzieliśmy w przypadku teorii ugruntowanej – studia przypadków mogą tworzyć podstawę do budowania ogólniejszych teorii nomotetycznych.

Michael Burawoy z zespołem (*Ethnography Unbound...* 1991) sugerują nieco inną relację między studium przypadku a teorią. Dla nich metoda rozszerzonych przypadków ma na celu

odkrywanie słabych punktów, a następnie modyfikowanie istniejących teorii społecznych. To podejście różni się istotnie od niektórych spośród dotychczas omówionych.

Przedstawiciele teorii ugruntowanej poszukują możliwości wyjścia w teren bez koncepcji wstępnych, opisujących to, co zamierzają odkryć, sugestie Burawoya natomiast są odwrotne: starać się „wyłożyć tak spójnie, jak to możliwe, to, co spodziewamy się odkryć w badanym miejscu *przed* wyjściem w teren (*Ethnography Unbound...* 1991, s. 9). Burawoy widzi metodę rozszerzonych przypadków jako sposób przebudowania lub udoskonalenia teorii zamiast jej potwierdzania lub odrzucania. Tak więc szuka on wszystkich przejawów konfliktu obserwacji z istniejącymi teoriami i tym, co nazywa „teoretycznymi lukami i przemilczeniami" (1991, s. 10). Taki stosunek do badań terenowych sprawia, że znajomość literatury przed podjęciem badań jest dla Burawoya i jego współpracowników koniecznością, natomiast zwolennicy teorii ugruntowanej byliby zmartwieni tym, że znajomość konkluzji innych badaczy mogłaby wpływać obciążająco na ich własne obserwacje i teorie.

Dla zilustrowania metody rozszerzonych przypadków posłużę się dwoma przykładami badań studentów Burawoya.

Negocjacje nauczyciel–uczeń

Leslie Hurst (1991) przystąpiła do badania wzorców interakcji między nauczycielami a uczniami pewnej szkoły średniej. Hurst weszła w teren uzbrojona w istniejące, sprzeczne teorie dotyczące „oficjalnych" funkcji szkoły. Niektóre z nich sugerowały, iż celem szkoły jest wspieranie mechanizmów ruchliwości społecznej, natomiast inne utrzymywały, że szkoła głównie odtwarza *status quo* w formie warstwowego podziału pracy. Oficjalne role przypisane nauczycielom i uczniom mogłyby być interpretowane w kategoriach każdego z tych ujęć.

Hurst uderzył jednak kontrast między tymi teoriami a typami interakcji, jakie obserwowała w klasie. W jej własnym doświadczeniu jako uczennicy nauczyciele dysponowali pełnią władzy nad umysłami, ciałami i duszami swoich uczniów. Ona zaś obserwowała coś zupełnie innego w szkole na osiedlu zamieszkanym przez niższą klasę średnią, w Berkeley w Kalifornii – Emerald Junior High School, gdzie była nauczycielem wolontariuszem. Wyjaśniała ona charakter zaobserwowanych negocjacji nauczyciel–uczeń, koncentrując się na rozdzieleniu funkcji szkoły, nauczyciela i rodziny.

Zgodnie z obserwacjami Hurst funkcją wypełnianą przez szkołę było kontrolowanie „ciał" uczniów – na przykład przez regulowanie ogólnych zasad ich poruszania się i działań wewnątrz szkoły. „Umysły" uczniów miały być kształtowane przez nauczycieli, natomiast rodziny uczniów ponosiły odpowiedzialność z ich „dusze"; tak więc oczekiwano, iż rodziny zajmą się socjalizacją uczniów w zakresie osobistych wartości, postaw, poczucia własności oraz dobrych obyczajów. Kiedy uczniowie przychodzili do szkoły niewyposażeni w te wartości, nauczyciel, zdaniem Hurst, „musi najpierw negocjować z uczniami pewien kompromis co do tego, jak uczniowie będą się zachowywać i co będzie uznawane w klasie za dobre obyczaje" (1991, s. 185).

Hurst objaśnia stałe negocjacje między nauczycielami a uczniami jako przejaw podziału na „ciało", które przypisane jest szkole, i „duszę" jako domenę rodziny. Nauczyciele, którzy mieli ograniczone uprawnienia do kontrolowania umysłów uczniów w klasie, używali form negocjacji z uczniami, by mogli „kontrolować (...) ciało ucznia i jego poczucie własności" (1991, s. 185) lub, jak nazywa to Hurst, niańczyć ciało i duszę ucznia.

Hurst mówi, że różni się od tradycyjnych ujęć socjologicznych w następujący sposób:

> „Nie patrzę na szkoły w sposób futurystyczny. Nie postrzegam szkół w kategoriach szkolenia, socjalizacji lub wciskania ludzi w przyszłe hierarchie. Podchodzenie do szkół w ten sposób powoduje przeoczenie negocjacyjnych, chaotycznych aspektów funkcjonowania klasy i doświadczenia edukacyjnego. Perspektywa futurystyczna dąży do nałożenia porządku i celu na doświadczenia szkoły, tracąc z oczu jej codzienną rzeczywistość".
>
> (1991, s. 186)

Podsumowując, ze studium Hurst wyłania się próba poprawienia tradycyjnego socjologicznego rozumienia edukacji przez dodanie idei, że klasa szkolna, szkoła i rodzina mają oddzielne funkcje, które z kolei mogą wyjaśniać powstanie w klasie „negocjowanego porządku".

Walka z AIDS

Katherine Fox (1991) przystąpiła do badania agencji, której celem była walka z epidemią AIDS przez dostarczanie prezerwatyw i substancji sterylizujących do czyszczenia igieł dla dożylnych narkomanów. Jest to dobry przykład odkrywania ograniczeń powszechnie stosowanych modeli wyjaśniania teoretycznego w dziedzinie rozumienia dewiacji – a konkretnie „modelu terapii" przewidującego, iż narkomani będą przybywać do kliniki i prosić o leczenie. Interakcje Fox z pracownikami socjalnymi agencji – większość z nich była częścią społeczności uzależnionych od narkotyków lub byłych prostytutek – były zaprzeczeniem tego modelu.

Na początku konieczne było zrozumienie subkultury narkomanów i wykorzystanie tej wiedzy do przygotowania bardziej realistycznych sposobów postępowania i programów. Docelowi użytkownicy musieli na przykład zostać przekonani, że pracownicy związani z programem są godni zaufania, że naprawdę są zainteresowani jedynie dostarczaniem środków odkażających i prezerwatyw. Docelowi użytkownicy musieli być pewni, że nie zostaną aresztowani.

Badania terenowe Fox nie ograniczały się do badania zażywających narkotyki. Badała również pracowników agencji, odkrywając, że program pracy socjalnej z zakażonymi miał odmienne znaczenie dla kierowników badań i dla pracowników socjalnych. Niektórzy z wolontariuszy, którzy faktycznie dostarczali środków odkażających i prezerwatyw, czuli się sfrustrowani z powodu tylko niewielkich zmian, jakich mogli dokonać. Wielu myślało, że program jest jedynie namiastką pomocy chorym na AIDS i rozwiązywania problemu uzależnienia od narkotyków. Niektórzy byli niezadowoleni z konieczności sporządzania notatek terenowych. Z kolei kierownicy potrzebowali sprawozdań i notatek terenowych, gdyż dzięki nim mogli uprawomocnić swoje badania w oczach federalnych i stanowych agencji finansujących projekt. Studium Fox pokazało, jak projekt badań nad AIDS ugrzązł w biurokratycznej inercji typowej dla organizacji formalnych: ich celem stało się podtrzymywanie własnego istnienia.

Oba badania ilustrują, jak działa metoda rozszerzonych przypadków. Badacz wkracza w teren, dysponując pełną wiedzą o istniejących teoriach, ale dążąc do odkrycia sprzeczności, które wymagają modyfikacji tych teorii.

Etnografia instytucjonalna

Etnografia instytucjonalna jest podejściem opracowanym pierwotnie przez Dorothy Smith (1978) w celu lepszego zrozumienia codziennych doświadczeń kobiet poprzez ujawnienie relacji władzy, które je kształtują. Obecnie metoda ta została rozszerzona na ideologie, które kształtują doświadczenia każdego uciskanego podmiotu.

Smith i inni socjologowie uważają, że jeśli badacz zapyta kobiety lub innych członków podporządkowanych grup o to, „jak to się dzieje", może ujawnić instytucjonalne zwyczaje, które kształtują ich rzeczywistość (M. L. Campbell 1998; D. Smith 1978). Celem takich badań jest ujawnienie form ucisku, które często zostają przeoczone w bardziej tradycyjnych typach badań.

Metodologia Dorothy Smith jest podobna do etnometodologii w tym, że to nie sami badani znajdują się w centrum badań. Etnograf instytucjonalny zaczyna od osobistych doświadczeń jednostek, ale dąży do odkrycia instytucjonalnych relacji władzy, które strukturyzują te doświadczenia i kierują nimi. W procesie tym badacz może odkrywać te aspekty społeczeństwa, które mogłyby zostać przeoczone przez badania, które za punkt wyjścia miałyby oficjalne cele instytucji.

To podejście łączy „poziom mikro" codziennych doświadczeń osobistych z „poziomem makro" instytucji. M. L. Campbell ujął to następująco:

> „Etnografia instytucjonalna, podobnie jak inne formy etnografii, opiera się na wywiadach,

obserwacjach i dokumentach jako na źródłach danych. Etnografia instytucjonalna odbiega od innych podejść etnograficznych poprzez traktowanie tych danych nie jako tematów lub przedmiotów zainteresowania, lecz jako sposobów «wejścia» w relacje społeczne w danym otoczeniu. Ta idea polega na uzyskaniu dostępu do specjalistycznej wiedzy ludzi".

(1998, s. 57)

Oto dwa przykłady tego podejścia.

Matkowanie, nauka szkolna i rozwój dziecka

Pierwszym przykładem etnografii instytucjonalnej są badania Alison Griffith (1995). Wspólnie z Dorothy Smith zgromadziła ona dane o relacjach między matkowaniem, nauką szkolną a rozwojem dzieci. Griffith rozpoczęła od przeprowadzenia wywiadów z matkami z trzech miast z południowego Ontario na temat ich codziennej pracy nad budowaniem relacji między rodzinami a szkołą. Był to punkt wyjścia dla kolejnych wywiadów z rodzicami, nauczycielami, pracownikami administracji szkolnej, pracownikami socjalnymi, psychologami szkolnymi i z pracownikami administracji centralnej.

Odkrycia Griffith wyjaśniły, jak dyskurs na temat matkowania przesunął w czasie swój punkt ciężkości z interakcji między matką a dzieckiem ku „dzieckocentrycznym" zaleceniom. Autorka dostrzegała wyraźne podobieństwa w dyskursie prowadzonym przez szkoły, środki przekazu (czasopisma i programy telewizyjne), państwo i specjalistów zajmujących się rozwojem dzieci.

Nauczyciele i specjaliści od rozwoju dzieci postrzegali rolę matek w kategoriach współpracy między matkami a szkołami, koniecznej, by dzieci odnosiły sukcesy nie tylko w szkole, ale również w życiu. Ze względu na nierówne zasoby nie wszystkie matki uczestniczyły w dyskursie na temat „właściwego" rozwoju dzieci w taki sam sposób. Griffith odkryła, że matki z klasy robotniczej były postrzegane jako słabiej „stymulujące" wysiłek związany z nauką szkolną, niż matki z klasy średniej. Griffith dowodzi, że ten dyskurs dotyczący rozwoju dzieci, wbudowany w instytucje szkolne, utrwala reprodukcję klasową poprzez uczynienie ideałów klasy średniej, dotyczących relacji między rodziną a szkołą, normą ogólną.

Przymusowa heteroseksualność

Drugim przykładem etnografii instytucjonalnej są podjęte przez Didi Khayatt (1995) badania, dotyczące instytucjonalizacji przymusowej heteroseksualności w szkołach i jej wpływu na uczennice lesbijki. W 1990 roku Khayatt rozpoczęła swoje badania od przeprowadzenia wywiadów z 12 lesbijkami w wieku od 15 do 24 lat z Toronto. Wychodząc od punktu widzenia tych młodych kobiet, poszerzyła następnie swoje badania o innych uczniów, nauczycieli, doradców i pracowników administracyjnych.

Khayatt odkryła, że szkolne zwyczaje administracyjne generowały *przymusową heteroseksualność*, która wytworzyła poczucie marginalizacji i nadwrażliwości wśród uczennic lesbijek. Szkoła nie karała na przykład dręczenia i przezywania uczniów homoseksualnych. Kwestia homoseksualności została wykluczona z programu, żeby nie wydawała się uczniom alternatywą wobec heteroseksualności.

W obu tych przykładach opisałem badania rozpoczęte z punktu widzenia kobiet matek i uczennic lesbijek. Jednak zamiast podkreślania punktu widzenia badanych, obie analizy skupiają się na relacjach władzy, które kształtują doświadczenia tych kobiet i ich rzeczywistość.

Badania uczestniczące

Ostatni paradygmat badań terenowych pozwala nam wrócić do naszego wcześniejszego omówienia relacji statusu i władzy, łączących badaczy z badanymi. W ramach paradygmatu **badań uczestniczących** (*participatory action research* – PAR) funkcją badacza jest służenie badanym – na ogół grupom upośledzonym – jako środek skutecznego działania w ich własnym interesie. Dyskryminowane osoby definiują swoje problemy, definiują pożądane środki zaradcze i obejmują kierownictwo w projektowaniu badania, które pomoże im zrealizować ich cele.

Początki tego podejścia sięgają rozwoju badań Trzeciego Świata, lecz szybko rozpowszechniło się ono w Europie i Ameryce Północnej (Gaventa 1991). Wywodzi się ono z ożywionej krytyki klasycznych badań społecznych. Tradycyjne badania, zgodnie z paradygmatem PAR, są postrzegane jako „model elitarystyczny" (Whyte, Greenwood, Lazes 1991), który redukuje badanych do „przedmiotu" badań. Zgodnie z wieloma orędownikami tej perspektywy rozróżnienie badacza i badanego powinno zniknąć. Dowodzą oni, że badani, których będą dotyczyły badania, powinni być również odpowiedzialni za ich kształt.

W tym podejściu tkwi ukryte przekonanie, że funkcją badań jest nie tylko wytwarzanie wiedzy, lecz jest ono także „narzędziem edukacji i rozwoju świadomości oraz mobilizowania do działania" (Gaventa 1991, s. 121–122). Zwolennicy badań uczestniczących zrównują dostęp do informacji z władzą i dowodzą, że ta władza jest utrzymywana w rękach dominującej klasy, płci, grupy etnicznej lub narodowej. Gdy ludzie postrzegają siebie jako badaczy, automatycznie odzyskują władzę nad wiedzą.

Przykłady tego podejścia obejmują badania struktur władzy w społecznościach, badania korporacyjne i ruchy typu „prawo do wiedzy" (Whyte, Greenwood, Lazes 1991). Oto dwa przykłady badań korporacyjnych, w których zostało użyte podejście PAR.

Korporacja Xerox

Badania uczestniczące przeprowadzono w korporacji Xerox z inicjatywy zarówno kierownictwa, jak i przywódców związkowych. Celem kierownictwa było obniżenie kosztów tak, żeby przedsiębiorstwo mogło rozwijać się na rynku o rosnącej konkurencyjności. Związki sugerowały nieco szerszy zakres: poprawę warunków pracy z jednoczesnym obniżeniem kosztów produkcji i wzrostem produktywności.

Menedżerowie spółki rozpoczęli od skupienia się na problemach na poziomie sklepów; mniej zajmowali się sprawami umów o pracę i problematyczną polityką kierownictwa. W tym czasie kierownictwo miało plan wprowadzenia programu „outsourcingu", który doprowadziłby do zwolnienia 180 robotników, związek zaś zaczął organizować sprzeciw wobec tego planu. Peter Lazes, konsultant wynajęty przez Xeroxa, spędził pierwszy miesiąc przekonując kierownictwo i związki do stworzenia „zespołu badania kosztów" (ZBK), w skład którego wchodziliby robotnicy z wydziału wytwarzającego wiązki przewodów.

Ośmiu pełnoetatowych robotników zostało wyznaczonych do ZBK na sześć miesięcy. Ich zadaniem było zbadanie możliwości przeprowadzenia zmian, które pozwoliłyby przedsiębiorstwu zaoszczędzić 3,2 miliona dolarów i zachować 180 miejsc pracy. Zespół miał dostęp do wszystkich informacji finansowych i był uprawniony do rozmawiania ze wszystkimi w przedsiębiorstwie. Ta strategia pozwoliła robotnikom na przedstawianie sugestii dotyczących sfery, która zwykle była im niedostępna. Jak pisze Whyte ze współpracownikami, „zmiana reguł gry umożliwiła ZBK wezwanie zarządzających do wyjaśnienia i uzasadnienia wszystkich świadczeń pracowniczych" (1991, s. 27). Dzięki zmianom zasugerowanym przez ZBK i wprowadzonym przez zarząd spółka zgodnie z zamierzeniami zaoszczędziła 3,2 miliona dolarów.

Kierownictwo było tak zadowolone z tych rezultatów, że rozszerzyło projekt ZBK z wydziału wiązek przewodów na trzy inne wydziały zagrożone przez konkurencję. Tam również kierownictwo było zadowolone z oszczędności, jakie udało się wypracować dzięki zespołom robotników.

Studium przypadku spółki Xerox jest interesującym przykładem badań uczestniczących, ponieważ pokazuje, że generowanie wiedzy nie zawsze musi być przedsięwzięciem elitarnym. „Eksperci" niekoniecznie muszą być fachowcami. Jak wskazują Whyte i współpracownicy, „u Xeroxa badania uczestniczące wywołały intensywny proces organizacyjnego uczenia się i pokierowały tym procesem – procesem, dzięki któremu przywódcy robotników i zarząd uczyli się od siebie wzajemnie oraz od konsultanta/animatora, on zaś uczył się od nich" (1991, s. 30).

Projekt Mondragón-Cornell

Projekt badań uczestniczących Mondragón-Cornell opiera się kompleksie spółdzielczym

Mondragón w hiszpańskim Kraju Basków. W skład kompleksu wchodzą spółdzielnie produkcyjne, spółdzielnia spożywców, bank spółdzielczy, spółdzielcza organizacja badań i rozwoju oraz inne wspierające i powiązane struktury (Whyte, Greenwood, Lazes 1991, s. 31).

U podłoża tego projektu badawczego leżały trzy sprawy: potrzeba restrukturyzacji organizacji, wzrost bezrobocia w kraju, wielka ekspansja kompleksu Mondragón. Projekt zaproponował socjolog William Foote Whyte, a kierowali nim wspólnie Davydd Greenwood, dyrektor Cornell Center for International Studies i Jose Luis Gonzalez, członek personelu w spółdzielni Mondragón. Whyte został zaangażowany na wczesnym etapie projektu. Greenwood i Gonzalez sprawowali funkcje animatorów i konsultantów podczas całego projektu, który finansowany był głównie z grantów wspólnego Hiszpańsko-Amerykańskiego Komitetu Wymiany Oświatowej i Kulturalnej.

Powołano główny zespół złożony z 15 członków spółdzielni, który przystąpił do badań sondażowych wśród około 50 innych członków. Na tym pierwszym etapie projektu badania krążyły wokół wykorzystania literatury i teorii, dotyczących kultury korporacyjnej i zestawiania tych koncepcji i teorii z rzeczywistością spółdzielni Mondragón. Wcześnie zespół doszedł do wniosku, że spółdzielnia powinna wykorzystywać *różnorodność swoich członków*, zamiast wprowadzać w życie uniformizujące zmiany organizacyjne.

To podejście do badań terenowych było szczególnie efektywne w ujawnianiu postaw podzielanych przez członków spółdzielni, takich jak zaangażowanie w organizację i zainteresowanie składnikami wynagrodzenia innymi niż stanowisko pracy i pewna płaca. Współpraca ze zróżnicowanymi pod względem wieku, doświadczenia, formalnego stanowiska i poziomu wykształcenia członkami spółdzielni przyniosła wzrost poziomu zaangażowania i poczucie, że potrzebowali oni rozwiązań łączących ich ideały z rzeczywistością. Ponieważ „konceptualizacja i zastosowanie szły ręka w rękę" (Whyte, Greenwood, Lazes 1991, s. 39) w analizach realizowanych w ramach tego projektu badań uczestniczących, pobudzało to również taki sam proces w zakresie poszukiwania rozwiązań.

Jak widać, pozornie prosty proces obserwowania działań społecznych w ich naturalnym przebiegu ma subtelne, ale ważne odmiany. Jak widzieliśmy w rozdziale 2, wszystkie nasze myśli funkcjonują w ramach paradygmatów i są przez nie kształtowane, niezależnie od tego, czy jesteśmy tego świadomi, czy też nie. Badacze prowadzący jakościowe badania terenowe niezwykle starannie formułują wiele paradygmatów tak, by wzbogacić obserwacje życia społecznego.

Prowadzenie jakościowych badań terenowych

W tym rozdziale do tej pory omawialiśmy tematy odpowiednie dla jakościowych badań terenowych, specjalne okoliczności, jakie towarzyszą wykonywaniu badań tego rodzaju, oraz podaliśmy próbki paradygmatów wymagających badań różnego typu. Przy okazji obejrzeliśmy kilka przykładów, które ilustrowały badania terenowe w działaniu. By uzupełnić obraz, zwracamy się teraz ku charakterystycznym ideom i technikom prowadzenia badań terenowych, rozpoczynając od tego, jak badacze przygotowują się do pracy w terenie.

Przygotowanie do pracy w terenie

Przyjmijmy przez chwilę, że zdecydowałeś się przeprowadzić badania terenowe w organizacji politycznej mieszczącej się w campusie uniwersyteckim. Następnie załóżmy, że nie jesteś członkiem tej grupy, że nie wiesz o niej zbyt wiele i że ujawnisz się wobec uczestników jako badacz. W tej części użyjemy tego i innych przykładów, aby omówić niektóre ze sposobów, jakich mógłbyś użyć, aby przygotować się do podjęcia bezpośredniej obserwacji.

Podobnie jak w przypadku wszystkich metod badawczych, wskazane jest rozpoczęcie badań od poszukiwania stosownej literatury, co pozwoli ci uzupełnić wiedzę na temat przedmiotu badań i dowiedzieć się, co inni stwier-

dzili na ten temat (poszukiwania biblioteczne są omówione szczegółowo w załączniku A).

W następnej fazie badań mógłbyś chcieć omówić problemy studenckiej grupy politycznej z tymi, którzy już ją badali, lub z kimś, kto ją może znać. W szczególności mogłoby okazać się użyteczne omówienie spraw grupy z jednym lub wieloma informatorami (co zostało omówione w rozdziale 7). Być może masz przyjaciela, który jest członkiem tej grupy, lub możesz poznać kogoś, kto nim jest. Ten aspekt twoich przygotowań byłby prawdopodobnie bardziej efektywny, gdyby twoje relacje z informatorem sięgały poza rolę badacza. W kontaktach z członkami grupy jako informatorami powinieneś zachować ostrożność, żeby początkowe rozmowy nie naraziły ani nie ograniczyły dalszych aspektów badań. Weź pod uwagę, że wrażenie, jakie zrobisz na informatorze, rola, w jakiej się ustawisz, mogą zostać przeniesione na twoje dalsze działania. Na przykład początkowe wrażenie, że możesz być tajnym agentem FBI, prawdopodobnie nie ułatwi późniejszej obserwacji grupy.

Powinieneś również ostrożnie traktować informacje, jakie uzyskujesz od informatorów. Chociaż mogą oni mieć bardziej bezpośrednią i osobistą wiedzę na temat przedmiotu badań niż ty, to, co „wiedzą", jest prawdopodobnie mieszaniną faktów i poglądów. Członkowie ugrupowania politycznego w naszym przykładzie raczej nie dostarczyliby całkowicie bezstronnej informacji (podobnie jak członkowie przeciwnych ugrupowań politycznych). Zanim więc podejmiesz pierwszy kontakt z członkami badanej grupy, powinieneś być już dobrze z nią obeznany i rozumieć ogólny kontekst filozoficzny, w jakim działa.

Jest wiele sposobów na nawiązanie pierwszego kontaktu z ludźmi, których zamierzasz badać. Sposób, w jaki to zrobisz, zależy po części od roli, jaką planujesz odegrać. Zwłaszcza jeśli zamierzasz wejść w rolę pełnego uczestnika, musisz znaleźć sposób utożsamienia się z ludźmi, którzy będą badani. Jeśli zamierzasz badać zmywaczy naczyń w restauracji, najprostszą metodą byłoby podjęcie pracy jako zmywacz. W przypadku studenckiego ugrupowania politycznego mógłbyś po prostu do niego dołączyć.

Wiele procesów społecznych, nadających się do badania za pomocą metody terenowej, ma wystarczająco otwarty charakter, by móc nawiązać kontakt z badanymi dość łatwo i bezpośrednio. Jeśli chcesz obserwować demonstrację masową, po prostu bądź tam. Jeśli chcesz obserwować wzorce nieprawidłowego przechodzenia przez jezdnię, postój trochę na zatłoczonych ulicach.

Zawsze gdy chcesz nawiązać z ludźmi kontakt bardziej formalny, ujawniając się jako badacz, musisz nawiązać z nimi dobre stosunki. Możesz kontaktować się z uczestnikiem, z którym dobrze się czujesz, i zyskać wsparcie tej osoby. Przy badaniu grup formalnych możesz zwrócić się do przywódców tych grup lub możesz zostać przedstawiony przez jednego z informatorów.

Prawdopodobnie będziesz miał do wyboru wiele możliwości nawiązania pierwszego kontaktu z grupą, musisz jednak zdawać sobie sprawę, że wybór, jakiego dokonasz, może wpłynąć na twoje późniejsze obserwacje. Załóżmy na przykład, że badasz uniwersytet i zaczynasz od administracji wyższego szczebla. Ten wybór prawdopodobnie będzie miał kilka ważnych konsekwencji. Po pierwsze, wstępne wrażenie dotyczące tego uniwersytetu do pewnego stopnia zostanie ukształtowane przez poglądy pracowników administracji, a będzie ono dość odmienne od spojrzenia studentów czy pracowników naukowych. To pierwsze wrażenie może mieć wpływ na sposób, w jaki obserwujesz i interpretujesz dalsze wydarzenia – szczególnie jeśli nie uświadamiasz sobie tego wpływu.

Po drugie, jeśli administracja pochwala twój projekt badawczy i zachęca studentów i pracowników naukowych do współpracy z tobą, grupy te będą prawdopodobnie postrzegać cię jako kogoś powiązanego z administracją, co może wpływać na to, co ci powiedzą. Na przykład pracownicy naukowi mogą niechętnie opowiadać o planach zorganizowania się poprzez związek zawodowy kierowców.

Nawiązując bezpośredni formalny kontakt z ludźmi, których chcesz badać, będziesz musiał wyjaśnić im jakoś cel twoich badań. Stajesz tutaj ponownie wobec dylematu etycznego.

Mówiąc im, jaki jest pełny cel badań, możesz całkowicie wykluczyć ich współpracę lub znacząco wpłynąć na ich zachowanie. Jednak przekazanie tylko tego, co, jak uważasz, jest akceptowalnym wyjaśnieniem, może oznaczać otwarte oszustwo. To, jaką decyzję podejmiesz w tej i w innych sprawach, będzie prawdopodobnie w dużym stopniu zależało od celu badań, natury tego, co badasz, od obserwacji, jakie zamierzasz zastosować i od podobnych czynników, jednak względy etyczne także muszą zostać wzięte pod uwagę.

Wcześniejsze badania terenowe nie proponowały w tym względzie żadnych stałych reguł metodologicznych ani etycznych. Pojawienie się ciebie jako badacza, niezależnie od celu, jaki przedstawisz, może spowodować ciepłe przyjęcie przez ludzi, którym schlebia, że naukowcy uważają ich za wystarczająco ważnych, by ich badać. Ale może też sprawić, że dotknie cię całkowity ostracyzm lub coś gorszego. Prawdopodobnie nie byłoby dobrym pomysłem wtargnięcie na spotkanie zorganizowanej grupy przestępczej i oświadczenie, że piszesz pracę semestralną o zorganizowanej przestępczości.

Wywiady jakościowe

Po części badania terenowe są sprawą dotarcia tam, gdzie toczą się wydarzenia i, po prostu, patrzenia oraz słuchania. Zgodnie ze słowami przypisywanymi gwieździe baseballu Yogiemu Berra: „możesz wiele zaobserwować tylko patrząc" – jeśli jesteś uważny. Zarazem, jak już zaznaczyłem, badania terenowe mogą wymagać bardziej aktywnych dociekań. Niekiedy właściwe jest zadawanie ludziom pytań i rejestrowanie odpowiedzi. Twoja bezpośrednia obserwacja trwających zamieszek będzie niepełna, jeżeli nie będziesz wiedział, dlaczego ludzie biorą w nich udział. Zapytaj o to kogoś.

Omówiliśmy już przeprowadzanie wywiadów w rozdziale 9 i wiele z tego, co tam powiedziano, stosuje się do przeprowadzania jakościowych wywiadów terenowych. Jednak przeprowadzanie wywiadów powiązanych z obserwacją terenową jest na tyle odmienne, że wymaga osobnego potraktowania. W sondażach kwestionariusze są ściśle ustrukturyzowane, jednak do badań terenowych lepiej nadają się słabiej ustrukturyzowane wywiady. Oto jak Herbert i Riene Rubinowie (1997, s. 202) opisują tę różnicę: „Projekt wywiadów jakościowych jest raczej elastyczny, etapowy i ewolucyjny (*flexible, iterative, continuous*) niż sztywny i niezmienny". Rozwijają to następująco:

> „Projekt wywiadów jakościowych realizuje się etapowo. Oznacza to, że powtarzając podstawowy proces gromadzenia informacji, ich analizowania, selekcjonowania i sprawdzania, przybliżasz się do klarowniejszego i bardziej przekonującego modelu badanego zjawiska (...).
>
> Ewolucyjność wywiadu jakościowego oznacza, że pytania są modyfikowane w toku przedsięwzięcia badawczego"[1].
>
> (Rubin, Rubin 1997, s. 205, 206)

W przeciwieństwie do sondażu, **wywiad jakościowy** jest interakcją między prowadzącym wywiad a respondentem. Prowadzący ma w nim ogólny plan badania, lecz nie jest to konkretny zestaw pytań, które należy zadać z użyciem konkretnych słów i w ustalonym porządku. Jest zarazem bardzo ważne, aby prowadzący wywiad jakościowy, podobnie jak ankieter prowadzący sondaż, był doskonale obeznany z pytaniami, które ma zadać. Dzięki temu wywiad będzie przebiegał gładko i naturalnie.

Wywiad jakościowy jest w zasadzie rozmową, której prowadzący nadaje ogólny kierunek i kładzie szczególny nacisk na pewne tematy, poruszane przez respondenta. Idealnie byłoby, gdyby wypowiedzi respondenta stanowiły większą część rozmowy. Jeśli mówisz więcej niż 5% czasu, to przypuszczalnie mówisz za wiele.

Steinar Kvale (1996, s. 3–5) przedstawia dwie metafory odnoszące się do przeprowadzania wywiadu: prowadzący wywiad jako „górnik" lub jako „podróżnik". Pierwszy model zakłada, że badany ma konkretną informację

[1] Herbert Rubin, Riene Rubin, *Jak zdążać do celu nie wiążąc sobie rąk. Projektowanie wywiadów jakościowych*, tłum. Marcin Ciechomski, w: *Ewaluacja w edukacji*, Leszek Korporowicz (red.), Warszawa 1997 (przyp. tłum.).

i że zadaniem przeprowadzającego wywiad jest jej wydobycie. W drugim modelu prowadzący wywiad:

> „Wędruje przez kraj i nawiązuje rozmowy z napotkanymi ludźmi. Podróżnik bada wiele obszarów kraju, jako nieznane terytoria lub z pomocą mapy, wędrując swobodnie po terenie... Badacz wędruje wraz z mieszkańcami, zadaje im pytania, które pozwalają badanym opowiedzieć swoje historie o świecie, w którym żyją".

Zadawanie pytań i notowanie odpowiedzi jest naturalnym zachowaniem ludzi i wydaje się, że jest wystarczająco proste, by dodać je do zestawu technik używanych przez badacza terenowego. Bądź jednak dość ostrożny. Formułowanie pytań to trudne zadanie. Wszyscy aż nazbyt często zadajemy pytania, które subtelnie wypaczają odpowiedzi, jakie otrzymujemy. Czasami wywieramy na naszego respondenta presję, by dobrze wypadł. Czasami umieszczamy pytanie w szczególnym kontekście, który całkowicie pomija najwłaściwsze odpowiedzi.

Załóżmy na przykład, że chcesz dowiedzieć się, dlaczego grupa studentów bierze udział w zamieszkach i rabunkach na terenie campusu. Możesz mieć skłonność do skupienia się na odczuciach studentów, związanych z ostatnimi regulacjami dziekana, zgodnie z którymi wymaga się zawsze noszenia przy sobie w campusie *Badań społecznych w praktyce*. (Według mnie słusznie). Chociaż mógłbyś zebrać wiele informacji na temat postaw studentów wobec niepopularnej regulacji, oni jednak mogą wszczynać zamieszki z jakiegoś innego powodu. Być może większość zwyczajnie przyłącza się, bo są tym podekscytowani. Prawidłowo przeprowadzone wywiady terenowe umożliwią ci wyjaśnienie tego.

Chociaż możesz zabrać się za robienie wywiadów z wystarczająco jasną ideą tego, o co chcesz pytać, jedną z charakterystycznych mocnych stron badań terenowych jest ich elastyczność. W szczególności odpowiedzi udzielone na twoje pierwsze pytania mogą kształtować pytania, które zadasz później. Tu nie chodzi po prostu o zadawanie wcześniej przygotowanych pytań i zapisywanie odpowiedzi. Przeciwnie, musisz zadać pytanie, wysłuchać uważnie odpowiedzi, zinterpretować jej znaczenie dla twoich ogólnych celów, a następnie sformułować kolejne pytanie, którego celem będzie pogłębienie wcześniejszych odpowiedzi lub skierowanie uwagi rozmówcy ku bardziej interesującemu cię obszarowi. Mówiąc krótko, powinieneś być zdolny do słuchania, myślenia i mówienia prawie jednocześnie.

Omawiane w rozdziale 9 techniki dopytywania dostarczyły użytecznych instrukcji, jak uzyskiwać pogłębione odpowiedzi bez obciążającego wpływu na późniejsze odpowiedzi. Ogólniej rzecz biorąc, badacze przeprowadzający wywiady terenowe powinni mieć umiejętność uważnego słuchania. Staraj się być raczej zainteresowany niż interesujący. Naucz się takich zwrotów, jak: „Jak to jest?", „W jaki sposób?", „Co masz na myśli?", „Co mogłoby być tego przykładem?". Ćwicz patrzenie i słuchanie wyczekująco i pozwól, by to osoba, z którą prowadzisz wywiad, wypełniała ciszę.

Jednocześnie nie możesz pozwolić sobie na bycie całkowicie pasywnym odbiorcą. Przystępujesz do wywiadów, dysponując kilkoma ogólnymi (lub szczegółowymi) pytaniami, na które chcesz uzyskać odpowiedzi, oraz kilkoma tematami, które chcesz poruszyć. Czasami będzie ci potrzebna umiejętność subtelnego ukierunkowywania przebiegu rozmowy.

Możemy się pod tym względem czegoś nauczyć od sztuk walki. Mistrz aikido nigdy nie opiera się ciosom przeciwnika, lecz raczej przyjmuje je, dołącza do nich i delikatnie kieruje je w bardziej odpowiednią stronę. Prowadzenie wywiadów terenowych wymaga analogicznych umiejętności. Zamiast starać się utrzymać w nurcie rozmowy nadanym przez respondenta, ucz się przyjmować to, co on lub ona właśnie powiedzieli, i przetwarzaj te wypowiedzi tak, by powrócić do kierunku rozmowy właściwego dla twoich celów. Większość ludzi uwielbia mówić do kogoś, kto jest rzeczywiście zainteresowany. Zatrzymując potok wypowiedzi, dajesz im do zrozumienia, że nie jesteś zainteresowany; prosząc o rozwinięcie w konkretnym kierunku, przeciwnie, pokazujesz, że jesteś.

Rozważmy hipotetyczny przykład, w którym interesuje cię, jak studenci college'u wybierają specjalizację:

TY: W czym się specjalizujesz?

RESPONDENT: W inżynierii.

TY: Aha. Jak to się stało, że wybrałeś inżynierię?

RESPONDENT: Mam wujka, który został wybrany na najlepszego inżyniera w Arizonie w 1981 roku.

TY: Ojej, to wspaniale.

RESPONDENT: No... Zarządzał budową nowego centrum miejskiego w Tucson. Pisali o tym w większości czasopism technicznych.

TY: Aha. Czy rozmawiałeś z nim o tym, że i ty zostaniesz inżynierem?

RESPONDENT: No... Powiedział, że został inżynierem przez przypadek. Potrzebował pracy po skończeniu szkoły średniej, więc poszedł na budowę jako robotnik niewykwalifikowany. Spędził tam osiem lat, zaczynając od najniższych stanowisk, aż zdecydował się iść na studia i wrócić na wyższe stanowisko.

TY: Czyli interesujesz się głównie inżynierią lądową, tak jak twój wujek, czy może jakąś inną dziedziną inżynierii?

RESPONDENT: Właściwie uczę się bardziej w kierunku elektroniki – szczególnie komputerów. Zacząłem grzebać w komputerach jeszcze w szkole średniej, a mój długofalowy plan jest taki...

Zauważ, jak zaczyna się wywiad, który następnie zbacza w kierunku opowieści o wujku respondenta. Pierwsza próba, by ponownie skoncentrować się na wyborze specjalizacji przez studenta („Czy rozmawiałeś z wujkiem...?"), nie powiodła się. Druga próba („Czyli interesujesz się...?") zakończyła się powodzeniem. Tym razem student dostarczył ci takiej informacji, jaka była ci potrzebna. Dla badacza terenowego ważne jest rozwijanie zdolności do „kontrolowania" rozmowy w taki właśnie sposób.

Herbert i Riene Rubinowie proponują kilka sposobów kontrolowania „moderowanej konwersacji", takich jak następujące:

> „Jeśli możesz ograniczyć liczbę głównych tematów, łatwiej będzie utrzymać strumień konwersacji płynący od jednego tematu ku drugiemu. Przejścia będą gładkie i logiczne. «Dotychczas rozmawialiśmy o matkach, teraz porozmawiajmy o ojcach» brzmi gwałtownie. Gładszym przejściem mogłoby być: «Wspominałeś, że twoja matka nie interesowała się tym, jak radziłeś sobie w szkole – czy twój ojciec był bardziej zaangażowany?». Bardziej raptowne przejście brzmi tak, jakby przeprowadzający wywiad miał plan, który chce wykonać, a nie jakby chciał słuchać tego, co ma do powiedzenia respondent".
>
> (1995, s. 123)

Ponieważ prowadzenie wywiadów terenowych jest tak podobne do normalnej rozmowy, badacze muszą stale przypominać sobie, że nie prowadzą zwykłej rozmowy. W normalnych rozmowach każdy z nas stara się okazać interesującą, godną uwagi osobą. Gdy przyjrzysz się sobie, gawędząc następnym razem z kimś, kogo za dobrze nie znasz, dostrzeżesz prawdopodobnie, że większość uwagi pochłania ci obmyślanie interesujących rzeczy, które powiesz – takiego wkładu do rozmowy, który zrobi dobre wrażenie. Często tak naprawdę nie słuchamy się nawzajem, ponieważ jesteśmy zbyt zajęci myśleniem, co zaraz powiemy. Pragnienie, by wydać się interesującym, jest dla osoby prowadzącej wywiad kontrproduktywne. Prowadzący musi sprawić, by interesująca była ta druga osoba, w tym celu okazując jej zainteresowanie – i raczej słuchając niż mówiąc. (Czyń tak w zwykłych rozmowach, a ludzie będą traktować cię jako wspaniałego rozmówcę).

John i Lyn Loflandowie (1995, s. 56–57) sugerują, by badacze, przeprowadzając wywiad, przyjmowali rolę „niekompetentnych w społecznie akceptowalnym stopniu". Oznacza to, że przedstawiasz siebie jako kogoś, kto nie rozumie sytuacji, w jakiej się znalazł, i wymaga pomocy, aby uchwycić najprostsze i oczywiste jej aspekty: „badacz naturalistyczny niemal z definicji jest kimś, kto nie rozumie. Jest «ignorantem» i trzeba go «nauczyć». Rola osoby przyglądającej się i zadającej pytania to wzorcowa rola *ucznia*" (Lofland, Lofland 1995, s. 56).

Prowadzenie wywiadów powinno być integralną częścią całego procesu badań terenowych. Dalej będę kładł nacisk na konieczność przeglą-

dania każdego wieczora notatek obserwacyjnych – żeby interpretować to, co zostało zaobserwowane, lepiej rozumieć badaną sytuację oraz określać, na co należy zwrócić baczniejszą uwagę w dalszym przebiegu obserwacji. W ten sam sposób należy przeglądać notatki z wywiadów, zapisując przede wszystkim pytania, które zadziałały, i odnajdując te wszystkie pytania, które należało zadać, ale nie zadano ich. Zacznij zadawać te pytania podczas następnych wywiadów. Jeżeli nagrałeś wywiady, przesłuchaj je, traktując to jako przygotowanie do kolejnych.

Steinar Kvale (1996, s. 88) wyszczególnił siedem etapów w pełnym procesie prowadzenia wywiadu:

1. *Określenie tematu*: wyjaśnienie celu wywiadu i zgłębianych pojęć.
2. *Projektowanie*: zaplanowanie procesu, za pomocą którego będą osiągane cele, włącznie z jego wymiarem etycznym.
3. *Prowadzenie wywiadu*: faktyczne przeprowadzenie wywiadu.
4. *Transkrypcja*: stworzenie tekstowego zapisu wywiadu.
5. *Analiza*: określenie znaczenia zgromadzonego materiału w powiązaniu z celem badań.
6. *Weryfikacja*: sprawdzanie rzetelności i trafności materiału.
7. *Raportowanie*: opowiedzenie innym, czego się dowiedzieliśmy.

Podobnie jak w przypadku wszystkich innych aspektów badań terenowych, prowadzenie wywiadów staje się coraz lepsze wraz z nabywanym doświadczeniem. Na szczęście jest to coś, co można ćwiczyć, gdy tylko ma się na to ochotę. Ćwicz na swoich przyjaciołach.

Grupy fokusowe

Nasze dotychczasowe omówienie badań terenowych koncentrowało się na badaniu ludzi, którzy żyją swoimi sprawami, badacze jednak zapraszają czasami ludzi do laboratorium, by tam prowadzić wywiady jakościowe połączone z obserwacją. Technika ta jest szczególnie często używana w związku z marketingiem produktów.

Wyobraź sobie, że zastanawiasz się nad wprowadzeniem nowego produktu. Załóżmy, że wynalazłeś nowy komputer, który nie tylko działa jako edytor tekstu, arkusz kalkulacyjny, narzędzie analizy danych itp., lecz zawiera w sobie również faks, odbiornik radiowy i TV, odtwarzacz CD, podwójny odtwarzacz kasetowy, kuchenkę mikrofalową, urządzenie do czyszczenia sztucznych szczęk i ekspres do kawy. By podkreślić, że jest komputerem i ekspresem, chcesz nazwać go „kompulator". Szacujesz, że nowy komputer będzie sprzedawany za około 28 tysięcy dolarów, i chcesz się dowiedzieć, czy ludzie będą chcieli go kupować. Twoją potrzebę mogłyby zaspokoić grupy fokusowe.

W grupie fokusowej[2] na ogół od 12 do 15 osób zbiera się razem w sali w celu nawiązania moderowanej dyskusji na jakiś temat – w tym przypadku akceptacji dla kompulatora i możliwości jego sprzedawania. Badani zostali dobrani na podstawie związku z przedmiotem poddawanym badaniu. Na przykład biorąc pod uwagę prawdopodobną cenę kompulatora, uczestnicy twojej grupy byliby zapewne rekrutowani wyłącznie z grup o wyższych dochodach. Także inne podobne okoliczności mogłyby wpływać na dobór uczestników.

Jest mało prawdopodobne, by uczestnicy grup fokusowych byli dobierani za pomocą rygorystycznych metod doboru losowego. Oznacza to, że uczestnicy nie są statystyczną reprezentacją żadnej sensownej populacji. Celem badań jest w większym stopniu eksploracja niż opis lub wyjaśnianie w jakimkolwiek wiążącym sensie. Najczęściej jednak w danych badaniach zbiera się więcej niż jedną grupę, mając na uwadze poważne niebezpieczeństwo, że pojedyncza grupa od 7 do 12 osób może być zbyt nietypowa, by na jej podstawie zrozumieć badane zjawisko w sposób pozwalający na uogólnienia.

William Gamson (1992) użył techniki grup fokusowych, by zbadać, jak obywatele Stanów Zjednoczonych wyrażają swoje opinie dotyczące kwestii politycznych. Wybrawszy cztery kwestie – akcję afirmatywną, energię nukle-

[2] W języku polskim grupy fokusowe określane są też jako „zogniskowane wywiady grupowe" lub po prostu „fokusy" (przyp. red. nauk.).

arną, gałęzie przemysłu przeżywające trudności i konflikt izraelsko-arabski – Gamson dokonał analizy treści relacji prasowych, by uchwycić tworzony przez media kontekst, w którym myślimy i rozmawiamy o polityce. Następnie zebrano grupy fokusowe, aby bezpośrednio obserwować proces omawiania przez ludzi spraw politycznych z przyjaciółmi.

Richard Krueger wylicza pięć zalet grup fokusowych:

1. Technika ta jest zorientowaną społecznie metodą badawczą, polegającą na zbieraniu danych z realnego życia w środowisku społecznym.
2. Jest elastyczna.
3. Ma wysoką trafność fasadową.
4. Szybko daje wyniki.
5. Jest tania.

(1988, s. 47)

Oprócz tych zalet, dynamika grupy często ujawnia takie aspekty tematu, których badacz by nie przewidział, i które nie pojawiłyby się w wywiadach z jednostkami. Na przykład w postronnej rozmowie para uczestników może sobie zacząć żartować na temat skutków usunięcia jednej litery z nazwy produktu. Zwrócenie na to uwagi może uchronić producenta przed późniejszym wielkim zakłopotaniem.

Krueger zauważa jednak także pewne wady metody grup fokusowych:

1. W grupie fokusowej kontrola sprawowana przez badacza jest słabsza niż w przypadku wywiadów indywidualnych.
2. Analiza danych jest trudna.
3. Moderator musi mieć specjalne umiejętności.
4. Różnice między grupami mogą sprawiać kłopot.
5. Zebranie grupy jest trudne.
6. Dyskusja musi być prowadzona w sprzyjającym otoczeniu.

(1988, s. 44–45)

Choć badania prowadzone przy wykorzystaniu metody grup fokusowych odbiegają od innych form terenowych badań jakościowych, to jednak są one kolejnym przykładem możliwości przeprowadzania badań społecznych w bezpośrednim kontakcie z tymi, których chcemy zumieć. David Morgan (*Successful...* 1993) wskazuje, że grupy fokusowe są doskonałym narzędziem do tworzenia pytań kwestionariusza do późniejszych badań sondażowych.

Rejestrowanie obserwacji

Największą zaletą metody badań terenowych jest obecność obserwującego, myślącego badacza w miejscu wydarzeń. Nawet magnetofony i kamery nie mogą uchwycić wszystkich istotnych aspektów procesów społecznych. Dlatego zarówno w obserwacji bezpośredniej, jak i w wywiadach istotne jest, żeby z tego, co się dzieje, sporządzać pełne i ścisłe notatki. Jeśli to możliwe, notatki z obserwacji rób na bieżąco. Jeśli jest to niewykonalne, zapisuj później, ale tak szybko, jak jest to możliwe.

W notatkach zawieraj zarówno obserwacje empiryczne, jak i ich interpretacje. Innymi słowy, rejestruj to, co „wiesz", że się wydarzyło, i to, co „myślisz", że się wydarzyło. Upewnij się, że właściwie oznaczasz te różne rodzaje notatek. Mógłbyś zanotować na przykład, że osoba X wyraziła głośno swój sprzeciw wobec propozycji zgłoszonej przez lidera grupy (obserwacja), że *myślisz*, że oznacza to, iż osoba X próbowała przejąć przywództwo w grupie (interpretacja), i że *myślisz*, że usłyszałeś komentarz lidera grupy, dotyczący tego wrażenia, jako odpowiedź na sprzeciw (obserwacja wstępna).

Oczywiście nie możesz liczyć na zaobserwowanie wszystkiego ani na zarejestrowanie wszystkiego, co obserwujesz. Tak jak twoje obserwacje będą jedynie próbą wszystkich możliwych obserwacji, tak twoje notatki będą jedynie próbą twoich obserwacji. Założeniem jest oczywiście rejestrowanie tych najbardziej adekwatnych. Ramka „Transkrypcja wywiadu z notatkami badacza" zawiera podany przez Sandrine Zerbib przykład wywiadu pogłębionego z kobietą reżyserem filmowym.

Niektóre z najważniejszych obserwacji mogą być antycypowane przed rozpoczęciem badań; inne staną się oczywiste wraz z postępem obserwacji. Czasami łatwiej jest robić notatki, zawczasu przygotowując standaryzowane formularze do rejestracji. W przypadku badania nieprawidło-

wego przechodzenia przez jezdnię można przewidzieć cechy pieszych, które najprawdopodobniej będą przydatne do analizy – wiek, płeć, klasa społeczna, grupa etniczna itd. – i przygotować formularz, w którym obserwacje tych zmiennych mogłyby być łatwo rejestrowane. Albo można opracować zawczasu symboliczny system stenograficzny przyspieszający zapisywanie. Do badania uczestnictwa publiczności w zgromadzeniu publicznym można by zbudować ponumerowaną siatkę odpowiadającą różnym częściom sali, w której odbywa się spotkanie; w ten sposób można by rejestrować rozmieszczenie uczestników łatwo, szybko i dokładnie.

Żadne z przedstawionych przygotowań wstępnych nie powinno ograniczać odnotowywania nieprzewidzianych wydarzeń oraz aspektów sytuacji. Wręcz przeciwnie, szybsze opracowywanie obserwacji, które dają się przewidzieć, może dać więcej swobody w obserwowaniu tych nieprzewidzianych.

Zapoznałeś się już z procesem sporządzania notatek, masz już także co najmniej nieformalne doświadczenie z badaniami terenowymi ogólnie. Podobnie jak dobre badania terenowe, dobre robienie notatek wymaga ostrożności i uwagi oraz specyficznych umiejętności. Kilka wskazówek znajduje się w dalszej części. (Możesz dowiedzieć się więcej z *Analyzing Social Settings* Johna i Lyn Loflandów [1995, s. 91–96]).

Po pierwsze, nie ufaj swojej pamięci bardziej, niż musisz; nie jest ona godna zaufania. Aby zilustrować tę uwagę, spróbuj zrobić eksperyment. Przypomnij sobie ostatnio widziane trzy lub cztery filmy, które naprawdę ci się podobały. Następnie podaj nazwiska pięciu aktorów lub aktorek. Kto miał najdłuższe włosy? Kto najchętniej zaczynał rozmowę? Kto najczęściej wysuwał propozycje, za którymi szła reszta? Jeżeli nie miałeś żadnych trudności z odpowiedziami na te pytania, to jak bardzo jesteś pewny odpowiedzi, których udzieliłeś? Czy założyłbyś się o 100 dolarów, żeby zespół bezstronnych sędziów ocenił prawdziwość tego, co sobie przypomniałeś?

Nawet jeśli jesteś dumnym posiadaczem fotograficznej pamięci, dobrze jest robić notatki podczas obserwacji albo zaraz po niej, tak szybko jak to możliwe. Jeśli robisz notatki podczas obserwacji, rób to w sposób dyskretny, ponieważ ludzie mogą różnie się zachowywać, jeśli zobaczą, że zapisujesz wszystko, co robią lub mówią.

Po drugie, zwykle dobrze jest robić notatki etapami. W pierwszym etapie warto jest zrobić notatki szkicowe (zawierające słowa i frazy), aby być na bieżąco z tym, co się dzieje. Następnie, na osobności, należy przepisać notatki, uzupełniając je o szczegóły. Jeśli zrobisz to krótko po tym, kiedy obserwowałeś wydarzenia, szkicowe notatki pozwolą ci przypomnieć sobie większość szczegółów. Im większe opóźnienie, tym mniejsze prawdopodobieństwo, że będziesz w stanie przypomnieć sobie zdarzenia dokładnie i w pełni.

Wiem, że ta metoda wygląda na logiczną, lecz stosowanie jej w praktyce wymaga samodyscypliny. Uważna obserwacja i sporządzanie notatek może być nużące, szczególnie jeśli wiąże się z podnieceniem lub napięciem oraz przeciąga się na długi czas. Jeżeli spędziłeś właśnie osiem godzin, obserwując zmagania ludzi z katastrofalną powodzią i robiąc notatki, to prawdopodobnie pierwsze, czego pragniesz potem, to trochę się przespać, wysuszyć ubranie lub napić się. Pewną inspiracją mogą stać się dla ciebie reporterzy gazet, którzy doświadczają udręk tego samego rodzaju, a następnie spisują swoje opowieści, by zdążyć przed wyznaczonym terminem.

Po trzecie, z pewnością będziesz zdziwiony tym, jak wiele masz do zarejestrowania. Czy na pewno wysiłek zapisywania wszystkich szczegółów, jakie możesz sobie przypomnieć zaraz po sesji obserwacyjnej, jest uzasadniony? Ogólna wskazówka mówi, że tak. Na ogół w badaniach terenowych nie możesz być zupełnie pewny, co jest istotne, a co nie jest, dopóki nie masz okazji do przejrzenia i przeanalizowania wielkiej ilości informacji. Powinieneś zatem rejestrować nawet rzeczy, które z początku nie wydają się istotne. Na koniec to one mogą okazać się ważne. Także sama czynność zapisywania szczegółów czegoś „nieistotnego" może ci pomóc przywołać z pamięci coś istotnego.

Musisz zdawać sobie sprawę z tego, że większość twoich notatek terenowych nie znajdzie odzwierciedlenia w raporcie końco-

Transkrypcja wywiadu z notatkami badacza

Czwartek, 26 sierpnia, godz. 12^{00}–13^{00}

B: Co stanowi wyzwanie dla kobiet reżyserów w codziennym doświadczeniu, w codziennym życiu?

J: Przetrwanie.

B: No dobrze. Czy mogłabyś to trochę rozwinąć? [Powinnam popracować nad moim planem wywiadu, tak by respondentka odpowiadała w bardziej rozbudowany sposób bez potrzeby dopytywania].

J: No, chodzi mi o to wszystko, co jest związane ze staraniem się o dostanie, no wiesz, ze staraniem o dostanie roboty i, no wiesz, z próbą zrobienia świetnej roboty, tak żeby znów zaproszono cię do następnych rzeczy. Zwłaszcza że jest tyle, no wiesz, kłopotów z reżyserowaniem przez kobiety. Jest dwa razy trudniej przebić się na to stanowisko, gdzie robisz nieprzeciętną robotę, bo... nie możesz robić po prostu średniej roboty, musisz [347] robić ją stale na najwyższych obrotach, ale czasami nie masz możliwości, żeby tak robić, bo albo nie masz dobrego producenta, albo jest tak wiele nacisków, że nie masz jasności spojrzenia, scenariusz jest nędzny, a ty musisz zrobić majstersztyk. No wiesz, zbierasz znacznie więcej dodatkowych ciosów niż przeciętny facet na twoim miejscu z podobnymi problemami, bo jesteś kobietą, a oni na to patrzą, i kobiety są bardziej widoczne niż mężczyźni... w wyjątkowych sytuacjach.

[Wydaje się, że Joy mówi o szczególnych cechach przemysłu filmowego. Tam nie ma zbyt wielu możliwości i żeby dalej pracować, musi zdobyć pewną pozycję. Jedynie przez stałe reżyserowanie może utrzymać lub poprawiać swoją pozycję. Uważa, że to jest jeszcze trudniejsze dla kobiet, ale nie wyjaśnia tego].

B: Hm... a jak jest na planie, czy twoje doświadczenia, czy czułaś... czy ludzie dawali ci do zrozumienia, że jesteś kobietą, a ty czułaś się traktowana inaczej? [Staram się skłonić ją do opowiedzenia o konkretnych i bardziej osobistych doświadczeniach bez naprowadzania na odpowiedź].

J: O, tak, tak, myślę, że... wiele kobiet mówiło, że mnie rozumieją, no wiesz, kiedy przychodzisz na plan po raz pierwszy, oni wszyscy przywykli do pracy jak dobrze naoliwione maszyny i mówią „o, kobieta, coś innego" i czasami mogą być okropni, mogą ci się sprzeciwiać jako reżyserowi i mogą, mogą cię sabotować, zaczynają długo świecić światło lub zmieniają plan lub robią coś... i wtedy ty tracisz czas, a to idzie do raportu, a raport idzie do głównego [368] biura, i wiesz, i tak dalej, i tak dalej, itd., itp. A ludzie na górze nie wiedzą, jaka jest sytuacja, i nie zamierzają wyrzucać operatora, który jest w ich branży na wieki wieków... oni nie chcą wiedzieć, że ten facet to naprawdę gnojek i zamienia twoje życie w horror. Nie chcą wiedzieć, więc wyłączają się, ona to kobieta, nie zatrudniajmy więcej żadnych kobiet, skoro on ma problemy z kobietami. No wiesz, więc to o to chodzi.

[Muszę przejrzeć literaturę na temat instytucjonalnej dyskryminacji. Wydaje się, że wyzwania, przed jakimi staje Joy, nie są sprawami konkretnej jednostki. Ona ma dylemat, bo czy się poskarży, czy nie, nie będzie traktowana na równi z mężczyznami. Wydaje się, że czas jest jedyną ilościową miarą tego, jak dobrze wykonuje ona swoją pracę i, tak jak w innych zawodach, fakt, iż jest kobietą, postrzegany jest jako ułomność. Przejrzeć literaturę o kobietach na wysokich stanowiskach w zarządzaniu. Muszę stale zadawać pytania na temat dynamiki relacji między moimi respondentkami a członkami ekipy z planu. Operator ma na planie najwyższy status po reżyserze. Zgłębić pozostałe wywiady pod względem przyczyn konfliktu między nimi].

[Metody (notatka na użytek własny do następnych wywiadów): starać się unikać wywiadów telefonicznych, chyba że na konkretną prośbę ze strony respondentki. Trudno jest ocenić, jak respondentka czuje się przy pytaniu. Mowa ciała jest mi potrzebna, ponieważ staję się niespokojna o przebieg wywiadu].

Uwaga: Liczba w nawiasach oznacza słowo, które w wywiadzie było niesłyszalne. Jest to liczba, jaką pokazał magnetofon, przy czym każdy wywiad zaczynał się przy stanie licznika 0. Pomaga ona badaczowi szybko zlokalizować urywek podczas przesłuchiwania wywiadu.

wym z projektu. Mówiąc bardziej szorstko, większość zrobionych przez ciebie notatek będzie „zmarnowana". Dla otuchy pomyśl, że nawet najbogatsza ruda złota zawiera jedynie około 30 gramów złota na tonę, co oznacza, że 99,997% rudy to odpad. A jednak owe 30 gramów złota można rozwalcować tak cienko, że może pokryć obszar 1,7 metra kwadratowego równoważny 685 stronom książki. Zbierz zatem tonę notatek, a planuj wybrać i użyć jedynie złoto.

Podobnie jak w innych aspektach badań terenowych (a pod tym względem – w ogóle wszystkich badań), biegłość przychodzi wraz z praktyką. Miłą rzeczą w badaniach terenowych jest to, że możesz rozpocząć ćwiczenie już teraz i kontynuować je w prawie każdej sytuacji. Nie musisz uczestniczyć w zorganizowanym projekcie badawczym, by ćwiczyć obserwację i sposoby jej rejestrowania. Na przykład możesz zacząć, zgłaszając się na ochotnika do protokołowania zebrań komisji. Lub po prostu wybrać słoneczny dzień i w zacienionym miejscu w campusie starać się obserwować i zapisywać pewne charakterystyczne cechy przechodzących ludzi. Możesz robić to samo w centrum handlowym lub na rogu ruchliwej ulicy. Pamiętaj, że obserwacja i rejestracja to umiejętności zawodowe i tak jak inne cenne umiejętności rozwijają się dzięki ćwiczeniu.

Mocne i słabe strony jakościowych badań terenowych

Podobnie jak wszystkie metody badawcze, także jakościowe badania terenowe mają charakterystyczne mocne i słabe strony. Jak już wcześniej wskazywałem, badania terenowe są szczególnie efektywne, gdy stosuje się je do badania subtelnych niuansów w postawach i zachowaniach oraz do badania procesów społecznych w czasie. Dlatego główna mocna strona tej metody leży w głębi rozumienia, jakie umożliwia. Podczas gdy innym metodom badawczym można zarzucić, że są „powierzchowne", ten zarzut rzadko bywa stawiany badaniom terenowym.

Elastyczność jest kolejną zaletą badań terenowych. Jak mówiliśmy wcześniej, projekt badań terenowych można modyfikować w dowolnym momencie. Co więcej, jesteś stale przygotowany, by zaangażować się w badania terenowe, gdy nadarzy się taka okazja, natomiast nie mógłbyś tak łatwo przystąpić do sondażu lub eksperymentu.

Badania terenowe mogą być również dość tanie. Inne metody badawcze w naukach społecznych mogą wymagać kosztownego sprzętu lub personelu badawczego, lecz badania terenowe na ogół mogą być podejmowane przez jednego badacza, wyposażonego w notes i ołówek. Nie oznacza to, że badania terenowe nigdy nie są kosztowne. Na przykład charakter projektu badawczego może wymagać zaangażowania wielu wyszkolonych obserwatorów. Może okazać się niezbędny kosztowny sprzęt do nagrywania. Lub możesz zechcieć podjąć obserwację uczestniczącą interakcji w drogich klubach nocnych Paryża.

Badania terenowe mają jednak również kilka słabych stron. Przede wszystkim, jako metoda raczej jakościowa niż ilościowa, nie są odpowiednim środkiem dla dochodzenia do statystycznych opisów wielkich populacji. Na przykład obserwacja codziennych dyskusji politycznych w pralniach samoobsługowych nie dostarczyłaby wiarygodnych szacunków przyszłych zachowań wyborczych całego elektoratu. Niemniej jednak takie badania mogłyby dać ważny wgląd w kształtowanie się postaw politycznych.

By lepiej ocenić badania terenowe, skupmy się na kwestii ich trafności i rzetelności. Przypomnijmy sobie, że trafność i rzetelność są cechami pomiaru. Trafność dotyczy tego, czy pomiar faktycznie mierzy to, co sądzimy, czy też może mierzy coś innego. Rzetelność natomiast jest sprawą niezawodności: czy uzyskasz te same rezultaty, przeprowadzając ten sam pomiar wiele razy? Zobaczmy, jak badania terenowe wyglądają pod tym względem.

Trafność

Wydaje się, że badania terenowe dostarczają mierników o większej trafności niż pomiary sondażowe i eksperymentalne, które są często kry-

tykowane jako powierzchowne i w rzeczywistości nietrafne. Przejrzyjmy kilka przykładów badań terenowych, by zobaczyć, dlaczego tak się dzieje.

„Bycie na miejscu" to skuteczna technika uzyskiwania wglądu w naturę ludzkich spraw w całej ich bogatej złożoności. Posłuchaj na przykład sprawozdania pielęgniarki na temat barier napotykanych przez pacjentów walczących z rakiem:

> „Wspólny lęk, który może hamować proces walki w przypadku osób z rakiem, może obejmować, co następuje:
> - Lęk przed śmiercią – dotyczy pacjenta i skutków jego śmierci dla znaczących innych.
> - Lęk przed nieudolnością – ponieważ rak może być chorobą chroniczną z ostrymi epizodami, które mogą dawać cykliczne okresy napięcia, różny stopień zdolności danej osoby do radzenia sobie z chorobą oraz ciągłego dostosowywania się mogą prowadzić do zależności od innych przy codziennych czynnościach, a w konsekwencji mogą stać się ciężarem.
> - Lęk przed alienacją – od znaczących innych i przedstawicieli służby zdrowia, wywołujący poczucie bezradności i beznadziei.
> - Lęk przed zarażeniem – obawa przed tym, że rak jest przenoszony i/lub dziedziczny.
> - Lęk przed utratą własnej godności – przed utratą kontroli nad wszystkimi funkcjami cielesnymi i byciem całkowicie zdanym na innych".
>
> (Garant 1980, s. 2167)

Obserwacje i konceptualizacje takie jak powyższe mają wartość same w sobie. Na dodatek mogą one dostarczać podstaw do dalszych badań – zarówno jakościowych, jak i ilościowych.

Spójrz teraz, co Joseph Howell ma do powiedzenia na temat „twardości" jako fundamentalnego składnika życia na Clay Street, waszyngtońskiej dzielnicy białej klasy robotniczej:

> „Większość ludzi z Clay Street postrzegała siebie jako bojowników zarówno w sensie symbolicznym, jak i dosłownym. Uważali się za silnych, niezależnych ludzi, którzy nie pozwoliliby, aby nimi pomiatano. Dla Bobbiego bycie bojownikiem oznaczało walkę z wydziałem spraw socjalnych i wyklinanie pracowników socjalnych i lekarzy przy każdej okazji. Oznaczało doprawianie piwa Barry'ego tabletkami nasennymi i walenie go miotłą po głowie. Dla Barry'ego oznaczało to besztanie swojego szefa i odmowę zawieszenia drzwi, co było czynem prowadzącym do wyrzucenia z pracy. Oznaczało przechodzenie rytuału pojedynku z Alem. Oznaczało pomiatanie Bubbą i czasami bycie szorstkim dla Bobbiego.
>
> June i Sam mieli mniej powodów do kłótni, choć przyciśnięci do muru obaj wspominali, że także mogliby walczyć. Bycie bojownikiem doprowadziło Toma do szybkiego konfliktu z braćmi Peg, doprowadziło do konfliktu Lesa z Lonniem, do konfliktu Arlene'a z Phyllisem na kręgielni itp."
>
> (1973, s. 292)

Nawet nie słysząc epizodów, referowanych w powyższym fragmencie przez Howella, ma się wyraźne wrażenie, że na Clay Street jest ciężko żyć. Ową „twardość" widać znacznie wyraźniej na podstawie tych obserwacji terenowych niż z zestawu statystyk dotyczących średniej liczby walk na pięści w tym okresie.

Te przykłady wskazują na większą trafność badań terenowych w porównaniu z sondażami i eksperymentami. Ten rodzaj całościowych pomiarów dostępnych badaczowi terenowemu dotyka głębi znaczenia takich pojęć, jak wspólne lęki pacjentów dotkniętych rakiem i „twardość" (albo takich jak liberalizm i konserwatyzm), które są ogólnie niedostępne sondażom i eksperymentom. Zamiast określania pojęć badacze terenowi zamieszczają zwykle szczegółowe przykłady.

Rzetelność

Badania terenowe mogą mieć jednak problem z rzetelnością. Załóżmy, że miałbyś określić poglądy polityczne swojego najlepszego przyjaciela zgodnie z tym wszystkim, co o nim wiesz. Twoja ocena poglądów przyjaciela miałaby znaczną trafność i jest z pewnością nieprawdopodobne, aby była powierzchowna. Nie mielibyśmy jednak pewności, że inny obserwator

scharakteryzowałby poglądy twojego przyjaciela w taki sam sposób jak ty, nawet gdyby dysponował taką samą ilością obserwacji.

Pogłębione pomiary w badaniach terenowych są często również bardzo osobiste. To, jak ocenię poglądy polityczne twojego przyjaciela, bardzo zależy od moich własnych poglądów, podobnie jak twoja ocena zależy od twoich poglądów politycznych. Możliwe więc, że ty opisałbyś swojego przyjaciela jako centrystę, chociaż ja mógłbym odnieść wrażenie, że obserwuję ziejącego ogniem radykała.

Jak już wcześniej sugerowałem, badacze, którzy używają technik jakościowych, są świadomi tej sprawy i podejmują starania, by uniknąć takich efektów. Poszczególni badacze często oddzielają swoje własne poglądy, a dzięki społecznej naturze nauki ich koledzy mogą im w tym pomóc. Mimo wszystko roztropnie jest być ostrożnym co do czysto opisowych pomiarów we własnych lub cudzych badaniach terenowych. Jeśli badacz donosi, że członkowie klubu są dość konserwatywni, to taka ocena jest w nieunikniony sposób związana z własnymi poglądami politycznymi badacza. Bardziej wiarygodne są oceny *porównawcze:* na przykład określenie, kto jest bardziej konserwatywny od kogo. Nawet gdybyśmy mieli różne poglądy polityczne, prawdopodobnie dosyć łatwo zgodzilibyśmy się co do uszeregowania członków grupy pod kątem ich względnego konserwatyzmu.

Jak widzieliśmy, badania terenowe mogą być potężnym narzędziem w rękach badaczy społecznych, narzędziem równoważącym mocne i słabe strony eksperymentów i sondaży. Pozostałe rozdziały części 3 prezentują dodatkowe sposoby obserwacji, dostępne badaczom społecznym.

Etyka badawcza w jakościowych badaniach terenowych

Jak wielokrotnie wskazywałem, wszystkie formy badań społecznych wiążą się z kwestiami etycznymi. Przez wejście badaczy w bezpośredni i często osobisty kontakt z badanymi, badania terenowe wywołują spory etyczne w szczególnie dramatyczny sposób. Oto kilka z tych kwestii wzmiankowanych przez Johna i Lyn Loflandów (1995, s. 63):

- Czy etyczne jest rozmawianie z ludźmi, którzy nie wiedzą, że będziesz rejestrować ich słowa?
- Czy etyczne jest uzyskiwanie informacji dla twoich celów od ludzi, których nienawidzisz?
- Czy etyczne jest dostrzeganie pilnej potrzeby pomocy i niereagowanie na to?
- Czy etyczne jest bycie w jakimś środowisku lub sytuacji i nieangażowanie się w pełni?
- Czy etyczne jest przyjmowanie wobec innych ludzi wyrachowanej postawy, aby działać zgodnie z planem w relacjach z innymi ludźmi?
- Czy etyczne jest stawanie po stronie jednych lub niestawanie po żadnej ze stron, kiedy istnieje podział na frakcje?
- Czy etyczne jest „płacenie" ludziom w transakcjach wymiany za dostęp do ich życia i umysłów?
- Czy etyczne jest „używanie" ludzi jako sojuszników lub informatorów w celu uzyskania dostępu do innych ludzi lub dla ulotnego zrozumienia?

Planowanie i przeprowadzanie badań terenowych w sposób odpowiedzialny wymaga zajmowania się tymi i innymi kwestiami etycznymi. Omówienie etyki badawczej będziemy kontynuować w rozdziale 18.

Główne zagadnienia

- Badania terenowe wymagają bezpośredniej obserwacji zjawisk społecznych w ich naturalnym środowisku. Badania terenowe są na ogół raczej jakościowe niż ilościowe.
- W badaniach terenowych obserwacja, przetwarzanie danych i analiza są przeplatającymi się cyklicznymi procesami.
- Badania terenowe są szczególnie dostosowane do tematów i procesów, które nie są łatwo przekładalne na analizy ilościowe, które najlepiej jest badać w ich naturalnym środowisku lub które zmieniają się w czasie. Wśród tych zagadnień znajdują się

działania zwyczajowe, epizody, spotkania, role, związki, grupy, organizacje, miejsca zamieszkania, światy społeczne i style życia lub subkultury.

- Do specyficznych zagadnień, związanych z badaniami terenowymi, należą rozmaitość ról, w jakie może wcielać się obserwator, oraz relacje badacza z badanymi. Jako badacz terenowy musisz zdecydować, czy obserwować jako ktoś z zewnątrz, czy jako uczestnik, czy pozwolić, aby ujawnić się jako badacz, czy też nie, a także jak ustalić swoje relacje z podmiotami.
- Badania terenowe mogą być prowadzone w każdym z wielu paradygmatów, takich jak naturalizm, etnometodologia, teoria ugruntowana, studium przypadku i metoda rozszerzonych przypadków, etnografia instytucjonalna i badania uczestniczące.
- Przygotowanie do wyjścia w teren wymaga przeprowadzenia badań podstawowych o charakterze przygotowawczym, w celu określenia, jak nawiązać kontakt z badanymi, i zdecydowania, jaki charakter będą miały relacje z badanymi.
- Badacze terenowi często prowadzą wywiady pogłębione, które są znacznie słabiej ustrukturyzowane niż wywiady przeprowadzane w badaniach sondażowych. Prowadzenie wywiadów jakościowych jest raczej moderowaną rozmową niż szukaniem określonych informacji. Skuteczne prowadzenie wywiadów wymaga umiejętności aktywnego słuchania i zdolności do dyskretnego kierowania rozmową.
- W celu skonstruowania grupy fokusowej badacze gromadzą badanych i obserwują ich interakcje podczas zgłębiania określonego zagadnienia.
- Jeśli to tylko możliwe, obserwacje terenowe powinny być rejestrowane na bieżąco. W przeciwnym przypadku powinny być rejestrowane tak szybko, jak to tylko możliwe.
- Do zalet badań terenowych należy możliwość pogłębionego rozumienia, elastyczność i (zwykle) niskie koszty.
- W porównaniu z sondażami i eksperymentami badania terenowe cechuje większa trafność, ale mniejsza rzetelność. Generalnie badania terenowe nie są właściwym sposobem uzyskiwania statystycznych opisów wielkich populacji.
- Odpowiedzialne prowadzenie badań terenowych wymaga zmierzenia się z kilkoma kwestiami etycznymi, które wynikają z bezpośredniego kontaktu badacza z badanymi.

Kluczowe pojęcia

naturalizm
etnografia
etnometodologia
teoria ugruntowana
studium przypadku
metoda rozszerzonych przypadków
etnografia instytucjonalna
badania uczestniczące
wywiad jakościowy
grupa fokusowa

Pytania i ćwiczenia

1. Pomyśl o jakiejś grupie lub działaniu, w którym uczestniczysz, lub z którym jesteś dobrze zaznajomiony. W dwóch, trzech akapitach opisz, jak człowiek z zewnątrz mógłby efektywnie zająć się badaniem tej grupy lub działania. Co powinien przeczytać, jakie kontakty nawiązać itp.?
2. Wybierz jakieś dwa paradygmaty, omawiane w tym rozdziale. Następnie opisz, jak twoje hipotetyczne badania z punktu 1 mogłyby być prowadzone zgodnie z każdym z nich. Porównaj, zwracając uwagę na różnice, w jaki sposób te paradygmaty mogłyby zostać wykorzystane w twoich badaniach.
3. W celu zbadania mocnych i słabych stron eksperymentu, sondażu i badań terenowych wybierz ogólny obszar badawczy (na przykład uprzedzenia, poglądy polityczne, edukację) i opisz krótko badania w tej dziedzinie, jakie mogłyby zostać prowadzone przy użyciu każdej z tych trzech metod. W każdym przypadku wyjaśnij, dlaczego wybrana metoda jest najbardziej odpowiednia dla opisywanych badań.
4. Powróć do przykładu wymyślonego w związku z punktem 1 i wymień pięć kwestii etycz-

nych, z którymi, jak sądzisz, zetknąłbyś się przy takich badaniach.

5. Używając Internetu znajdź raport z badań, przeprowadzonych metodą teorii ugruntowanej. Podsumuj schemat badań i główne wyniki.

Dodatkowe lektury

Computer-Aided Qualitative Data Analysis: Theory, Methods, and Practice, Udo Kelle (red.), Sage, Thousand Oaks, Cal. 1995. Międzynarodowa grupa uczonych opisuje swoje doświadczenia z wieloma programami komputerowymi, stosowanymi do analizy danych jakościowych.

Contemporary Field Research, Robert M. Emerson (red.), Little Brown, Boston 1988. Różnorodny i interesujący zbiór artykułów o tym, jak badania terenowe przyczyniają się do wyjaśniania, o roli teorii w badaniach tego rodzaju, o wynikających z nich spraw osobistych i dotyczących relacji między ludźmi oraz o kwestiach etycznych i politycznych.

Experiencing Fieldwork: An Inside View of Qualitative Research, William B. Shaffir, Robert A. Stebbins (red.), Sage, Newbury Park, Cal. 1991. Kilku praktyków badań terenowych omawia naturę tej sztuki i przywołuje doświadczenia z terenu. Można tu zarówno „poczuć" metodę, jak i nauczyć się pewnych technik.

Gubrium Jaber F., Holstein James A., *The New Language of Qualitative Method*, Oxford University Press, New York 1997. Książka ta dostarcza niezbędnych podstaw dla zrozumienia niektórych z głównych podejść lub tradycji jakościowych badań terenowych.

Handbook of Qualitative Research, Norman K. Denzin, Yvonna S. Lincoln (red.), Sage, Thousand Oaks, Cal. 1994. Ten podręcznik jest obszernym zbiorem artykułów, obejmujących kwestie odnoszące się do szerokiej dziedziny badań jakościowych. Książka ta ukazała sie również w trzech tomach: T. 1 *The Landscape of Qualitative Research: Theories and Issues*; T. 2 *Strategies of Qualitative Inquiry*; T. 3 *Interpreting Qualitative Materials*.

Johnson Jeffrey C., *Selecting Ethnographic Informants*, Sage, Newbury Park, Cal. 1990. Autor omawia rozmaite strategie, które odnoszą się do zadania, jakim jest dobór próby w badaniach terenowych.

Kvale Steinar, *InterViews: An Introduction to Qualitative Research Interviewing*, Sage, Thousand Oaks, Cal. 1996. Dokładna prezentacja wywiadów pogłębionych. Oprócz przedstawienia technik Kvale umieszcza prowadzenie wywiadów w kontekście postmodernizmu i innych systemów filozoficznych.

Lofland John, Lofland Lyn, *Analyzing Social Settings*, wyd. 3, Wadsworth, Belmont. Niedościgła prezentacja metod badań terenowych od A do Z. W tej szczególnie czytelnej książce udało się zarysować powiązania między logiką odkrycia naukowego a najistotniejszymi praktycznymi stronami obserwacji, komunikacji, rejestracji, porządkowania, raportowania i całej reszty spraw, związanych z badaniami terenowymi. W dodatku książka zawiera bogactwo odwołań do przykładów badań terenowych.

Our Sociological Eye: Personal Essays on Society and Culture, Arthur Shostak (red.), Alfred, Port Washington, New York 1997. Bogactwo wglądu w badania społeczne, zachwycający zbiór relacji badawczych z pierwszej ręki, przedstawiający konkretny wewnętrzny ogląd procesu myślowego w badaniach socjologicznych, szczególnie w badaniach terenowych.

Silverman David, *Doing Qualitative Research: A Practical Handbook*, Sage, Thousand Oaks, Cal. 1999. Ta książka skupia się na procesie zbierania i interpretacji danych jakościowych.

Strauss Anselm, Corbin Juliet, *Basic of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage, Newbury Park, Cal. 1990. Jeśli zdecydowałeś się zastosować podejście teorii ugruntowanej, jest to bardzo ważna książka, którą należy przeczytać przed zbieraniem danych, ale także podczas ich analizy.

Successful Focus Groups: Advancing the State of the Art, David L. Morgan (red.), Sage, Newbury Park, Cal. 1993. Ten zestaw artykułów o zastosowaniach techniki grup fokusowych

zwraca uwagę na wiele, zazwyczaj pomijanych, aspektów.

Uwe Flick, *An Introduction to Qualitative Research*, Sage, Thousand Oaks, Cal. 1998. To książka dająca dobre wprowadzenie do obszernej dziedziny badań jakościowych.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 11

Badania niereaktywne

Przegląd holograficzny

Rozdział ten jest przeglądem trzech metod badań niereaktywnych: analizy treści, analizy istniejących danych statystycznych i analiz historyczno-porównawczych. Każda z tych metod pozwala badaczowi na poznawanie życia społecznego z dystansu, bez oddziaływania na nie.

Wprowadzenie

Analiza treści
Tematy odpowiednie dla analizy treści
Dobór próby w analizie treści
Kodowanie w analizie treści
Przykład analizy treści
Mocne i słabe strony analizy treści

Analiza istniejących danych statystycznych
Badania nad samobójstwami Durkheima
Jednostki analizy
Problem trafności
Problem rzetelności
Źródła istniejących danych statystycznych

Analizy historyczno-porównawcze
Przykłady analiz historyczno-porównawczych
Źródła danych historyczno-porównawczych
Techniki analityczne

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Z wyjątkiem pełnego obserwatora w badaniach terenowych, każdy z dotychczas omówionych rodzajów obserwacji wymaga od badacza pewnego stopnia ingerencji w to, co jest badane. Najbardziej oczywiste jest to w przypadku eksperymentów, nieznacznie ustępują im badania sondażowe. Jak widzieliśmy, nawet badania terenowe mogą zmieniać sytuacje poddawane procesowi badawczemu.

Co najmniej jeden z wcześniej przedstawionych w tej książce przykładów jest całkowicie pozbawiony tego niebezpieczeństwa. Przeprowadzona przez Durkheima analiza samobójstw w żaden sposób nie wpływała na samobójstwa (patrz rozdział 5). Jego badania są przykładem **badań niereaktywnych**, czyli inaczej metod badania zachowań społecznych, nie wpływających na te zachowania. Jak zobaczymy, mierniki niereaktywne mogą być zarówno jakościowe, jak i ilościowe.

W tym rozdziale przedstawione zostaną trzy metody badań niereaktywnych: analiza treści, analiza istniejących danych statystycznych i analizy historyczno-porównawcze. W przypadku analizy treści badacze analizują jakąś klasę wytworów społecznych, najczęściej dokumentów pisanych, takich jak artykuły redakcyjne w gazetach. Badania Durkheima są przykładem wykorzystania istniejących statystyk. Jak się przekonasz, otaczają cię wielkie masy danych, czekających na wykorzystanie dla zrozumienia życia społecznego. Wreszcie, analizy historyczno-porównawcze, forma badań o budzącej respekt historii w naukach społecznych, są obecnie znowu popularne. Podobnie jak badania terenowe, analizy historyczno-porównawcze to metoda jakościowa, w której jednym z głównych źródeł obserwacji i analiz są zapisy historyczne. Nazwa metody zawiera słowo „porównawcze", ponieważ badacze społeczni – w przeciwieństwie do historyków, którzy mogą po prostu opisać konkretny zestaw wydarzeń – starają się raczej odkryć wspólne prawidłowości, powtarzające się w różnym czasie i różnych miejscach.

W ramach przygotowania do zapoznania się z tymi trzema metodami badawczymi, chciałbym skierować twoją uwagę ku wspaniałej książce, która powinna uwrażliwić cię na ogólny potencjał miar niereaktywnych. Z tej książki, prócz innych rzeczy, zaczerpnąłem termin *miary niereaktywne* (*unobtrusive measures*).

W 1966 roku Eugene Webb z trzema współpracownikami opublikował niewielką, oryginalną książeczkę na temat badań społecznych (wznowioną w 1981 roku). Stała się ona pozycją klasyczną. Skupiała się na idei badań niereaktywnych. Webb i jego współpracownicy potraktowali swobodnie zadanie poznawania ludzkich zachowań poprzez obserwowanie, co ludzie nieumyślnie pozostawiają po sobie. Czy chciałbyś wiedzieć, jakie wystawy są najpopularniejsze w muzeum? Mógłbyś przeprowadzić sondaż, lecz ludzie mogliby powiedzieć ci to, co według nich chciałbyś usłyszeć lub to, co mogłoby ukazać ich jako myślących i poważnych. Mógłbyś stać przy różnych wystawach i liczyć przechodzących widzów, lecz ludzie mogliby przychodzić, aby zobaczyć, co robisz. Webb ze współpracownikami sugerują, by sprawdzać zużycie i zniszczenia podłogi przed różnymi ekspozycjami. Te, przed którymi klepka jest najbardziej zniszczona, są prawdopodobnie najbardziej popularne. Chcesz wiedzieć, jaka ekspozycja jest popularna wśród małych dzieci? Szukaj śladów śliny na szklanych gablotach. By uzyskać pojęcie o najbardziej popularnych stacjach radiowych, mógłbyś umówić się z mechanikiem samochodowym, by kontrolował ustawienia częstotliwości w samochodach przyprowadzanych do naprawy.

Możliwości są nieograniczone. Podobnie jak detektyw prowadzący śledztwo w sprawie przestępstwa, badacz społeczny poszukuje wskazówek. Jeśli zatrzymasz się, by się uważnie rozejrzeć, przekonasz się, że wskazówki dotyczące zachowań społecznych są wszędzie wokół ciebie. W pewnym sensie wszystko, co widzisz, stanowi odpowiedź na jakieś ważne pytanie nauk społecznych – ty musisz tylko wymyślić to pytanie.

Chociaż problemy trafności i rzetelności pojawiają się w miarach niereaktywnych, nieco pomysłowości wystarczy, by sobie z nimi poradzić lub ustawić je we właściwych proporcjach. Zachęcam do zajrzenia do książki Webba. To

przyjemna lektura, która może być źródłem zachęty i inspiracji do prowadzenia badań społecznych z wykorzystaniem już istniejących danych. Skierujmy teraz naszą uwagę ku trzem niereaktywnym metodom, często wykorzystywanym przez badaczy społecznych, rozpoczynając od analizy treści.

Analiza treści

Jak wspominałem we wprowadzeniu do tego rozdziału, **analiza treści** jest badaniem zarejestrowanych ludzkich przekazów. Do form przekazu odpowiednich do takiego badania należą książki, czasopisma, strony www, wiersze, gazety, piosenki, obrazy, przemówienia, listy, listy elektroniczne, wiadomości na tablicach informacyjnych w Internecie, prawa i konstytucje, podobnie jak części, z których się one składają, lub złożone z nich zbiory. Shulamit Reinharz (1992, s. 146–147) wskazuje, że feministyczne badaczki używają analizy treści do badania „książek dla dzieci, bajek, tablic reklamowych, książek feministycznych, zarówno beletrystyki, jak i literatury faktu, dziecięcych prac plastycznych, mody, kartek pocztowych, podręczników dziewczęcych drużyn skautowych, działań artystycznych, retoryki gazetowej, rejestrów szpitalnych, publikacji badawczych, podręczników ze wstępu do socjologii oraz cytatów, żeby wymienić tylko część".

Tematy odpowiednie dla analizy treści

Analiza treści jest szczególnie dobrze dostosowana do badania przekazów i do udzielania odpowiedzi na klasyczne pytanie w badaniach przekazów: „kto mówi, co, do kogo, dlaczego, jak i z jakim skutkiem?". Czy popularne powieści francuskie częściej mówią o miłości niż powieści w Stanach Zjednoczonych? Czy brytyjska muzyka popularna lat 60. była bardziej cyniczna politycznie niż popularna muzyka niemiecka w tym samym okresie? Czy kandydaci polityczni, którzy podnosili głównie kwestie źródeł utrzymania, byli wybierani częściej niż kandydaci podnoszący kwestie zasad wyższego rzędu? Każde z tych pytań dotyczy pewnego tematu społecznych badań naukowych: pierwsze mogło dotyczyć charakteru narodowego, drugie poglądów politycznych, a trzecie procesu politycznego. Chociaż takie tematy można badać przez obserwację pojedynczych ludzi, analiza treści proponuje inne podejście.

Przykładem wczesnej analizy treści jest praca Idy B. Wells. W 1891 roku Wells, której rodzice byli niewolnikami, chciała sprawdzić szeroko upowszechnione przekonanie, że czarni mężczyźni na Południu byli linczowani głównie za gwałty na białych kobietach. Zanalizowała artykuły gazetowe o 728 linczach z ostatnich dziesięciu lat. W zaledwie jednej trzeciej przypadków ofiary linczu były choćby oskarżane o gwałt, jeszcze rzadziej udowodniono im winę. Przede wszystkim ludzi tych oskarżano o to, że są bezczelni, nie znają „swojego miejsca" (Reinharz 1992, s. 146).

W naszych czasach bestseller książkowy *Megatrendy 2000* (Naisbitt, Aburdene 1997) stosuje analizę treści dla określenia głównych trendów w życiu współczesnych Stanów Zjednoczonych. Autorzy regularnie monitorowali tysiące gazet lokalnych w każdym miesiącu, chcąc odkryć lokalne i regionalne trendy, i publikując je w serii kwartalnych sprawozdań. Książka rozpatruje niektóre z zaobserwowanych trendów w skali ogólnokrajowej.

Są tematy, dla których analiza treści jest najbardziej odpowiednią metodą badawczą. Załóżmy, że interesuje cię przemoc w telewizji. Może przypuszczasz, że producenci produktów dla mężczyzn są bardziej od innych skłonni sponsorować programy telewizyjne zawierające przemoc. Analiza treści byłaby najlepszym sposobem, by to sprawdzić.

Oto krótko, co należy zrobić. Po pierwsze trzeba zbudować definicje operacyjne dwóch kluczowych dla twoich badań zmiennych: *produkty dla mężczyzn* i *przemoc*. W dalszej części tego rozdziału, poświęconej kodowaniu, omówimy kilka sposobów, jak można to zrobić. Ostatecznie potrzebny byłby plan, który pozwoliłby oglądać telewizję, klasyfikować sponsorów i oceniać stopień przemocy w poszczególnych programach.

Następnie należy zadecydować, co oglądać. Prawdopodobnie zadecydowałbyś: (1) jakie stacje oglądać, (2) przez jaki okres i (3) w jakich godzinach. Następnie zaopatrzyłbyś się w zapasy piwa i chipsów, i zacząłbyś oglądać, klasyfikować i rejestrować. Mając skompletowane obserwacje mógłbyś zanalizować zebrane dane i sprawdzić, czy producenci artykułów dla mężczyzn sponsorowali więcej przelewu krwi niż inni reklamodawcy.

Analiza treści jako sposób obserwacji wymaga przemyślanego posługiwania się tym, „co" jest komunikowane. Analiza danych zebranych tą metodą, podobnie jak w przypadku innych metod, odnosi się do pytań „dlaczego" i „z jakim skutkiem".

Dobór próby w analizie treści

W badaniach przekazów, podobnie jak w badaniach ludzi, często nie można obserwować bezpośrednio wszystkiego, co chciałoby się zgłębić. W twoich badaniach, dotyczących przemocy telewizyjnej i sponsorowania, odradzam oglądanie wszystkiego, co jest nadawane. Byłoby to niemożliwe, a twój mózg prawdopodobnie wysiadłby wskutek krótkiego spięcia, zanim doszedłbyś do swojego odkrycia. Zwykle więc wskazane jest dobranie próby. Zacznijmy od przypomnienia pojęcia jednostek analizy, a następnie dokonamy przeglądu niektórych technik doboru próby, które można zastosować do jednostek analizy w analizie treści.

Jednostki analizy

W rozdziale 4 wspominałem, że określenie właściwych jednostek analizy – pojedynczych jednostek, na temat których formułujemy stwierdzenia opisowe i wyjaśniające – może być skomplikowanym zadaniem. Jeśli na przykład zamierzamy obliczyć średni dochód rodziny, jednostką analizy będzie pojedyncza rodzina. Musimy jednak zapytać poszczególnych członków rodziny, ile pieniędzy zarabiają. Tak więc pojedyncze osoby będą jednostkami *obserwacji*, chociaż indywidualna rodzina pozostanie jednostką analizy. Możemy też chcieć porównać wskaźniki przestępczości w różnych miastach, uwzględniając ich wielkość, region geograficzny, skład rasowy i inne różnice. Chociaż cechy tych miast będą częściowo funkcją zachowań i cech pojedynczych mieszkańców, jednostkami analizy byłyby ostatecznie miasta.

Złożoność tej kwestii jest często lepiej widoczna w przypadku analizy treści niż w przypadku innych metod badawczych, szczególnie kiedy jednostka obserwacji różni się od jednostki analizy. Kilka przykładów powinno wyjaśnić tę różnicę.

Załóżmy, że chcemy dowiedzieć się, czy prawo karne, czy też prawo cywilne w większym stopniu inaczej traktuje kobiety i mężczyzn. W tym przykładzie pojedyncze akty prawne byłyby zarazem jednostkami obserwacji i jednostkami analizy. Moglibyśmy dobrać próbę stanowych przepisów prawa karnego i cywilnego, a następnie zaklasyfikować każdy przepis ze względu na to, czy inaczej traktuje kobiety i mężczyzn, czy też nie. W ten sposób moglibyśmy określić, czy prawo karne, czy też cywilne bardziej różnicuje ze względu na płeć.

W nieco inny sposób badalibyśmy, czy stany, które ustanawiają prawa różnicujące poszczególne grupy rasowe, również częściej niż inne stany ustanawiają prawa różnie traktujące kobiety i mężczyzn. Chociaż badanie tego zagadnienia również wymagałoby kodowania pojedynczych aktów prawnych, to jednostką analizy w tym przypadku byłby pojedynczy stan, a nie prawo.

Załóżmy teraz, że jesteśmy zainteresowani malarstwem figuratywnym. Jeśli zechcemy porównać względną popularność malarstwa figuratywnego i niefiguratywnego, to jednostką analizy będzie pojedynczy obraz. Z kolei jeśli chcemy sprawdzić, czy malarstwo figuratywne jest bardziej charakterystyczne dla zamożnych, czy dla zubożałych malarzy, dla wykształconych czy niewykształconych, kapitalistycznych czy socjalistycznych, to jednostkami analizy będą poszczególni malarze.

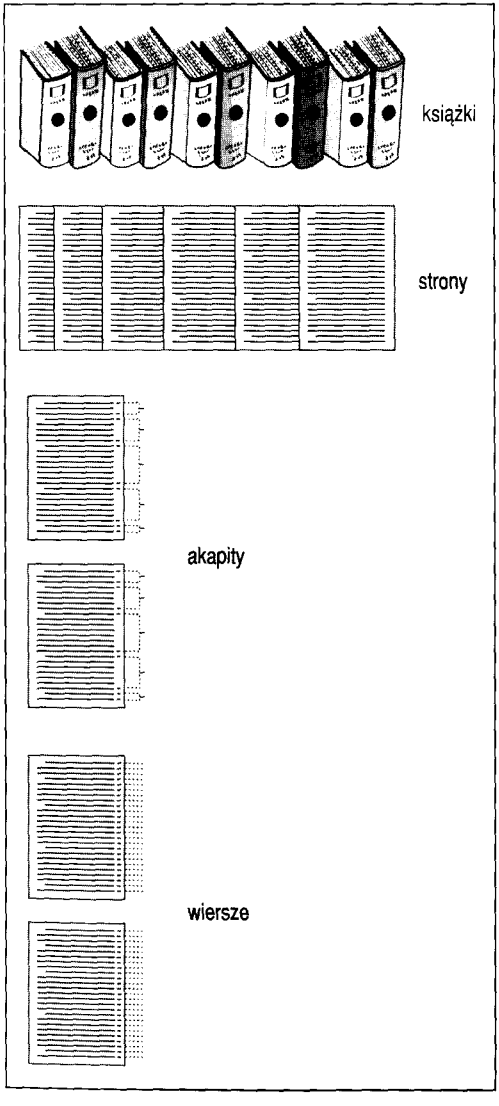
Podkreślmy wyraźnie: dobór próby w dużym stopniu zależy od tego, co jest jednostką analizy. Jeśli jednostką analizy są pojedynczy pisarze, należy włączyć do badania wszystkich pisarzy, których dotyczy pytanie badawcze, lub dobrać ich próbę. Jeśli jednostką analizy są książki, powinniśmy dobrać próbę książek, nie biorąc pod

uwagę ich autorów. Bruce Berg (1989, s. 112–113) wskazał, że gdy planuje się analizę pewnej ilości materiałów tekstowych, jednostką analizy mogą być słowa, tematy, postacie, akapity, całe pozycje (takie jak książki lub listy), pojęcia, semantyka, lub ich kombinacje. Rysunek 11.1 ilustruje niektóre z tych możliwości.

Nie chcę powiedzieć, że dobór próby powinien opierać się wyłącznie na jednostkach analizy. W rzeczywistości możemy często dobierać podpróby – próby z podkategorii – dla każdej pojedynczej jednostki analizy. Tak więc jeśli pisarze są jednostką analizy, możemy: (1) dobrać próbę pisarzy z całej populacji pisarzy, (2) dobrać próbę książek napisanych przez każdego z dobranych pisarzy i (3) dobrać fragmenty z każdej dobranej książki do obserwacji i kodowania.

Rysunek 11.1

Kilka możliwych jednostek analizy w analizie treści



Na koniec spójrzmy na bardziej skomplikowany przykład: badanie przemocy w telewizji i sponsorów. Co jest jednostką analizy dla następującego pytania badawczego: „Czy producenci produktów dla mężczyzn są bardziej niż inni reklamodawcy skłonni sponsorować programy telewizyjne zawierające przemoc?". Czy jednostką analizy jest program telewizyjny? Czy sponsor? Czy przypadek przemocy?

W najprostszym schemacie badawczym nie byłby to żaden z nich. Chociaż można ustrukturyzować te badania na różne sposoby, najprostszy schemat opierałby się na reklamach jako jednostkach analizy. Mógłbyś użyć dwóch rodzajów jednostek obserwacji: reklam i programu (mieszczącego się między reklamami). Obserwując obydwie jednostki, klasyfikowałbyś reklamy ze względu na to, czy dotyczą produktów dla mężczyzn, a programy ze względu na zawartą w nich przemoc. Klasyfikacje programów odnosiłbyś do pojawiających się przy nich reklam. Rysunek 11.2 dostarcza przykładu rodzaju zapisu, jaki mógłbyś otrzymać.

Zwróć uwagę, że w schemacie badawczym, zilustrowanym na rysunku 11.2, wszystkie reklamy nadawane w jednej przerwie reklamowej zostały wzięte w klamrę i otrzymały te same wyniki. Również liczba przypadków przemocy, zapisanych jako następujące po jednej przerwie reklamowej, jest taka sama jak liczba poprzedzająca następną przerwę. Ten prosty schemat pozwala nam zaklasyfikować każdą reklamę według jej sponsora i związanego z nią stopnia przemocy. Tak więc na przykład pierwsza reklama Grunt Aftershave została zakodowana jako odpowiadająca produktowi dla mężczyzn i powiązana z dziesięcioma przypadkami przemocy. Reklamę Buttercup Bra zakodowano jako przedstawiającą produkt nie dla mężczyzn i nie związaną z żadnymi przypadkami przemocy.

Na ilustracji umieszczono cztery reklamy produktów dla mężczyzn ze średnią 7,5 przy-

Rysunek 11.2

Przykład arkusza wyników dla przemocy w telewizji

Sponsor	Produkt dla mężczyzn?			Liczba przypadków przemocy	
	tak	nie	?	przed	po
Grunt Aftershave	✓			6	4
Brute Jock Straps	✓			6	4
Bald-No-More Lotion	✓			4	3
Grunt Aftershave	✓			3	0
Snowflake Toothpaste		✓		3	0
Godliness Cleanser		✓		3	0
Big Thumb Hammers			✓	0	1
Snowflake Toothpaste		✓		1	0
Big Thumb Hammers			✓	1	0
Buttercup Bras		✓		0	0

padków przemocy dla każdej z nich. Cztery reklamy sklasyfikowane jako zdecydowanie nie dotyczące produktów dla mężczyzn mają średnią 1,75, dwie zaś, które mogły, ale nie musiały dotyczyć produktów dla mężczyzn, miały średnio 1 przypadek przemocy na każdą z nich. Gdyby ten wzorzec zróżnicowań utrzymał się dla znacznie większej liczby obserwacji, prawdopodobnie moglibyśmy wyciągnąć wniosek, że wytwórcy produktów dla mężczyzn są bardziej od innych producentów skłonni sponsorować programy telewizyjne zawierające przemoc.

Celem tej ilustracji jest wykazanie, jak jednostki analizy kształtują zbieranie danych i ich analizę. Powinieneś mieć jasność co do swoich jednostek analizy przed zaplanowaniem strategii doboru próby, ale w tym przykładzie nie możesz po prostu dobrać próby reklam. Jeśli nie masz dostępu do planu emisji stacji nadającej, nie będziesz wiedział, kiedy reklamy mają się pojawić. Co więcej, podobnie jak reklamy, musisz oglądać programy z nimi związane. W efekcie musisz przygotować schemat doboru próby, który będzie zawierał wszystko, czego potrzebujesz do obserwacji.

Przy projektowaniu próby może być potrzebne określenie populacji, z której będzie ona dobierana. W naszym przykładzie, jakie stacje telewizyjne będziesz obserwował? Jaki będzie okres badania – liczba dni? I w jakich godzinach każdego dnia będziesz prowadził obserwację? Następnie, ile reklam chcesz obserwować i zakodować do analizy? Pooglądaj przez chwilę telewizję i zorientuj się, ile reklam pojawia się w każdej godzinie programu, następnie zastanów się, ilu godzin obserwacji będziesz potrzebować (i możesz na to poświęcić).

W tym momencie jesteś gotowy, aby zaprojektować schemat doboru próby. Uwaga praktyczna: nie musisz dobierać próby spośród

różnych stacji; jeśli masz pomocników, każdy z was mógłby oglądać inny kanał w tym samym czasie. Ale załóżmy, że pracujesz sam. Twój ostateczny operat losowania, za pomocą którego dobierzesz próbę i będziesz na niej prowadził obserwacje, może wyglądać mniej więcej tak:

7 stycznia, kanał 2, 19^{00}–21^{00}
7 stycznia, kanał 4, 19^{00}–21^{00}
7 stycznia, kanał 9, 19^{00}–21^{00}
7 stycznia, kanał 2, 21^{00}–23^{00}
7 stycznia, kanał 4, 21^{00}–23^{00}
7 stycznia, kanał 9, 21^{00}–23^{00}
8 stycznia, kanał 2, 19^{00}–21^{00}
8 stycznia, kanał 4, 19^{00}–21^{00}
8 stycznia, kanał 9, 19^{00}–21^{00}
8 stycznia, kanał 2, 21^{00}–23^{00}
8 stycznia, kanał 4, 21^{00}–23^{00}
8 stycznia, kanał 9, 21^{00}–23^{00}
9 stycznia, kanał 2, 19^{00}–21^{00}
9 stycznia, kanał 4, 19^{00}–21^{00}
itd.

Zauważ, że w tym przykładzie podjąłem za ciebie kilka decyzji. Po pierwsze założyłem, że jedynymi kanałami odpowiednimi dla twoich badań są kanały 2, 4 i 9. Założyłem, że uznałeś godziny najlepszego czasu antenowego 19^{00}–23^{00} za najbardziej odpowiednie i że praca będzie wykonywana w dwugodzinnych odcinkach. Siódmego stycznia jako datę rozpoczęcia badania wybrałem na chybił trafił. Oczywiście w praktyce wszystkie decyzje tego rodzaju powinny być podejmowane po starannym rozważeniu, co będzie odpowiednie dla tych właśnie badań.

Jeśli masz już jasność co do jednostek analizy i najlepiej dostosowanych do nich obserwacji oraz stworzyłeś oparat losowania podobny do tego, który zaprezentowałem, dobór próby jest prosty i bezpośredni. Wybór dostępnych procedur jest taki sam, jak opisany w rozdziale 7: dobór losowy prosty, systematyczny, warstwowy itd.

Techniki doboru próby

Jak mogliśmy zobaczyć, w analizie treści dzieł pisanych dobór próby może występować na każdym pojedynczym poziomie lub na wszystkich poziomach, łącznie z kontekstem powiązanym z dziełem. Dobór próby może dotyczyć również innych form przekazu na każdym z właściwych im poziomów pojęciowych.

W analizie treści moglibyśmy użyć każdej z konwencjonalnych technik doboru próby omówionych w rozdziale 7. Możemy wylosować prostą lub systematyczną próbę powieściopisarzy francuskich lub amerykańskich, próbę praw uchwalonych w stanie Missisipi lub monologów szekspirowskich. Możemy wybrać (zaczynając od losowego punktu startowego) co 23 akapit z *Wojny i pokoju* Tołstoja lub ponumerować wszystkie piosenki nagrane przez Beatlesów i wybrać losową próbę 25 z nich.

Warstwowy dobór próby jest także odpowiedni dla analizy treści. By analizować politykę wydawniczą amerykańskich gazet, możemy na przykład najpierw pogrupować wszystkie gazety pod kątem regionów kraju lub wielkości miejscowości, w której się one ukazują, częstotliwości publikowania lub średniego nakładu. Następnie możemy dobrać do analizy próbę losową gazet warstwową lub systematyczną. Później możemy dobrać próbę artykułów redakcyjnych z każdej wybranej gazety, być może powarstwowanych chronologicznie.

Grupowy dobór próby jest równie odpowiedni dla analizy treści. Właściwie jeśli pojedyncze artykuły redakcyjne byłyby naszymi jednostkami analizy, selekcja gazet na pierwszym etapie doboru próby byłaby doborem grupowym. W analizie przemówień politycznych możemy zacząć od doboru próby polityków; każdy polityk mógłby reprezentować „grupę" przemówień politycznych. Badania reklam telewizyjnych opisane wcześniej są kolejnym przykładem próby grupowej.

Trzeba powtórzyć, że dobieranie próby nie musi być zakończone, kiedy dojdziemy do jednostek analizy. Jeśli powieści są w danym badaniu jednostkami analizy, możemy wybrać próbę powieściopisarzy, podpróby powieści autorstwa każdego z wybranych pisarzy i podpróbę akapitów z każdej z powieści. Następnie możemy analizować treść tych akapitów w celu opisania samych powieści. (*Uwaga*: badacze mówią o próbach pobieranych z prób jako o „podpróbach").

Przejdźmy teraz do kodowania i klasyfikacji materiału z obserwacji. Część 4 omawia

operowanie takimi klasyfikacjami w celu wyciągania wniosków o charakterze opisowym i wyjaśniającym.

Kodowanie w analizie treści

Ogólnie rzecz biorąc, istotą analizy treści jest **kodowanie**. *Kodowanie* to proces przekształcania surowych danych na dane w postaci standaryzowanej. W analizie treści przekazy – ustne, pisane lub inne – są kodowane lub klasyfikowane według pewnych ram pojęciowych. Tak więc na przykład artykuły redakcyjne w gazetach mogą być zakodowane jako liberalne lub konserwatywne. Audycje radiowe mogą być zakodowane jako propagandowe lub niepropagandowe, powieści jako romantyczne lub nie, malarstwo jako figuratywne lub niefiguratywne, a przemówienia polityczne jako naszpikowane atakami personalnymi lub ich pozbawione. Przypomnij sobie, że terminy takie jak te mogą być przedmiotem wielu interpretacji, a badacz musi jasno określić ich definicje.

Kodowanie w analizie treści wymaga także działania zgodnego z logiką konceptualizacji i operacjonalizacji, omówioną w rozdziale 5. Tak jak w innych metodach badawczych, musisz udoskonalić ramy pojęciowe i wypracować konkretne związane z nimi metody obserwacji.

Treści jawne i ukryte

We wcześniejszym omówieniu badań terenowych stwierdziliśmy, że badacz staje przed fundamentalnym wyborem między głębią wyjaśniania a jego ścisłością. Często odpowiada to wyborowi między – odpowiednio – trafnością a rzetelnością. Na ogół badacze terenowi opowiadają się za głębią, woląc oprzeć swoje sądy na szerokim zakresie obserwacji i informacji, nawet ryzykując, że inny obserwator mógłby inaczej ocenić tę samą sytuację. Badania sondażowe, dzięki zastosowaniu standaryzowanych kwestionariuszy, reprezentują drugą skrajność: pełną ścisłość, nawet jeśli ścisłe mierniki zmiennych mogą nie odzwierciedlać odpowiednio trafnie tych zmiennych. Badacz posługujący się analizą treści ma jednak w tej sprawie pewną możliwość wyboru.

Kodowanie **treści jawnych** – widocznej, zewnętrznej treści przekazu – jest analogiczne do używania standaryzowanego kwestionariusza. Na przykład dla określenia, ile jest erotyki w pewnych powieściach, można po prostu zliczyć, ile razy występuje w każdej powieści słowo *miłość* lub ile razy średnio pojawia się ono na stronie. Można też użyć listy słów, takich jak *miłość*, *pocałunek*, *przytulenie* i *pieszczota*; każde z nich może służyć jako wskaźnik erotycznego charakteru powieści. Ta metoda miałaby zaletę polegającą na łatwości i rzetelności kodowania oraz dostarczenia czytelnikowi raportów badawczych precyzyjnej informacji, jak był mierzony erotyzm. Niemniej jednak ma ona wadę związaną z trafnością. Z całą pewnością określenie *powieść erotyczna* ma bogatsze i głębsze znaczenie niż liczba użytych powtórzeń słowa *miłość*.

Alternatywą jest kodowanie **treści ukrytych** przekazu: jego głębokiego znaczenia. W przywoływanym tu przykładzie mógłbyś przeczytać całą powieść lub próbę dobraną spośród akapitów albo stron i całościowo ocenić, na ile ta powieść jest erotyczna. Chociaż twoja ocena mogła być oczywiście uwarunkowana przez występowanie słów, takich jak *miłość* i *pocałunek*, nie zależałaby jednak w pełni od częstotliwości ich występowania.

Rzecz jasna, ta druga metoda wydaje się bardziej przydatna do wydobywania głębokiego znaczenia przekazu, lecz dzieje się to kosztem rzetelności i ścisłości. Zwłaszcza w sytuacjach, gdy więcej niż jedna osoba koduje powieść, może dojść do zastosowania nieco innych definicji lub standardów. Urywek tekstu, który jeden koder potraktuje jako erotyczny, drugiemu wcale się taki nie wyda. Nawet jeśli sam wykonasz całe kodowanie, nie ma gwarancji, że twoje definicje i standardy pozostaną stałe podczas całego przedsięwzięcia. Ponadto czytelnik twojego raportu badawczego mógłby nie być pewien, jakiej definicji używałeś.

Najlepszym rozwiązaniem tego dylematu jest używanie gdziekolwiek to możliwe obu metod. Na przykład Carol Auster była zainteresowana zmianami w socjalizacji młodych kobiet w Żeńskiej Organizacji Skautowej. Aby to zbadać, zastosowała analizę treści podręczników dla dziewcząt skautek w kolejnych zmienianych

wydaniach. Auster była szczególnie zainteresowana poglądem, że kobiety powinny ograniczać się do czynności domowych. Jej analiza treści jawnych wskazywała na zmianę: „zauważyłam, że w 1913 roku 23% odznak skupiało się na życiu rodzinnym, w 1963 roku było tak natomiast jedynie w przypadku 13% odznak, a w 1980 roku – 7% odznak" (1985, s. 361).

Analiza treści ukrytych również wskazywała na emancypację w tej organizacji, podobną do tej, jaka zachodziła w całym społeczeństwie amerykańskim. Jednym ze wskaźników była zmiana mundurów: „przejście od spódnic do spodni może odzwierciedlać uznanie bardziej aktywnej fizycznie roli kobiet, jak również różnorodności wizerunków zewnętrznych, dostępnych współczesnej kobiecie" (Auster 1985, s. 362). Wspierający dowód został odnaleziony w pojawieniu się odznak, takich jak „detektyw naukowy", „przestrzeń powietrzna" i „pani złota rączka".

Konceptualizacja i tworzenie kategorii kodowych

W przypadku wszystkich metod badawczych konceptualizacja i operacjonalizacja na ogół wymagają powiązania kwestii teoretycznych i obserwacji empirycznych. Jeśli na przykład uważasz, że pewne artykuły redakcyjne są liberalne, a inne konserwatywne, zastanów się, dlaczego tak uważasz. Przeczytaj kilka artykułów redakcyjnych i zadaj sobie pytanie, które z nich są liberalne, a które konserwatywne. Czy orientacja polityczna konkretnego artykułu wyraźniej zaznaczała się w jego jawnej treści, czy też w jego tonie? Czy twoja decyzja opierała się na używaniu pewnych określeń (na przykład *lewicowy*, *faszystowski* itp.), czy na poparciu dla poszczególnych kwestii czy polityków lub przeciwstawianiu się im?

W tym zadaniu należy zastosować zarówno metodę indukcyjną, jak i dedukcyjną. Gdy testuje się założenia teoretyczne, to one powinny decydować o doborze wskaźników pojęć. Gdy zaczyna się od określonych obserwacji empirycznych, należy dążyć do wyprowadzenia z nich ogólnych reguł powiązanych z tymi obserwacjami, a następnie do zastosowania tych reguł do innych obserwacji empirycznych.

Bruce Berg (1989, s. 111) lokuje tworzenie kodów w kontekście teorii ugruntowanej i porównuje je do układania układanki:

> „Kodowanie i inne fundamentalne procedury, związane z budowaniem teorii ugruntowanej, niewątpliwie są ciężką pracą i muszą być traktowane poważnie, jednak podobnie jak wielu ludzi cieszy się, kończąc skomplikowane układanie puzzli, wielu badaczy znajduje ogromną satysfakcję w kodowaniu i analizie. Kiedy badacz (...) zaczyna widzieć coraz pełniejszy obraz wyłaniający się z połączonych fragmentów układanki, proces ten może być bardzo ekscytujący".

Podczas tych czynności pamiętaj, że definicja operacyjna każdej zmiennej jest złożona z wartości, które ją tworzą. Co więcej, wartości te powinny być wyczerpujące i wykluczające się wzajemnie. Na przykład artykuł redakcyjny w gazecie nie może być opisany jednocześnie jako liberalny i konserwatywny, chociaż można by zapewne dopuścić, że niektóre uplasują się gdzieś między tymi określeniami. Dla twoich celów wystarczyłoby zakodowanie powieści jako erotycznych lub nieerotycznych, lecz mógłbyś również rozważyć, czy niektóre nie mogłyby być antyerotyczne. Obrazy mogą być zaklasyfikowane jako figuratywne lub nie, jeśli zaspokaja to twoje cele badawcze, ale możesz też chcieć klasyfikować je jako impresjonistyczne, abstrakcyjne, alegoryczne itd.

Musisz uświadomić sobie, że w analizie treści mogą być stosowane różne poziomy pomiaru. Można na przykład zastosować kategorie nominalne: liberalny i konserwatywny, w celu scharakteryzowania artykułów redakcyjnych w gazetach, lub można zechcieć użyć bardziej subtelnego rankingu porządkowego, od skrajnie liberalnego do skrajnie konserwatywnego. Pamiętaj jednak, że poziom pomiaru, wynikający z przyjętej przez ciebie metody kodowania – nominalny, porządkowy, interwałowy czy ilorazowy – nie musi odzwierciedlać charakteru twoich zmiennych. Jeśli słowo *miłość* pojawi się 100 razy w powieści A i 50 razy w powieści B, masz prawo powiedzieć, że słowo miłość pojawia się dwa razy częściej w powieści A, lecz nie że powieść A jest dwa razy bardziej erotyczna

niż powieść B. Podobnie to, że badany zgadza się z dwukrotnie większą liczbą stwierdzeń antysemickich, zawartych w kwestionariuszu, niekoniecznie sprawia, że jest on dwa razy bardziej antysemitą.

Zliczanie i zachowywanie zapisów

Jeśli zamierzasz oceniać swoje dane z analizy treści w sposób ilościowy, operację kodowania trzeba dostosować do przetwarzania danych. Oznacza to po pierwsze, że końcowy produkt kodowania musi mieć charakter liczbowy. Jeśli zliczasz częstość pewnych słów, zwrotów lub innej treści jawnej, kodowanie z konieczności będzie liczbowe. Lecz nawet jeśli kodujesz treści ukryte na podstawie całościowych ocen, trzeba będzie zapisać twoje decyzje dotyczące kodowania w postaci liczbowej: 1 = bardzo liberalny, 2 = umiarkowanie liberalny, 3 = umiarkowanie konserwatywny itd.

Po drugie, sposób zapisywania musi czynić jasne rozróżnienia jednostek analizy i jednostek obserwacji, zwłaszcza jeśli są one różne. Początkowe kodowanie oczywiście musi odnosić się do jednostek obserwacji. Jeśli na przykład powieściopisarze są jednostkami analizy, a zamierzasz scharakteryzować ich za pomocą analizowania treści ich powieści, początkowy zapis będzie odpowiadał powieściom jako jednostkom obserwacji. Następnie możesz łączyć wyniki dla pojedynczych powieści, by scharakteryzować każdego pisarza jako jednostkę analizy.

Po trzecie, w trakcie zliczania będzie na ogół ważne zarejestrowanie podstawy dokonywania obliczeń. Niewiele dałaby nam informacja o liczbie obrazów realistycznych namalowanych przez danego malarza, gdybyśmy nie wiedzieli, ile obrazów namalował w sumie; malarza uznalibyśmy za realistycznego, gdyby znaczny procent jego obrazów stanowiły obrazy tego rodzaju. Równie niewiele powiedziałaby nam informacja, że słowo *miłość* występuje 87 razy w powieści, gdybyśmy nie wiedzieli, z ilu słów składa się cała powieść. Kwestię podstawy obserwacji najłatwiej jest rozwiązać, jeśli każdy przypadek kodowany jest w kategoriach jednej z wartości składających się na zmienną. Na przykład zamiast po prostu zliczać liczbę liberalnych artykułów redakcyjnych w danym zbio-

Rysunek 11.3

Próbka arkusza wyników (fragment)

Numer identyfikacyjny gazety	Liczba ocenianych artykułów redakcyjnych	SUBIEKTYWNA OCENA 1) bardzo liberalny 2) umiarkowanie liberalny 3) pośredni 4) umiarkowanie konserwatywny 5) bardzo konserwatywny	Liczba artykułów „izolacjonistycznych"	Liczba artykułów „pro-oenzetowskich"	Liczba artykułów „anty-oenzetowskich"
001	37	2	0	8	0
002	26	5	10	0	6
003	44	4	2	1	2
004	22	3	1	2	3
005	30	1	0	6	0

rze, lepiej jest kodować każdy z nich pod kątem jego politycznej orientacji, nawet jeśli trzeba zakodować „bez wyraźnej orientacji".

Załóżmy, że chcemy opisać i wyjaśnić politykę wydawniczą różnych gazet. Rysunek 11.3 przedstawia część arkusza wyników, który można by otrzymać w rezultacie kodowania artykułów redakcyjnych gazet. Zauważ, że jednostkami analizy są gazety. Każdej gazecie został przypisany numer identyfikacyjny, by ułatwić automatyczne przetwarzanie. Druga kolumna zawiera miejsce na liczbę artykułów redakcyjnych, zakodowanych dla każdej gazety. To będzie istotny fragment informacji, gdyż chcemy móc stwierdzić na przykład: „ze wszystkich artykułów redakcyjnych 22% popierało Narody Zjednoczone", a nie tylko: „było osiem artykułów popierających Narody Zjednoczone".

Jedna z kolumn na rysunku 11.3 służy przypisaniu całościowej subiektywnej oceny polityki wydawniczej gazety (może być ona później porównana z poszczególnymi obiektywnymi miernikami). Pozostałe kolumny to miejsce dla zapisania liczby artykułów redakcyjnych odzwierciedlających konkretne stanowisko redakcji. W rzeczywistej analizie treści mogłoby tam być miejsce na rejestrację innych stanowisk redakcyjnych oraz nieredakcyjnych informacji na temat każdej gazety, takich jak region, w którym jest wydawana, jej nakład itd.

Analiza danych jakościowych

Nie zawsze wynikiem analizy treści są obliczenia. Czasami najbardziej odpowiednia jest jakościowa ocena materiału, tak jak w badaniach zmian w umundurowaniu i w języku podręczników Żeńskiej Organizacji Skautowej, przeprowadzonych przez Carol Auster.

Bruce Berg (1989, s. 123–125) omawia „testowanie przypadków negatywnych" jako technikę jakościowego testowania hipotez. Po pierwsze, zgodnie z tradycją teorii ugruntowanej, rozpoczyna się od badania danych, które mogą dostarczyć ogólnej hipotezy. Powiedzmy, że za pomocą przeglądania protokołów z zebrań badasz przywództwo nowego stowarzyszenia w społeczności lokalnej, aby zobaczyć, kto zgłaszał wnioski, które potem zostały przyjęte. Początkowe badania danych sugerowały, że zamożniejsi członkowie najczęściej podejmują role przywódcze.

Drugi etap analizy to poszukiwanie danych o wszystkich przypadkach, które mogłyby podważyć hipotezę początkową. W tym przypadku będziesz wyszukiwał biedniejszych członków, których wnioski zostały przyjęte, oraz bogatych, którym się to nigdy nie udało. Po trzecie, musisz przejrzeć każdy przypadek podważający hipotezę i albo (1) zrezygnować z tej hipotezy, albo (2) przyjrzeć się, jak można ją zmodyfikować.

Powiedzmy, że w trakcie analizy przypadków sprzecznych z hipotezą zauważyłeś, że każdy z niezamożnych przywódców ma wykształcenie wyższe, natomiast każdy z zamożnych nie-przywódców ma bardzo niskie formalne wykształcenie. Możesz zrewidować swoją hipotezę, tak by uwzględniała ona zarówno wykształcenie, jak i poziom zamożności jako drogi do przywództwa w stowarzyszeniu. Być może odkryjesz pewne progi dla przywództwa (pracę umysłową, pewien poziom dochodów, pewien poziom wykształcenia), powyżej których ci z największymi pieniędzmi, najwyższym wykształceniem lub jednym i drugim są najbardziej aktywnymi liderami.

Proces ten jest przykładem tego, co Barney Glaser i Anselm Strauss (1967) nazywali *indukcją analityczną*. Jest to indukcja w tym sensie, że zaczyna się od obserwacji, i jest analityczna, ponieważ wychodzi poza opis, by znaleźć wzorce i związki między zmiennymi.

Są oczywiście niebezpieczeństwa związane z tą formą analizy, podobnie jak z wszystkimi pozostałymi. Najważniejsze ryzyko wiąże się z błędnym zaklasyfikowaniem obserwacji, tak by potwierdzały wyłaniające się hipotezy. Na przykład można błędnie wnioskować, że nie-przywódca nie ukończył szkoły wyższej, lub można zdecydować, że praca brygadzisty w fabryce jest „wystarczająco bliska" pracy umysłowej.

Berg (1989, s. 124) proponuje techniki służące unikaniu takich błędów:

1. Jeśli liczba przypadków jest wystarczająca, wybierz losowo po kilka z każdej katego-

rii, w celu uniknięcia wybierania tylko tych, które najlepiej potwierdzają hipotezy.

2. Daj co najmniej trzy przykłady na poparcie każdego twierdzenia, jakie formułujesz na temat danych.

3. Poddawaj swoje interpretacje analityczne dokładnemu sprawdzeniu przez inne osoby, nie włączone w projekt badawczy, by przekonać się, czy zgodzą się z tobą.

4. Opisuj wszystkie niezgodności, które odkryjesz – każdy przypadek, który nie przystaje do twoich hipotez. Zauważ, że niewiele prawidłowości społecznych jest spójnych w stu procentach, tak więc możesz odkryć coś ważnego, nawet jeśli nie ma to zastosowania do absolutnie całego życia społecznego. Powinieneś jednak uczciwie informować o tym swoich czytelników.

Przykład analizy treści

Kilka projektów badawczych wskazywało, że kobiety są stereotypowo przedstawiane w telewizji. R. Stephen Craig (1992) poprowadził te badania o krok dalej w celu zbadania obrazów zarówno kobiet, jak i mężczyzn, śledząc programy telewizyjne w różnych okresach.

Dla zbadania stereotypów płci w reklamach telewizyjnych Craig wybrał próbę 2209 reklam sieciowych, emitowanych w kilku okresach między 6 a 14 stycznia 1990 roku.

> „Część obejmująca dni robocze (w tej próbie były to okresy od poniedziałku do piątku w godzinach 14^{00}–16^{00}) składała się wyłącznie z oper mydlanych i została wybrana ze względu na wysoki odsetek żeńskiej widowni. Część weekendowa (dwa kolejne sobotnie i niedzielne popołudnia podczas sportowych transmisji telewizyjnych) została wybrana ze względu na wysoki odsetek widowni męskiej. Wieczorny czas najlepszej oglądalności (od poniedziałku do piątku w godzinach 21^{00}–23^{00}) został wybrany jako baza do porównań z badaniami prowadzonymi w przeszłości i z innymi".
>
> (1992, s. 199)

Każda z reklam została zakodowana na kilka sposobów. „Postacie" były kodowane jako:

Wszyscy dorośli mężczyźni
Wszystkie dorosłe kobiety
Wszyscy dorośli, płeć mieszana
Dorośli mężczyźni z dziećmi lub nastolatkami (bez kobiet)
Dorosłe kobiety z dziećmi lub nastolatkami (bez mężczyzn)
Mieszana płeć i wiek

W dodatku koderzy Craiga notowali, która postać była podczas reklamy najdłużej na ekranie – oznaczono ją jako „główną postać wizualną" – podobnie jak role odgrywane przez postacie (współmałżonek, znakomitość, rodzic), typ reklamowanego produktu (na przykład produkt do ciała, alkohol), otoczenie (kuchnia, szkoła, przedsiębiorstwo) i komentarz narratora.

Tabela 11.1

Procent dorosłych pierwszoplanowych postaci wizualnych, z uwzględnieniem płci, pojawiających się w reklamach w trzech porach dnia

	Dzień	Wieczór	Weekend
Dorośli mężczyźni	40	52	80
Dorosłe kobiety	60	48	20

Źródło: Stephen R. Craig, *The Effect of Television Day Part on Gender Portrayals in Television Commercials: A Content Analysis*, „Sex Roles" 1992, 26, nr 5/6, s. 204.

Tabela 11.1 wskazuje różnice w porach, kiedy kobiety i mężczyźni pojawiają się w reklamach. Kobiety pojawiają się najczęściej podczas dnia (przy okazji emitowanych w ciągu dnia oper mydlanych), mężczyźni dominują podczas reklam weekendowych (w trakcie emitowanych w tym czasie programów sportowych), natomiast podczas wieczornego czasu najlepszej oglądalności i mężczyźni, i kobiety byli reprezentowani równomiernie.

Craig odnalazł też inne różnice w sposobie przedstawiania kobiet i mężczyzn.

> „Dalsze analizy wskazywały, że główne postacie męskie były proporcjonalnie częściej niż żeńskie przedstawiane jako znakomitości i profesjonaliści, niezależnie od pory dnia, podczas gdy kobiety proporcjonalnie częściej przedstawione były w roli osób przeprowadzających wy-

wiady/prezenterek, matek/żon lub obiektów seksualnych/modelek, niezależnie od pory dnia. (...) Z większym prawdopodobieństwem kobiety pojawiały się jako obiekty seksualne/modelki podczas weekendów niż w zwykłe dni".

(1992, s. 204)

Badania wykazały również, że różne produkty były reklamowane podczas różnych okresów. Jak można się domyślać, prawie wszystkie reklamy nadawane w ciągu dnia dotyczyły ciała, żywności lub produktów domowych. Tych produktów dotyczyła tylko co trzecia reklama w weekendy. Zamiast tego reklamy weekendowe kładły nacisk na produkty samochodowe (29%), produkty i usługi dla biznesu (27%) lub alkohol (10%). Reklam alkoholu praktycznie nie było podczas wieczorów i w ciągu dnia.

Jak można było oczekiwać, kobiety były najczęściej przedstawiane w otoczeniu domowym, mężczyźni zaś poza domem. Inne odkrycia dotyczyły różnych ról odgrywanych przez kobiety i przez mężczyzn.

„Kobiety, które pojawiały się w reklamach weekendowych, prawie nigdy nie były przedstawiane bez mężczyzn i rzadko jako główna postać reklamy. Generalnie były widziane w rolach uzupełniających wobec mężczyzn (na przykład recepcjonistki hotelowej, sekretarki lub stewardessy) lub jako obiekty seksualne lub modelki, kiedy to ich jedyną funkcją wydawało się być nadawanie reklamie aspektu erotycznego".

(1992, s. 208)

Chociaż niektóre odkrycia Craiga mogą wydawać się mało zaskakujące, trzeba pamiętać, że „wiedza potoczna" nie zawsze odpowiada rzeczywistości i dlatego zawsze warto sprawdzać powszechnie podzielane przypuszczenia. Nawet jeśli sądzimy, że znamy daną sytuację, zawsze przydatne jest poznanie bardziej specyficznych detali, takich jak te dostarczane przez analizę treści, podobną do zaprezentowanej wyżej.

Mocne i słabe strony analizy treści

Chyba największą zaletą analizy treści jest związana z nią oszczędność zarówno czasu, jak i pieniędzy. Student może dokonać analizy treści, podczas gdy na przykład przeprowadzenie sondażu mogłoby być niewykonalne. Nie wymaga ona licznego personelu badawczego ani specjalnego sprzętu. Jeśli tylko masz dostęp do nadającego się do zakodowania materiału, możesz podjąć się analizy treści.

Kolejną zaletą analizy treści jest bezpieczeństwo. Jeśli spostrzeżesz, że sknociłeś sondaż lub eksperyment, możesz być zmuszony do powtórzenia całego projektu badawczego z wszystkimi związanymi z tym kosztami, wyrażonymi w czasie i pieniądzu. Jeśli sknociłeś badania terenowe, ponowne przeprowadzenie projektu może być niemożliwe, badane wydarzenia mogą już nie mieć miejsca. W analizie treści jest zwykle łatwiej powtórzyć część badań niż w innych metodach badawczych. Co więcej, możesz być zmuszony do rekodowania jedynie części swoich danych, a nie całości.

Trzecią zaletą analizy treści jest to, że dopuszcza badanie długotrwałych procesów. Na przykład można skupić się na wizerunkach Afroamerykanów zawartych w amerykańskich powieściach z lat 1850–1860 lub można badać zmianę ich wizerunku od 1850 roku do współczesności.

W końcu analiza treści ma zaletę wszystkich mierników niereaktywnych, a mianowicie tę, że badacz rzadko ma jakikolwiek wpływ na przedmiot badań. Ponieważ powieści zostały już napisane, obrazy namalowane, przemówienia wygłoszone, analiza treści nie może mieć na nie wpływu.

Analiza treści ma również wady. Jest mianowicie ograniczona do zapisanych przekazów. Takie przekazy mogą być ustne, pisane lub graficzne, lecz muszą być zapisane w sposób pozwalający na analizę.

Jak widzieliśmy, analiza treści ma zarówno zalety, jak i wady w sensie trafności i rzetelności. Problemy z trafnością mogą się pojawić, jeśli przedmiotem badań nie jest sam proces komunikacji jako taki.

Przechodząc na drugą stronę bilansu, konkretność materiałów badanych w analizie treści wzmacnia prawdopodobieństwo jej rzetelności. Można zawsze kodować i rekodować, a nawet rekodować ponownie, jeśli chce się uzys-

kać pewność, że kodowanie jest spójne. W badaniach terenowych przeciwnie, z reguły nie można niczego zrobić po fakcie, by zapewnić większą rzetelność obserwacji i kategoryzacji.

Zostawmy teraz analizę treści i przejdźmy do pokrewnej metody badawczej: analizy istniejących danych. Chociaż w tym przypadku analizowaną treścią są bardziej liczby niż przekazy, myślę, że dostrzeżesz podobieństwo do analizy treści.

Analiza istniejących danych statystycznych[1]

Często można lub trzeba podejmować badania z zakresu nauk społecznych poprzez wykorzystanie oficjalnych lub quasi-oficjalnych statystyk. Nie jest to tym samym, co analiza wtórna, kiedy uzyskuje się kopię danych kogoś innego i przeprowadza swoje własne analizy statystyczne. W tym podrozdziale zamierzamy spojrzeć na sposoby wykorzystania analiz danych, przeprowadzonych wcześniej przez innych.

Metoda ta jest szczególnie istotna ze względu na to, że istniejące statystyki powinny być zawsze brane pod uwagę, co najmniej jako uzupełniające źródło danych. Jeśli planowałeś sondaż na temat postaw politycznych, dobrze zrobiłbyś, badając i prezentując swoje odkrycia w kontekście wzorców głosowania, frekwencji wyborczej lub podobnych statystyk, powiązanych z obszarem twoich zainteresowań badawczych. Jeśli przeprowadziłeś badania ewaluacyjne, dotyczące eksperymentalnego programu poprawiającego nastroje wśród pracowników linii montażowej, statystyki absencji, zwolnień chorobowych itp. byłyby dla ciebie prawdopodobnie interesujące i ujawniłyby istotne treści, związane z danymi pochodzącymi z twoich badań. Tak więc istniejące dane statystyczne często mogą dostarczyć kontekstu historycznego czy pojęciowego, w którym można umieścić oryginalne badania.

[1] W polskiej literaturze metodologicznej spotkać się można z równoważnym określeniem „analiza danych zastanych" (przyp. red. nauk.).

Istniejące statystyki mogą również być źródłem zasadniczych danych dla naukowych badań społecznych. Doskonałym przykładem są klasyczne badania wspomniane na początku tego rozdziału, *Le Suicide* Émile'a Durkheima ([1897] 1951[2]). Przyjrzyjmy się bliżej pracy Durkheima, zanim rozważymy niektóre ze szczególnych problemów prezentowanych przez tę metodę.

Badania nad samobójstwami Durkheima

Dlaczego ludzie się zabijają? Niewątpliwie każdy przypadek samobójstwa ma swoją własną historię i wyjaśnienie, jednak wszystkie takie przypadki mogłyby bez wątpienia zostać pogrupowane zgodnie z pewnymi wspólnymi przyczynami: niepowodzenie finansowe, kłopoty w miłości, okrycie się hańbą i inne rodzaje problemów osobistych. Gdy francuski socjolog Émile Durkheim zajął się kwestią samobójstw, miał na myśli trochę inne pytanie. Chciał odkryć warunki środowiska, szczególnie zaś warunki społeczne, które zachęcają lub zniechęcają do tego kroku.

Im dłużej Durkheim badał dostępne rejestry, tym więcej dostrzegał wzorców różnic. Jedną z pierwszych rzeczy, które przyciągnęły jego uwagę, była względna stałość wskaźnika samobójstw. Patrząc na poszczególne kraje odkrył, że wskaźniki samobójstw są mniej więcej takie same każdego roku. Odkrył również, że nieproporcjonalnie wiele samobójstw jest popełnianych latem, co doprowadziło go do hipotezy, że temperatura może mieć jakiś związek z samobójstwem. Jeśli tak by było, wskaźnik samobójstw powinien być wyższy w krajach Europy Południowej aniżeli w krajach o klimacie umiarkowanym. Durkheim odkrył jednak, że najwyższy wskaźnik występował w krajach o centralnym położeniu geograficznym, temperatura więc nie mogła posłużyć za wyjaśnienie.

Badał rolę wieku (najczęstszym wiekiem samobójcy było 35 lat), płeć (mężczyźni przeważali

[2] W Polsce ukazały się fragmenty tej pracy w tłumaczeniu Tadeusza Szawiela w: *Logika analizy socjologicznej* pod redakcją Antoniego Sułka, Warszawa 1979 (przyp. tłum.).

nad kobietami w stosunku mniej więcej cztery do jednego) i wiele innych czynników. Ostatecznie ogólny wzorzec wyłonił się z różnych źródeł.

Jeśli chodzi o stabilność samobójstw w czasie, Durkheim stwierdził, że wzorzec ten nie był całkowicie stabilny. Zdarzały się nagłe wzrosty wskaźnika podczas okresów zamieszania politycznego, jakie występowało w różnych krajach europejskich około 1848 roku. Ta obserwacja zaprowadziła go ku hipotezie, że samobójstwa mogły mieć coś wspólnego z „wyłomami w równowadze społecznej". Mówiąc inaczej, stabilność i integracja społeczna są, jak się wydaje, ochroną przed samobójstwami.

Ta ogólna hipoteza była udowadniana i precyzowana w trakcie analizowania przez Durkheima różnych zestawów danych. Wskaźniki samobójstw drastycznie różniły się w różnych krajach Europy. W Saksonii na przykład wskaźnik ten był około 10 razy większy niż we Włoszech, lecz względny ranking różnych krajów był niezmienny w czasie. W miarę jak Durkheim rozważał inne odmienności w różnych krajach, ostatecznie zauważył uderzający wzorzec: kraje zdominowane przez protestantów miały systematycznie wyższy wskaźnik samobójstw niż kraje katolickie. Kraje zdominowane przez protestantów miały 190 samobójstw na milion mieszkańców, kraje mieszane protestancko-katolickie 96, a zdominowane przez katolików 58 (Durkheim [1897] 1951, s. 152).

Chociaż wydawało się, że wskaźnik samobójstw jest zatem powiązany z religią, Durkheim był zdania, że jakiś inny czynnik, taki jak poziom rozwoju ekonomicznego i kulturalnego mógłby wyjaśniać obserwowane różnice między krajami. Gdyby religia miała faktyczny wpływ na samobójstwa, różnice religijne powinny zostać ujawnione również *wewnątrz* danych krajów. By sprawdzić ten pomysł, Durkheim po pierwsze zauważył, że niemieckie państwo Bawaria równocześnie było w większości katolickie oraz miało najniższy wśród badanych krajów wskaźnik samobójstw, natomiast zdecydowanie protestanckie Prusy miały znacznie wyższy wskaźnik samobójstw. Nie chcąc na tym poprzestać, Durkheim przebadał prowincje składające się na każde z tych państw.

Tabela 11.2

Wskaźniki samobójstw w różnych niemieckich prowincjach, uporządkowane pod względem przynależności religijnej

Religijny charakter prowincji	Samobójstwa na milion mieszkańców
Prowincje Bawarii (1867–1875)*	
Poniżej 50% katolików	
Palatynat Reński	167
Środkowa Frankonia	207
Górna Frankonia	204
Średnia	192
Od 50 do 90% katolików	
Dolna Frankonia	157
Szwabia	118
Średnia	135
Powyżej 90% katolików	
Górny Palatynat	64
Górna Bawaria	114
Dolna Bawaria	19
Średnia	75
Prowincje Prus (1883–1890)	
Powyżej 90% protestantów	
Saksonia	309,4
Szlezwig	312,9
Pomorze	171,5
Średnia	264,6
Od 68 do 89% protestantów	
Hanower	212,3
Hesja	200,3
Brandenburgia i Berlin	296,3
Prusy Wschodnie	171,3
Średnia	220,0
Od 40 do 50% protestantów	
Prusy Zachodnie	123,9
Śląsk	260,2
Westfalia	107,5
Średnia	163,6
Od 28 do 32% protestantów	
Poznań	96,4
Nadrenia	100,3
Hohenzollern	90,1
Średnia	95,6

* Uwaga: Ludność poniżej 15 roku życia została pominięta.

Źródło: Zaadaptowane z: Émile Durkheim, *Suicide*, Free Press, Glencoe, Ill. [1897] 1951, s. 153.

Tabela 11.2 pokazuje, co stwierdził. Jak widać, zarówno w Bawarii, jak i w Prusach prowincje z najwyższym udziałem protestantów miały jednocześnie najwyższe wskaźniki samobójstw. Stopniowo Durkheim nabrał pewności, że religia grała znaczącą rolę w sprawie samobójstw.

Powracając w końcu na ogólniejszy poziom teoretyczny Durkheim powiązał odkrycia dotyczące religii z wcześniejszymi obserwacjami na temat wzrastającej stopy samobójstw w okresach niepokojów politycznych. Mówiąc najprościej, Durkheim sugerował, że wiele samobójstw jest produktem *anomii*, „braku norm" lub ogólnego poczucia niestabilności społecznej i dezintegracji. Podczas okresów konfliktów politycznych ludzie mogą czuć, że dawne zwyczaje panujące w społeczeństwie upadają. Stają się zniechęceni i poddają się depresji, a samobójstwo jest jedną z ich odpowiedzi na dotkliwy dyskomfort. Spoglądając z innej strony, integracja społeczna i solidarność – odzwierciedlająca się w osobistym poczuciu, że jest się częścią spójnej, trwałej całości społecznej – dają ochronę przed depresją i samobójstwem. Oto miejsce, gdzie religie wytwarzają różnicę. Katolicyzm, jako znacznie bardziej ustrukturyzowany i zintegrowany system religijny, dawał ludziom większe poczucie spójności i stabilności, aniżeli czynił to luźniej ustrukturyzowany protestantyzm.

Z tych teorii Durkheim stworzył pojęcie *samobójstwa anomijnego*. Co bardziej istotne, dodał on, jak wiesz, pojęcie *anomii* do słownika nauk społecznych.

Ten opis klasycznych badań Durkheima jest oczywiście bardzo uproszczony. Każdy, kto uczy się badań społecznych, dobrze wyszedłby na przestudiowaniu oryginału. Jednak dla naszych celów podejście Durkheima stanowi dobry przykład możliwości badań, tkwiącej w ogromnej ilości danych, gromadzonych i opisywanych regularnie przez agencje rządowe i inne organizacje.

Jednostki analizy

Często w przypadku analizy istniejących danych statystycznych jednostką analizy nie jest pojedyncza osoba. Na przykład Durkheim musiał pracować z jednostkami polityczno-geograficznymi: krajami, regionami, państwami i miastami. Taka sama sytuacja pojawiłaby się prawdopodobnie, gdybyś miał zbadać wskaźniki przestępczości, wypadków czy chorób. Ze względu na swój charakter większość istniejących danych statystycznych jest zagregowana: opisują one grupy.

Zagregowany charakter istniejących danych statystycznych może stanowić pewien problem, mimo iż nie jest to problem nie do pokonania. Na przykład, jak widzieliśmy, Durkheim zamierzał określić, czy protestanci, czy katolicy byli bardziej skłonni do popełniania samobójstw. Trudność stanowił fakt, że żaden z dostępnych Durkheimowi rejestrów nie wskazywał na wyznanie ludzi, którzy popełnili samobójstwo. Ostatecznie więc nie mógł on powiedzieć, czy samobójstwa popełniali częściej protestanci niż katolicy, jednak tak wnioskował. Ponieważ protestanckie regiony i państwa miały wyższy poziom samobójstw niż katolickie, wniosek nasunął się sam.

W konkluzjach takiego typu tkwi jednak pewne niebezpieczeństwo. Jest możliwe, że wzorce zachowań na poziomie grupowym nie odzwierciedlają podobnych wzorców na poziomie indywidualnym. Ten rodzaj pomyłek określany jest mianem błędu ekologicznego w rozumowaniu, który został omówiony w rozdziale 4. Na przykład w przypadku badań Durkheima mogłoby być tak, że to katolicy popełniali samobójstwa na obszarach zdominowanych przez protestantów. Być może katolicy na obszarach zdominowanych przez protestantów byli tak bardzo prześladowani, że prowadziło ich to do rozpaczy i samobójstw. W takiej sytuacji byłoby możliwe, że kraje protestanckie miałyby wyższe wskaźniki samobójstw, lecz to nie protestanci popełnialiby samobójstwa.

Durkheim uniknął niebezpieczeństwa błędu ekologicznego na dwa sposoby. Po pierwsze, jego ogólne wnioski opierały się w takim samym stopniu na ścisłej dedukcji teoretycznej, co na faktach empirycznych. Odpowiedniość między teorią a faktami zmniejszała prawdopodobieństwo kontrwyjaśnień, takich jak powyższe. Po drugie, przez zakrojone na szeroką skalę ponowne testowanie swoich

konkluzji na różne sposoby, Durkheim wzmacniał prawdopodobieństwo, że są one poprawne. Wskaźniki samobójstw były wyższe w krajach protestanckich niż w katolickich, wyższe w protestanckich regionach krajów katolickich niż w katolickich regionach krajów protestanckich itd. Powtarzalność wyników zwiększała wagę materiału empirycznego, wspierającego jego konkluzje.

Problem trafności

Zawsze, gdy opieramy badania na danych, które już istnieją, jesteśmy oczywiście ograniczeni tym, co istnieje. Istniejące dane często nie obejmują dokładnie tego, czym jesteśmy zainteresowani, a nasze pomiary mogą nie być w pełni trafną reprezentacją zmiennych i pojęć, co do których chcemy wyciągać wnioski.

Dwie cechy nauki są używane przy zajmowaniu się problemem trafności w analizie istniejących danych statystycznych: *logiczne wnioskowanie* i *powtarzalność*. Strategia Durkheima stanowi przykład logicznego wnioskowania. Chociaż nie mógł on określić wyznania ludzi, którzy popełniali samobójstwa, wnioskował, że większość samobójców w regionie zdominowanym przez protestantów powinna być protestantami.

Powtarzalność może być ogólnym rozwiązaniem problemu trafności w badaniach społecznych. Przywołajmy wcześniejsze omówienie wzajemnej wymienialności wskaźników (rozdział 5). Płacz na smutnych filmach niekoniecznie jest trafną miarą współczucia, podobnie jak odkładanie piskląt do gniazd lub przekazywanie pieniędzy na cele charytatywne. Żadna z tych rzeczy brana samodzielnie nie świadczyłaby o tym, że jedną grupę (powiedzmy kobiety) cechuje więcej współczucia niż drugą (mężczyźni). Lecz gdyby kobiety okazywały się bardziej współczujące niż mężczyźni we wszystkich z tych miar, mogłoby to zwiększyć ciężar dowodów wspierających ten wniosek. W przypadku analizy istniejących danych statystycznych nieco pomysłowości i wnioskowania pozwala zwykle znaleźć kilka niezależnych testów dla danej hipotezy. Jeśli, jak się wydaje, wszystkie testy potwierdzają hipotezę, wówczas ciężar dowodów empirycznych wzmacnia trafność pomiaru.

Problem rzetelności

Analiza istniejących danych statystycznych zależy w dużym stopniu od jakości samych danych: czy dokładnie opisują to, co mają opisywać? Może to być czasami istotny problem, ponieważ na przykład ważne oficjalne tabele z danymi statystycznymi są czasami rażąco niedokładne.

Rozważmy badania dotyczące przestępczości. Ponieważ w takich badaniach bardzo wiele zależy od oficjalnych statystyk przestępczości, ten zestaw danych został poddany krytycznej ocenie. Rezultaty nie były zbyt zachęcające. Aby to zilustrować, załóżmy, że byłbyś zainteresowany śledzeniem długoterminowych trendów w używaniu marihuany w Stanach Zjednoczonych. Oficjalne statystyki, dotyczące liczby aresztowanych za sprzedawanie lub posiadanie marihuany, są, jak się wydaje, sensownym miernikiem jej używania, prawda? Niekoniecznie.

Na początku stajesz przed istotnym problemem trafności. Przed ustanowieniem Marijuana Tax Act w 1937 roku „trawa" była w Stanach Zjednoczonych legalna, więc rejestry aresztowanych nie będą trafną miarą jej używania. Lecz nawet jeśli ograniczyłbyś swoje badania do okresu po 1937 roku, nadal miałbyś problemy z rzetelnością, których źródłem jest sama natura egzekwowania prawa i prowadzenia rejestrów przestępstw.

Na przykład egzekwowanie prawa jest przedmiotem rozmaitych nacisków. Okrzyki publicznego oburzenia przeciw marihuanie, zapewne pod wodzą głośnych grup obywateli, w rezultacie często przynoszą policyjną rozprawę z „nielegalnym handlem narkotykami" – szczególnie podczas roku wyborczego lub budżetowego. Sensacyjne historie publikowane w prasie mogą mieć podobny efekt. W dodatku rozmiar innych spraw, jakie stoją przed policją, może wpływać na aresztowania związane z marihuaną.

Śledząc wzorzec aresztowań za narkotyki w Chicago między 1942 a 1970 rokiem, Lois De-

Fleur (1975) pokazał, że oficjalne rejestry znacznie lepiej ukazują historię praktyk policji i wywieranych na nią nacisków politycznych niż historię używania narkotyków. Na innym poziomie analizy Donald Black ze współpracownikami (1970) analizowali czynniki wpływające na podejmowanie przez policję decyzji o aresztowaniu winowajcy bądź jego warunkowym wypuszczeniu. W końcu, na oficjalne statystyki przestępstw wpływa to, czy konkretni winowajcy są dobrze, czy biednie ubrani, czy są grzeczni, czy też obrażają oficerów policji itd. Biorąc pod uwagę przestępstwa nie odnotowane, czasami szacowane jako aż dziesięciokrotność liczby przestępstw znanych policji, rzetelność statystyk kryminalnych stanie się jeszcze bardziej wątpliwa.

Powyższe uwagi odnoszą się do statystyk przestępstw na poziomie lokalnym. Często użyteczne jest analizowanie krajowych statystyk kryminalnych, takich jak prezentowane w corocznym Jednolitym Raporcie Kryminalnym (*Uniform Crime Reports*) FBI. Na poziomie krajowym pojawiają się dodatkowe problemy. Różne sądy lokalne różnie definiują przestępstwa, a uczestnictwo w programach FBI jest dobrowolne, tak więc dane są niekompletne.

Również proces przechowywania rejestrów wpływa na dane dostępne badaczom. Za każdym razem, gdy udoskonalany jest system przechowywana rejestrów, na przykład przez komputeryzację, wskaźniki przestępstw dramatycznie rosną. Może się to zdarzać nawet

Czy Ameryka jest numerem 1?

19 września 1999 roku stacja telewizyjna ABC nadała specjalny program, prowadzony przez Johna Stossela, który miał na celu zbadanie, na którym miejscu w rankingu krajów świata znajdują się Stany Zjednoczone. W miarę trwania programu stawało się jasne, że USA mają się dobrze – są prawdopodobnie na pierwszym miejscu – oraz że kluczem do naszego sukcesu jest przede wszystkim leseferystyczny system kapitalistyczny. Aby wzmocnić wymowę tego ostatniego punktu, Stossel wskazał na inne historie sukcesu, w których sukces także wynikał z leseferystycznego kapitalizmu.

Według Stossela Hongkong wyróżniał się wśród narodów świata jako lider w gospodarce wolnorynkowej. Jako dowód sukcesu Hongkongu Stossel doniósł, iż ma on „jedyny rząd na świecie, który wypracowuje nadwyżkę, dużą nadwyżkę". Co myślisz o tej konkluzji? Czy jest ona dla ciebie przekonująca?

Oto co miała do powiedzenia organizacja nadzorująca media Fairness and Accuracy in Reporting (FAIR) na temat twierdzenia Stossela:

> „Jak wie każdy, kto zwraca uwagę na politykę Waszyngtonu, budżet rządu Stanów Zjednoczonych od ponad roku wykazuje nadwyżkę; w ubiegłym roku wyniosła ona 70 mld. dolarów. Poza Stanami Zjednoczonymi na liście krajów z nadwyżką budżetową w ubiegłym roku znajdowały się: Wielka Brytania, Kanada, Australia, Dania, Finlandia, Islandia, Irlandia, Nowa Zelandia, Norwegia i Szwecja".

Dalej Stossel porównał Hongkong (historię kapitalistycznego sukcesu) z alternatywą wobec wolnej gospodarki: „stagnacja i często ubóstwo. Rozważmy Chiny, obecnie grzęznące w ubóstwie typowym dla Trzeciego Świata. Były one kiedyś światowym liderem". I znowu FAIR sugeruje inną ocenę:

> „W rzeczywistości o gospodarce Chin można powiedzieć wszystko poza tym, że jest w «stagnacji». Jak oznajmił w ubiegłorocznym przemówieniu Lawrence Summers z Departamentu Skarbu, «od momentu zapoczątkowania tam reform [ekonomicznych] w 1980 roku, Chiny są najszybciej rozwijającą się gospodarką w historii». Chociaż Chiny przejęły pewne elementy gospodarki rynkowej, duża część ich przedsiębiorstw jest ciągle w posiadaniu rządu".

W mediach i gdzie indziej często można napotkać stwierdzenia dotyczące faktów, które, jak się wydaje, opierają się na analizach statystycznych. Jednak zawsze dobrze jest sprawdzić fakty.

Źródło: Fairness and Accuracy in Reporting, „Action Alert: ABC News Gives up on Accuracy?" 28 września 1999, dostępne na: http://www.fair.org/activism/stossel-america.html/.

wówczas, gdy liczba przestępstw popełnionych, zgłoszonych i tych, gdzie przeprowadzono dochodzenie, nie wzrasta.

Pierwszym ratunkiem badaczy przed problemami, związanymi z rzetelnością w analizie istniejących danych statystycznych, jest sama świadomość, że taki problem może wystąpić. Dociekanie sposobu zbierania i zestawiania danych może umożliwić oszacowanie natury i stopnia nierzetelności tak, że będzie można ocenić jej potencjalny wpływ na twoje zamiary badawcze. Stosując także wnioskowanie logiczne i powtarzanie, zazwyczaj można poradzić sobie z tym problemem. Ramka „Czy Ameryka jest numerem 1?" daje przykład tego, co można odkryć przez staranne badanie istniejących statystyk.

Źródła istniejących danych statystycznych

Można stworzyć całą książkę, jedynie zestawiając źródła danych dostępnych do analizy. W tym podrozdziale chcę wymienić kilka z nich i wskazać ci kierunek, w którym powinieneś pójść, aby znaleźć dalsze, związane z twoimi zainteresowaniami badawczymi.

Niewątpliwie najbardziej wartościową książką, jaką możesz kupić, jest coroczny *Statistical Abstract of the United States*, wydawany przez Departament Handlu USA[3]. Jest to bezsprzecznie najlepsze źródło danych na temat Stanów Zjednoczonych, zawierające statystki poszczególnych stanów i (mniej obszernie) miast, jak również informacje o kraju jako całości. Gdzie jeszcze możesz znaleźć liczbę strajkujących zakładów w kolejnych latach, wartość podatków od nieruchomości w ważniejszych miastach, ilość ścieków w kraju, liczbę właścicieli przedsiębiorstw na terenie kraju oraz setki innych tak przydatnych informacji? By jeszcze bardziej uprościć życie, Hoover's Business Press oferuje tę samą książkę po niższej cenie w miękkiej okładce. Nie należy mylić tej komercyjnej wersji, zatytułowanej *The American Almanac*, z innymi almanachami, które są mniej rzetelne i mniej użyteczne w społecznych badaniach naukowych. Jeszcze lepiej jest kupić *Statistical Abstract* na płycie CD-ROM, co ułatwi poszukiwanie i przenoszenie danych.

Agencje federalne – departamenty pracy, rolnictwa, transportu itd. – publikują niezliczone serie danych. By dowiedzieć się, co jest dostępne, pójdź do biblioteki, poszukaj działu z dokumentami rządowymi i poświęć kilka godzin na szperanie wśród półek. Możesz też odwiedzić stronę www U.S. Governement Printing Office (http://www.access.gpo.gov) i porozglądać się na niej. Sieć jest najnowszym ułatwieniem w dostępie do istniejących danych statystycznych. Oto kilka stron w sieci, ilustrujących bogactwo tego nowego zasobu:

- Bureau of the Census (Biuro Spisowe) http://www.census.gov/
- Bureau of Labor Statistics (Biuro Statystyk Pracy) http://stats.bls.gov/blshome.html
- Department of Education (Departament Oświaty) http://www.ed.gov/
- Bureau of Transportation Statistics (Biuro Statystyk Transportu) http://www.bts.gov/
- Federal Bureau of Investigation (Federalne Biuro Śledcze) http://www.fbi.gov/
- Central Intelligence Agency (Centralna Agencja Wywiadowcza) http://www.cia.gov/
- State and Local Governments (władze stanowe i lokalne) http://www.piperinofo.com/state/index.cfm
- The World Bank (Bank Światowy) http://www.worldbank.org/

Statystyki światowe są dostępne za pośrednictwem Organizacji Narodów Zjednoczonych. Rocznik Demograficzny ONZ przedstawia najważniejsze dane statystyczne w ujęciu rocznym (urodzenia, zgony i inne dane dotyczące ludności) dla poszczególnych krajów świata. Inne publikacje przedstawiają rozmaite inne typy danych. Znowu wycieczka do biblioteki, połączona z poszukiwaniami w sieci, najlepiej uświadomi ci, co jest dostępne.

Ilość danych dostarczanych przez agencje pozarządowe jest równie zdumiewająca, jak tych, które finansujesz z podatków. Izby handlowe często publikują dane dotyczące biznesu,

[3] Jego polskim odpowiednikiem jest Rocznik Statystyczny, wydawany przez Główny Urząd Statystyczny (przyp. tłum.).

Cierpienie na świecie

W 1992 roku Komitet Kryzysu Ludnościowego (*Population Crisis Committee*), organizacja non-profit zaangażowana w walkę z eksplozją demograficzną, przeprowadziła analizę względnego stopnia cierpienia społeczeństw świata. Każdy kraj o liczbie ludności nie mniejszej niż jeden milion został oceniony pod względem następujących dziesięciu wskaźników – wynik 10 dla każdego z tych wskaźników oznaczał, że w danym kraju występują pod tym względem największe trudności:

oczekiwana długość życia;
dzienna racja kalorii na osobę;
procent ludności z dostępem do czystej wody pitnej;
proporcja szczepionych niemowląt;
wskaźnik uczęszczania do szkoły średniej;
produkt narodowy brutto;
inflacja;
liczba telefonów na 1000 osób;
wolność polityczna;
prawa obywatelskie.

Oto lista krajów świata, uszeregowanych ze względu na te wskaźniki. Pamiętaj, że wysoki wynik jest znakiem największego ogólnego cierpienia.

Skrajne cierpienie ludzi

93 Mozambik
92 Somalia
89 Afganistan, Haiti, Sudan
88 Zair
87 Laos
86 Gwinea, Angola
85 Etiopia, Uganda
84 Kambodża, Sierra Leone
82 Czad, Gwinea, Bissau
81 Ghana, Birma
79 Malawi
77 Kamerun, Mauretania
76 Rwanda, Wietnam, Liberia
75 Burundi, Kenia, Madagaskar, Jemen

Znaczne cierpienie ludzi

74 Wybrzeże Kości Słoniowej
73 Bhutan, Burkina Faso, Republika Środkowej Afryki
71 Tanzania, Togo
70 Lesotho, Mali, Niger, Nigeria
69 Gwatemala, Nepal
68 Bangladesz, Boliwia, Zambia
67 Pakistan
66 Nikaragua, Papua Nowa Gwinea, Senegal, Swaziland, Zimbabwe
65 Irak
64 Gambia, Kongo, Salwador, Indonezja, Syria
63 Komory, Indie, Paragwaj, Peru
62 Benin, Honduras
61 Liban, Chiny, Gujana, Republika Południowej Afryki
59 Egipt, Maroko
58 Ekwador, Sri Lanka
57 Botswana
56 Iran
55 Surinam
54 Algieria, Tajlandia
53 Republika Dominikany, Meksyk, Tunezja, Turcja
51 Libia, Kolumbia, Wenezuela
50 Brazylia, Oman, Filipiny

Umiarkowane cierpienie ludzi

49 Wyspy Salomona
47 Albania
45 Vanuatu
44 Jamajka, Rumunia, Arabia Saudyjska, Seszele, (była) Jugosławia
43 Mongolia
41 Jordania
40 Malezja, Mauritius
39 Argentyna
38 Kuba, Panama
37 Chile, Urugwaj, Korea Północna
34 Kostaryka, Korea Południowa, Zjednoczone Emiraty Arabskie
33 Polska
32 Bułgaria, Węgry, Katar
31 (dawny) Związek Radziecki
29 Bahrain, Hongkong, Trynidad i Tobago
28 Kuwejt, Singapur
25 Czechosłowacja, Portugalia, Tajwan

Minimalne cierpienie ludzi

21 Izrael
19 Grecja
16 Wielka Brytania
12 Włochy
11 Barbados, Irlandia, Hiszpania, Szwecja
8 Finlandia, Nowa Zelandia
7 Francja, Islandia, Japonia, Luksemburg
6 Austria, Niemcy
5 Stany Zjednoczone
4 Australia, Norwegia
3 Kanada, Szwajcaria
2 Belgia, Holandia
1 Dania

podobnie robią prywatne grupy konsumentów. Ralph Nader[4] dysponuje informacjami o bezpieczeństwie drogowym, a Common Cause[5] omawia politykę i sprawy rządowe. George Gallup publikuje bibliografię badań opinii publicznej, prowadzonych za pomocą sondaży Gallupa od 1935 roku.

Organizacje takie jak Biuro do Spraw Ludności (*Population Reference Bureau*) publikują szereg danych demograficznych dotyczących Stanów Zjednoczonych i świata, których mógłby użyć badacz zajmujący się analizami wtórnymi. Wydawany przez Biuro *World Population Data Sheet* (Zbiór Danych na Temat Ludności Świata) i *Population Bulletin* (Biuletyn Ludności) są zasobami intensywnie wykorzystywanymi przez badaczy społecznych. Dane o wskaźnikach społecznych można znaleźć w czasopiśmie *SINET: A Quarterly Review of Social Reports and Research on Social Indicators, Social Trends, and the Quality of Life* (Kwartalny Przegląd Raportów Społecznych i Badań Wskaźników Społecznych, Trendów Społecznych i Jakości Życia).

Zasoby, które wymieniłem, są jedynie małą częścią z tysięcy dostępnych. Przy tak wielkiej ilości już zgromadzonych danych nie można zasłaniać się brakiem funduszy na kosztowne zbieranie danych jako przyczyną rezygnacji z prowadzenia dobrych i użytecznych badań społecznych.

Dostępność istniejących statystyk umożliwia również tworzenie pewnych dość wyrafinowanych miar. Dołączona ramka „Cierpienie na świecie" opisuje analizę opublikowaną przez Komitet Kryzysu Ludnościowego (*Population Crisis Committee*), sporządzoną na podstawie danych dostępnych w praktyce rządowej.

[4] Ralph Nader jest amerykańskim działaczem konsumenckim, prawnikiem, autorem wielu publikacji. W latach 60. był także wykładowcą uniwersyteckim. W 2000 roku startował w wyborach prezydenckich jako kandydat Partii Zielonych. Więcej informacji o Naderze można znaleźć na jego stronie http://www.nader.org/ (przyp. tłum.).

[5] Common Cause to istniejąca od 1970 roku w USA pozarządowa i pozapartyjna organizacja zajmująca się lobbingiem obywatelskim (patrz www.commoncause.org) (przyp. red. nauk.).

Analizy historyczno-porównawcze

Badania historyczno-porównawcze różnią się zasadniczo od omawianych dotąd metod, chociaż w pewnym stopniu nakładają się na badania terenowe, analizę treści i analizę istniejących danych statystycznych. Wymagają one jednak użycia metod historycznych przez socjologów, politologów, a także innych przedstawicieli nauk społecznych.

Pomimo że w rozdziale 4 omawialiśmy schematy badań dynamicznych, jak dotąd nasz przegląd metod badawczych skupiał się przede wszystkim na badaniach zakotwiczonych w jednym punkcie w czasie i w jednej społeczności, czy była nią mała grupa, czy naród. Mimo trafnego przedstawienia głównego nurtu współczesnych badań z zakresu nauk społecznych, to skoncentrowanie pomija fakt, że badacze społeczni są też zainteresowani śledzeniem rozwoju form społecznych w czasie oraz międzykulturowymi porównaniami procesów rozwojowych. Po opisaniu kilku najważniejszych przykładów badań historycznych i porównawczych w przeszłości, w podrozdziale tym zostaną też omówione kluczowe elementy tej metody.

Przykłady analiz historyczno-porównawczych

Twórca terminu *sociologie* August Comte postrzegał tę nową dyscyplinę jako ostatni etap w historycznym rozwoju idei. W swoim najogólniejszym ujęciu odmalował on ewolucyjny obraz przejścia ludzkości od oparcia się na religii, przez metafizykę, ku nauce. W bardziej szczegółowym ujęciu portretował on naukę jako ewoluującą od rozwoju biologii i innych nauk przyrodniczych, przez rozwój psychologii, aż ku rozwojowi naukowej socjologii.

Bardzo wielu późniejszych przedstawicieli nauk społecznych zwracało się również w stronę szerokich procesów historycznych. Kilku badało historyczny rozwój form społecznych, od prostych do złożonych, od społeczeństw wiejsko-rolniczych do miejsko-przemysłowych. Na przykład amerykański antropolog Lewis Morgan postrzegał rozwój jako postępujący od „dzikości", przez „barbarzyństwo", ku „cywilizacji"

(1870). Robert Redfield, kolejny antropolog, bliższy współczesności, pisał o przejściu od „społeczeństwa ludowego" do „społeczeństwa miejskiego" (1941). Émile Durkheim widział ewolucję społeczną w dużej mierze jako proces coraz większego podziału pracy ([1893] 1964). W bardziej szczegółowej analizie Karol Marks badał historyczny rozwój systemów ekonomicznych od form pierwotnych, poprzez feudalne, aż do kapitalistycznych ([1867] 1967). W tym kontekście cała historia to, jak pisał, historia walki klas – „posiadającej", walczącej o utrzymanie swoich korzyści i „nieposiadającej", walczącej o poprawę losu. Wybiegając spojrzeniem poza kres kapitalizmu, Marks dostrzegał rozwój socjalizmu, a w końcu komunizmu.

Jednak nie wszystkie badania historyczne w naukach społecznych miały to ewolucyjne zabarwienie. Właściwie niektóre interpretacje danych historycznych, dokonywane przez nauki społeczne, wskazują raczej na istnienie wielkich cyklów rozwojowych niż na postęp liniowy. Żaden uczony nie reprezentuje lepiej tego punktu widzenia niż Pitirim A. Sorokin. Sorokin, uczestnik rewolucji październikowej, służył jako sekretarz premiera Kiereńskiego. Jednak zarówno Kiereński, jak i Sorokin wypadli z łask, a Sorokin rozpoczął swoją drugą karierę – jako amerykański socjolog.

Comte odczytywał historię jako rozwój od religii do nauki, natomiast Sorokin (1937–1940) sugerował, że społeczeństwa wahają się cyklicznie między dwoma punktami widzenia, które nazwał „ideacyjnym" i „zmysłowym". Zmysłowy punkt widzenia definiuje rzeczywistość w terminach doświadczeń zmysłowych, ideacyjny przeciwnie, większy nacisk kładzie na czynniki duchowe i religijne. Odczytywanie zapisów historycznych przez Sorokina wskazało następnie, że przejście między punktami ideacyjnym a zmysłowym odbywało się przez trzeci punkt widzenia, który nazywał on „idealistycznym". Ten trzeci sposób postrzegania łączył elementy zmysłowego i ideacyjnego w zintegrowany, racjonalny pogląd na świat.

Te przykłady wskazują na pewne tematy, jakie rozpatrują badacze zajmujący się analizami historyczno-porównawczymi. By lepiej zrozumieć, z czym wiążą się badania historyczno-porównawcze, spójrzmy nieco bardziej szczegółowo na kilka przykładów.

Weber i rola idei

W swoich analizach historii gospodarczej Karol Marks przedstawił punkt widzenia determinizmu ekonomicznego. Zgodnie z tym, co postulował, czynniki ekonomiczne określają charakter wszystkich innych aspektów społeczeństwa. Analizy Marksa wykazały na przykład, że funkcją kościołów europejskich było uprawomocnianie i wspieranie kapitalistycznego *status quo* – religia była narzędziem, którym posługiwali się możni w celu utrzymywania dominacji nad słabymi. „Religia jest westchnieniem uciśnionego stworzenia", napisał Marks w słynnym ustępie, „sercem nieczułego świata jak jest duszą bezdusznych stosunków. Religia jest opium ludu" (*Karl Marks: Selected Writings* 1956, s. 27)[6].

Max Weber, niemiecki socjolog, miał inny pogląd. Nie przecząc temu, że czynniki ekonomiczne mogły wpływać i wpływały na inne aspekty społeczeństwa, Weber argumentował, że determinizm ekonomiczny nie wyjaśnia wszystkiego. Tak naprawdę, powiada Weber, formy ekonomiczne mogły wywodzić się z idei pozaekonomicznych. W swoich badaniach z socjologii religii Weber badał zasięg, w jakim instytucje religijne były raczej źródłem zachowań społecznych niż zwykłym odbiciem warunków ekonomicznych. Jego najbardziej znane stwierdzenie tej strony zagadnienia jest zawarte w pracy *Etyka protestancka a duch kapitalizmu* z 1905 roku[7]. Oto krótki przegląd tez Webera.

Jan Kalwin (1509–1564), francuski teolog, był ważną postacią w protestanckiej reformacji chrześcijaństwa. Kalwin nauczał, że ostateczne zbawienie lub potępienie każdej jednostki zostało wcześniej postanowione przez Boga; ten

[6] Cytat pochodzi z *Przyczynku do krytyki heglowskiej filozofii prawa* Karola Marksa, 1844, wyd. pol. Karol Marks, Fryderyk Engels, *Dzieła*, t. 1, s. 458, tłum. Leszek Kołakowski, Warszawa 1976 (1 wyd. polskie K. Marks, F. Engels *Wybrane pisma filozoficzne 1844–1846*, KiW, 1949) (przyp. tłum.).

[7] Wydanie polskie: *Etyka protestancka a duch kapitalizmu*, tłum. Jan Miziński, Lublin 1994 (przyp. tłum.).

pogląd jest nazywany *predestynacją*. Kalwin sugerował również, że Bóg przekazuje swoją decyzję ludziom, obdarzając ich powodzeniem lub niepowodzeniem podczas ich ziemskiej egzystencji. Bóg daje każdej osobie ziemskie „powołanie" – zajęcie lub zawód – i objawia jej sukces lub porażkę za jego pośrednictwem. Jak na ironię, ten punkt widzenia skłonił zwolenników Kalwina do poszukiwania dowodu ich nadchodzącego zbawienia przez ciężką pracę, oszczędzanie pieniędzy i ogólne dążenie do sukcesu ekonomicznego.

Zgodnie z analizami Webera kalwinizm stanowił ważny bodziec dla rozwoju kapitalizmu. Zamiast „marnować" swoje pieniądze na doczesne wygody kalwiniści reinwestowali je w swoje przedsięwzięcia ekonomiczne, dostarczając w ten sposób kapitału niezbędnego do rozwoju kapitalizmu. W dochodzeniu do tej interpretacji początków kapitalizmu, Weber zbadał oficjalną doktrynę wczesnych kościołów protestanckich, przestudiował nauczanie Kalwina i innych przywódców kościoła i zbadał inne powiązane dokumenty historyczne.

W trzech innych pracach Weber przeprowadził szczegółowe analizy historyczne judaizmu ([1934] 1952) oraz religii Chin ([1934] 1951) i Indii ([1934] 1958). Weber chciał się między innymi dowiedzieć, dlaczego kapitalizm nie rozwinął się w dawnych społeczeństwach Chin, Indii i Izraela. W żadnej z tych trzech religii nie znalazł on nauczania, które wspierałoby akumulację i reinwestowanie kapitału – co wzmocniło jego wnioski na temat roli protestantyzmu w tym zakresie.

Religia japońska i kapitalizm

Teza Webera dotycząca protestantyzmu i kapitalizmu stała się klasyką nauk społecznych. Nic więc dziwnego, że inni badacze usiłowali testować ją w innych warunkach historycznych. Jednak żadna z tych analiz nie była bardziej interesująca niż badania Roberta Bellaha poświęcone rozwojowi kapitalizmu w Japonii w końcu XIX i w początkach XX wieku, *Tokugawa Religion* (1957).

W czasie studiów Bellah zainteresował się Weberem i społeczeństwem japońskim. Uwzględniając oba te obszary zainteresowań, trudno się dziwić, że w 1951 roku planował jako temat swojej rozprawy doktorskiej „nie mniej niż «Studium etyki gospodarczej w Japonii», aby dotrzymać kroku studiom Webera o Chinach, Indiach i judaizmie: *The Economic Ethic of the World Religions* (*Etyka gospodarcza religii świata*)" (przywołane w Bellah 1967, s. 168). Pierwotnie Bellah naszkicował projekt swoich badań w następujący sposób:

> „Problemy muszą być konkretne i ograniczone – bez próby pisania historii ogólnej – ponieważ rozpiętość czasowa wynosi kilka stuleci. Praca terenowa w Japonii na temat faktycznej etyki gospodarczej praktykowanej przez ludzi w różnych sytuacjach z, jeśli to możliwe, dopasowanymi próbami kontrolnymi ze Stanów Zjednoczonych (kwestionariusze, wywiady itp.)".
>
> (1967, s. 168)

Oryginalny plan Bellaha wymagał zatem badań sondażowych wśród współczesnych Japończyków i Amerykanów. Nie uzyskał on jednak wsparcia finansowego, koniecznego dla realizacji badań zgodnie z pierwotnym wyobrażeniem. Zamiast tego więc zagłębił się w zapisy historyczne dotyczące japońskiej religii, poszukując korzeni rozwoju kapitalizmu w tym kraju.

W ciągu kilku lat poszukiwań Bellah odkrył liczne tropy. W referacie semestralnym na ten temat z 1952 roku Bellah miał wrażenie, że znalazł odpowiedź w samurajskim kodeksie Bushido i w konfucjanizmie, praktykowanym przez klasę samurajów:

> „Znajdujemy tutaj, jak sądzę, prawdziwy rozwój owego świeckiego ascetyzmu, co najmniej równy temu odnalezionemu w Europie. Idąc dalej, dla tej klasy idea obowiązku w pracy wymagała osiągnięć bez ograniczeń, pochodzących z tradycji, lecz ograniczonych możliwościami człowieka, czy to w roli biurokraty, czy lekarza, nauczyciela bądź naukowca lub w innej roli otwartej dla samuraja".
>
> (cyt. w: Bellah 1967, s. 171)

Samurajowie stanowili jednak tylko część społeczeństwa japońskiego, więc Bellah stale

rozglądał się wśród ogółu japońskich religii. Jego znajomość japońskiego nie była jeszcze wówczas bardzo dobra, zamierzał jednak czytać teksty religijne w oryginale. W tej sytuacji oraz pod narastającą presją czasu Bellah zdecydował się skupić uwagę na pojedynczej grupie: Shingaku, osiemnasto- i dziewiętnastowiecznym kupieckim ruchu religijnym. Zauważył, że Shingaku miał podwójny wpływ na rozwój kapitalizmu. Proponował on postawę wobec pracy podobną do kalwińskiego pojęcia „powołania", a przez to sprawiał, że działalność gospodarcza stała się bardziej akceptowalnym powołaniem dla Japończyków. Początkowo bowiem handel miał w Japonii bardzo niską pozycję.

W innych fragmentach swoich analiz Bellah badał religijną i polityczną rolę cesarza oraz ekonomiczny wpływ pojawiającego się okresowo kultu cesarza. Ostatecznie badania Bellaha wskazały na wiele różnych czynników religijnych i filozoficznych, które położyły fundamenty pod rozwój kapitalizmu w Japonii. Wydaje się mało prawdopodobne, aby osiągnął on choć w przybliżeniu taką głębię rozumienia, gdyby mógł realizować swój pierwotny plan przeprowadzania wywiadów na porównywalnych próbach obywateli Stanów Zjednoczonych i Japonii.

Te przykłady badań historyczno-porównawczych powinny dać ci pewne wyobrażenie o możliwościach tej metody. Przejdźmy teraz do przeglądu źródeł i technik stosowanych w tej metodzie.

Źródła danych historyczno-porównawczych

Jak już widzieliśmy, omawiając istniejące dane statystyczne, niewyczerpane są dane nadające się do analiz historyczno-porównawczych. Historycy mogli już wcześniej opisać to, co zamierzasz badać, ich analizy zaś mogą być wprowadzeniem do tematu, punktem wyjścia dla głębszych poszukiwań.

Najprawdopodobniej ostatecznie zechcesz wyjść poza konkluzje innych i analizować pewne „dane surowe", by móc wyciągać własne wnioski. Takie dane różnią się, oczywiście, w zależności od przedmiotu badań. W studium religii Tokugawa dokonanym przez Bellaha na dane surowe składały się kazania nauczycieli Shingaku. Gdy William I. Thomas i Florian Znaniecki (1918) badali proces przystosowywania się polskich chłopów, przybywających do Stanów Zjednoczonych w początkach XX wieku, badali listy pisane przez emigrantów do ich rodzin w Polsce. (Uzyskiwali listy za pośrednictwem ogłoszeń w prasie). Inni badacze analizowali stare pamiętniki. Jednak tego rodzaju źródła osobiste jedynie muskają powierzchnię. Omawiając procedury studiowania historii życia rodzinnego, Ellen Rothman wskazuje na następujące źródła:

> „Oprócz źródeł osobistych istnieją ogólnodostępne zapisy, które również odsłaniają historię rodzinną. Gazety są szczególnie bogatym źródłem danych, dotyczących edukacyjnych, prawnych i rekreacyjnych aspektów życia rodzinnego w przeszłości, postrzeganego z perspektywy lokalnej. Czasopisma odzwierciedlają ogólniejsze wzorce życia rodzinnego; studenci często postrzegają je jako interesujące do badania ze względu na dane o postrzeganych i oczekiwanych głównych wartościach rodzinnych. Czasopisma oferują wiele różnych rodzajów źródeł jednocześnie: materiały wizualne (ilustracje i reklamy), komentarze (artykuły redakcyjne i kolumny z poradami) i prozę. Czasopisma popularne są szczególnie bogate w te dwa ostatnie rodzaje źródeł. Porady w wielu kwestiach, dotyczących rodziny – od właściwego sposobu dyscyplinowania dzieci po ekonomiczne używanie tapet – wypełniają kolumny czasopism od początków XIX wieku do dzisiaj. Opowieści wskazujące na wspólne doświadczenie lub postrzeganie życia rodzinnego pojawiają się równie stale".
>
> (1981, s. 53)

Organizacje na ogół prowadzą własną dokumentację, więc jeśli badasz rozwój pewnej organizacji – tak jak na przykład Bellah badał Shingaku – powinieneś przejrzeć jej oficjalne dokumenty: statuty, oświadczenia polityczne, przemówienia przywódców itd. Raz, kiedy badałem rozwój współczesnej japońskiej grupy religijnej – Sokagakkai – odkryłem nie tylko tygodniki i czasopisma, wydawane przez gru-

pę, ale także wydany zbiór wszystkich przemówień, wygłoszonych przez jej pierwszych przywódców. Za pomocą tych źródeł mogłem śledzić zmiany we wzorcach rekrutacji, dokonujące się w czasie. Na początku wyznawcy mieli nakaz wciągnięcia do grupy całego świata. Później nacisk przesunął się szczególnie na Japonię. Gdy uzyskano znaczącą liczbę członków w Japonii, powrócił nacisk na wciąganie całego świata (Babbie 1966).

Często oficjalne dokumenty rządowe dostarczają danych przydatnych do analiz. Aby móc lepiej ocenić historię relacji międzyrasowych w Stanach Zjednoczonych, Leon A. Higginbotham Jr. (1978) przebadał akty prawne i sprawy sądowe związane z rasą z około 200 lat. Sam będąc pierwszym Afroamerykaninem mianowanym na sędziego federalnego, Higginbotham zauważył, że zamiast chronić Afroamerykanów, prawo sankcjonowało fanatyzm i ucisk. W najstarszych sprawach sądowych panowała duża niejasność co do tego, czy Afroamerykanie są umówionymi służącymi czy właściwie niewolnikami. Późniejsze sprawy sądowe i prawa wyjaśniły tę sprawę – ograniczając Afroamerykanów do bycia czymś mniej niż ludźmi.

Źródła danych dla analiz historycznych są zbyt rozległe, by omówić je tu nawet w zarysie, jednak przykłady, na jakie spojrzeliśmy, powinny zasugerować pewne pomysły. Jednak bez względu na to, z jakich zasobów skorzystasz, na miejscu jest kilka przestróg.

Jak widzieliśmy, w przypadku istniejących danych statystycznych, nie można ufać dokładności zapisów oficjalnych czy nieoficjalnych, pierwotnych czy wtórnych. Twoje zabezpieczenie polega na powtarzalności: w przypadku badań historycznych oznacza to potwierdzanie. Jeśli kilka źródeł wskazuje na ten sam zestaw „faktów", twoje zaufanie do nich może zasadnie wzrosnąć.

Jednocześnie powinieneś zawsze być ostrożny ze względu na możliwe obciążenia twoich źródeł danych. Jeśli wszystkie dane na temat rozwoju jakiegoś ruchu politycznego pochodzą od samego ruchu, jest mało prawdopodobne, byś uzyskał wszechstronny obraz. Pamiętniki zamożnej szlachty średniowiecznej mogą nie dać ci dokładnego obrazu ogólnego życia w tych czasach. Tam, gdzie to możliwe, uzyskuj dane z rozmaitych źródeł reprezentujących różne punkty widzenia. Oto wypowiedź Bellaha na temat jego analiz Shingaku:

> „Ktoś mógłby przekonywać, że to, na co zdecydowali się zwrócić uwagę zachodni naukowcy, nie będzie wolne od obciążeń. Jednak fakt, że znalazł się tu materiał pochodzący od zachodnich naukowców o różnorodnych zainteresowaniach, z wielu krajów i z okresu blisko stulecia, redukuje prawdopodobieństwo obciążeń". (1967, s. 179)

Kwestie podniesione przez Bellaha są ważne. Jak wskazują Ron Aminzade i Barbara Laslett w ramce „Czytanie i ocena dokumentów", sztuką jest wiedzieć, jak odnosić się do takich dokumentów i co z nich można uzyskać.

Nawiasem mówiąc, taki krytyczny przegląd, do jakiego Aminzade i Laslett nawołują w odniesieniu do interpretacji dokumentów historycznych, jest przydatny w wielu innych obszarach życia poza prowadzeniem badań historyczno-porównawczych. Weź pod uwagę zastosowanie niektórych z tych pytań do prezydenckich konferencji prasowych, do reklam lub (uwaga!) do podręczników akademickich. Żadne z tych źródeł nie oferuje bezpośredniego oglądu rzeczywistości; autorami wszystkich są ludzie i dotyczą one także ludzi.

Techniki analityczne

Analiza danych historyczno-porównawczych to kolejny wielki temat, którego nie mogę tutaj omówić wyczerpująco. Ponadto ze względu na to, że badania historyczno-porównawcze są zwykle metodą jakościową, nie ma prostego zestawienia kolejnych działań, jakie należy podjąć przy analizie danych historycznych. Mimo to należy uczynić kilka uwag.

Max Weber używał niemieckiego terminu *verstehen* – „rozumienie" w odniesieniu do zasadniczej cechy badań społecznych. Miał on na myśli to, że badacz musi być w stanie przyjąć rozumowo okoliczności, poglądy i uczucia osób badanych, by móc właściwie interpretować ich działania. Z pewnością pojęcie to sto-

Czytanie i ocena dokumentów

Ron Aminzade i Barbara Laslett

Uniwersytet Minnesota

Celem poniższych uwag jest zapoznanie cię z tym rodzajem pracy interpretacyjnej, jaką wykonują historycy, oraz z krytycznym podejściem, jakim się oni kierują wobec swoich źródeł. Powinno to pomóc ci docenić niektóre umiejętności, rozwijane przez historyków podczas ich starań dotyczących zrekonstruowania przeszłości z tego, co pozostawiła, oceny statusu dowodowego różnych typów dokumentów oraz określenia zakresu dopuszczalnego wnioskowania i interpretacji. Oto niektóre z pytań dotyczących dokumentów, zadawane przez historyków:

1. Kto napisał te dokumenty? Dlaczego zostały napisane? Dlaczego przetrwały przez te wszystkie lata? Jakich metod użyto, by zdobyć informacje zawarte w tych dokumentach?
2. Jakie obciążenia mogą występować w tych dokumentach i jak mógłbyś zabrać się za ich sprawdzenie lub poprawienie? Jak duża lub reprezentatywna jest próba jednostek, wydarzeń itp., uwzględnionych w tym dokumencie? Jakie były ograniczenia instytucjonalne i ogólne procedury organizacyjne, w ramach których ten dokument został przygotowany? Do jakiego stopnia dokument mówi więcej o rodzajach aktywności instytucjonalnej niż o zjawisku, będącym przedmiotem badań? Jaki czas upłynął między obserwacją dokumentowanych zjawisk a relacją, jaką świadek o nich złożył? Jak bardzo poufne lub publiczne było przeznaczenie dokumentu? Jaką rolę w przedstawieniu materiału zawartego w dokumencie grała etykieta, konwencja i zwyczaj? Jak zniekształcona mogłaby być twoja wizja przeszłości, gdybyś polegał wyłącznie na danych zawartych w tym dokumencie? Jakie inne rodzaje dokumentów mógłbyś przejrzeć w poszukiwaniu danych dotyczących tej samej kwestii?
3. Jakie są kluczowe kategorie i pojęcia zastosowane przez piszącego dokument do uporządkowania prezentowanych informacji? Jaki rodzaj wybiórczości lub przemilczeń jest rezultatem tych kategorii myślowych?
4. Na jaki rodzaj kwestii i sporów teoretycznych i debat rzucają światło te dokumenty? Na jaki rodzaj pytań historycznych i/lub socjologicznych pomagają one odpowiedzieć? Jaki rodzaj uzasadnionych wniosków można sformułować na podstawie informacji zawartych w tych dokumentach? Jaki rodzaj uogólnień można poczynić na bazie informacji zawartych w tych dokumentach?

suje się do badań historyczno-porównawczych. To właśnie rozumienie ze strony obdarzonego wyobraźnią badacza może tchnąć życie w analizowane dane i nadać im znaczenie.

Badacz prowadzący analizy historyczno-porównawcze musi odnaleźć prawidłowości wśród wielkiej ilości szczegółów, opisujących przedmiot badań. Często przyjmują one formę, którą Weber nazywał *typami idealnymi*: modeli pojęciowych, złożonych z zasadniczych cech zjawisk społecznych. Tak więc na przykład sam Weber przeprowadził znaczące badania dotyczące biurokracji. Obserwując wiele rzeczywistych biurokracji, Weber ([1925] 1946) wyszczególnił cechy zasadnicze biurokracji w ogóle: obszary kompetencyjne, hierarchicznie ustrukturyzowaną władzę, dokumenty pisane itd. Weber nie ograniczył się do prostego skatalogowania cech wspólnych wszystkim rzeczywistym biurokracjom, jakie obserwował. Raczej, by stworzyć teoretyczny model „doskonałej" (typ idealny) biurokracji, musiał w pełni zrozumieć zasadnicze elementy działania biurokracji. Rysunek 11.4 przedstawia bardziej współczesny obraz graficzny pewnych pozytywnych i negatywnych aspektów biurokracji jako powszechnego zjawiska społecznego.

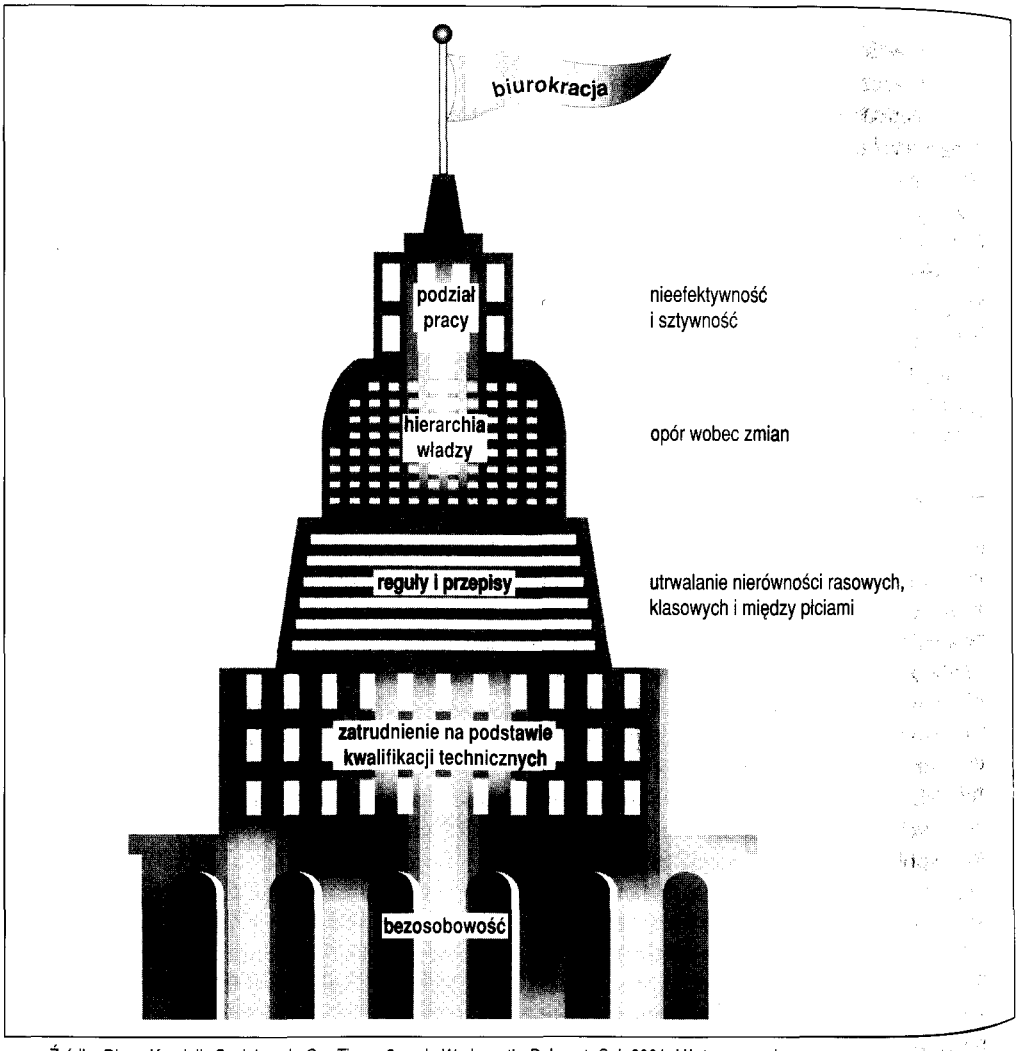
Często badania historyczno-porównawcze są ukształtowane przez konkretny paradygmat teoretyczny. Naukowcy marksistowscy mogą zatem podejmować analizy historyczne poszczególnych sytuacji – takich jak historia Latynosów i Latynosek w Stanach Zjednoczonych – by określić, czy mogą one zostać wyjaśnione

w kategoriach marksistowskiej wersji teorii konfliktu. Czasami badacze posługujący się analizą historyczno-porównawczą próbują powtarzać wcześniejsze badania w nowych sytuacjach – na przykład badania Bellaha religii Tokugawa w kontekście badań Webera nad religią i gospodarką.

Chociaż badania historyczno-porównawcze są często postrzegane jako technika raczej jakościowa niż ilościowa, wcale nie musi tak być. Analitycy historyczni czasami używają danych w szeregach czasowych, takich jak dane na temat ludności, wskaźniki przestępstw, bezrobocia, śmiertelności niemowląt itd., by śledzić zmieniające się w czasie okoliczności. Jednak analiza takich danych wymaga czasami dobrego przygotowania. Na przykład Larry Isaac i Larry Griffin (1989) omawiają użycie pewnej odmiany technik regresji (zobacz rozdział 17) do określania znaczących punktów przełomowych w procesach historycznych oraz do określania okresów, w których występowały pewne związki

Rysunek 11.4

Niektóre pozytywne i negatywne aspekty biurokracji



Źródło: Diana Kendall, *Sociology in Our Times*, 3 wyd., Wadsworth, Belmont, Cal. 2001. Użyte za zgodą.

między zmiennymi. Krytykując tendencję do traktowania historii jako równomiernie rozwijającego się procesu, autorzy koncentrują uwagę na statystycznych związkach między uzwiązkowieniem a częstością strajków, wykazując, że związek ten znacząco zmieniał się w czasie.

Isaac i Griffin podnoszą kilka ważnych kwestii dotyczących związków między teorią, metodami badawczymi a „faktami historycznymi", do których się one odnoszą. Ich analizy raz jeszcze ostrzegają przed naiwnym założeniem, że historia przedstawiona w dokumentach jest tym samym, co faktycznie się wydarzyło.

Główne zagadnienia

- Miary niereaktywne są sposobem badania zachowań społecznych bez wpływania na ich proces.
- Analiza treści jest metodą badań społecznych, dostosowaną do badania komunikowania się ludzi poprzez wytwory społeczne. Badacze mogą używać jej do badania nie tylko procesów porozumiewania się, lecz także innych aspektów zachowań społecznych.
- Powszechnie stosowane jednostki analizy w analizie treści obejmują elementy przekazów – słowa, akapity, książki itd. Niekiedy w analizie treści właściwe jest zastosowanie standardowych technik losowego doboru próby.
- Analiza treści wymaga kodowania – transformacji danych surowych na kategorie, oparte na jakimś schemacie pojęciowym. Kodowanie może dotyczyć zarówno treści jawnych, jak i ukrytych. Określenie treści ukrytych wymaga oceny ze strony badacza.
- Do interpretacji danych, pochodzących z analizy treści, odpowiednie są zarówno techniki ilościowe, jak i jakościowe.
- Zalety analizy treści obejmują oszczędność, bezpieczeństwo i możliwość badania procesów, trwających przez długi czas. Jej wady są następujące: jest ograniczona do zapisanych przekazów i mogą jej dotyczyć problemy związane z rzetelnością i trafnością.
- Wiele rządowych i pozarządowych agencji dostarcza zagregowanych danych statystycznych, które można wykorzystać do badania różnych aspektów życia społecznego.
- Problem trafności w analizie istniejących danych statystycznych często można rozwiązać poprzez wnioskowanie logiczne i powtarzalność.
- Istniejących danych statystycznych często dotyczą problemy związane z rzetelnością, należy więc korzystać z nich ostrożnie.
- Badacze społeczni używają metod historyczno-porównawczych do odnajdywania prawidłowości w historii różnych kultur.
- Chociaż często traktowane jako metoda jakościowa, badania historyczno-porównawcze mogą wykorzystywać techniki ilościowe.

Kluczowe pojęcia

badania niereaktywne
analiza treści
kodowanie
treść jawna
treść ukryta
badania historyczno-porównawcze

Pytania i ćwiczenia

1. W dwóch lub trzech akapitach nakreśl schemat analizy treści, mającej określić, która partia, demokratyczna czy republikańska, bardziej wspiera podstawowe prawa konstytucyjne (wolność słowa, wolność wyznania, ochrona przed samooskarżeniem itd.). Nie zapomnij określić jednostek analizy i metod doboru próby.
2. Opisz schemat kodowania, którego mógłbyś użyć w analizie treści w punkcie 1.
3. Posługując się siecią internetową, dowiedz się, jak wiele krajów ma większą „oczekiwaną długość życia" niż Stany Zjednoczone (możesz spróbować poprzez Population Reference Bureau pod adresem: http://www.prb.org/).
4. Max Weber podjął szeroko zakrojone badania nad niektórymi z głównych religii światowych. Stwórz bibliografię jego prac z tego zakresu.
5. Zlokalizuj sekcję Amerykańskiego Towarzystwa Socjologicznego (*American Sociolo-*

gical Association), działającą pod nazwą „Socjologia porównawcza i historyczna". Streść jakiś artykuł z biuletynu tej sekcji.

Dodatkowe lektury

Baker Vern, Lambert Charles, *The National Collegiate Athletic Association and the Governance of Higher Education*, „Sociological Quarterly" 1990, 31(3), s. 403–421. Historyczna analiza czynników, prowadzących do powstania i ukształtowania NCAA.

Berg Bruce L., *Qualitative Research Methods for the Social Sciences*, 3 wyd., Allyn and Bacon, Boston 1998. Książka zawiera wspaniałe materiały, dotyczące miar niereaktywnych, w tym rozdział o analizie treści. Berg skupiając się na badaniach jakościowych, pokazuje powiązania logiczne między podejściem jakościowym a ilościowym.

Comparative Methodology: Theory and Practice in International Social Research, Øyen Else (red.), Sage, Newbury Park, Cal. 1990. Oto rozmaite punkty widzenia różnych aspektów badań porównawczych. Stosownie do tego współpracownicy pochodzą z wielu różnych krajów.

Evans William, *Computer-Supported Content Analysis: Trends, Tools, and Techniques*, „Social Science Computer Review" 1996, 14(3), s. 269–279. Znajdziemy tutaj przegląd aktualnego oprogramowania komputerowego do analizy treści, takiego jak CETA, DICTION, INTEXT, MCCA, MECA, TEXTPACK, VBPro i WORDLINK.

U.S. Bureau of the Census. 1996. *Statistical Abstract of the United States, 1996, National Data Book and Guide to Sources*. Washington DC: U.S. Government Printing Office. Jest to absolutnie najlepsza dostępna oferta książkowa (poza niniejszą). Chociaż setki stron tabel statystycznych nie są ekscytującą lekturą przed snem – druk jest dosyć mały – jest to absolutnie podstawowy tom źródłowy dla każdego badacza społecznego. Obecnie ten dokument jest również dostępny w wersji na płycie CD-ROM.

Webb Eugene T., Campbell Donald T., Schwartz Richard D., Sechrest Lee, Grove Janet Belew, *Nonreactive Measures in the Social Sciences*, Houghton Mifflin, Boston 1981. Kompendium miar niereaktywnych. Zawiera zapisy fizyczne, różne źródła archiwalne i obserwacje. Dobre omówienie związanych z tym typem badań zagadnień etycznych i ograniczeń takich miar.

Weber Robert Philip, *Basic Content Analysis*, Sage, Newbury Park, Cal. 1990. Wspaniała książka dla początkujących, omawiająca projektowanie i realizację analizy treści. Przedstawiono w niej zarówno ogólne zagadnienia, jak i konkretne techniki.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 12

Badania ewaluacyjne

Przegląd holograficzny

W rozdziale tym poznasz najszybciej rozwijający się sposób wykorzystania badań społecznych: ewaluację interwencji społecznych. Po jego lekturze będziesz w stanie ocenić, czy programy takie kończą się sukcesem, czy porażką.

Wprowadzenie

Tematy odpowiednie dla badań ewaluacyjnych

Sformułowanie problemu: zagadnienia związane z pomiarem
Określenie wyników
Mierzenie tła eksperymentu
Określenie interwencji
Określenie populacji
Nowe mierniki a mierniki już istniejące
Operacjonalizacja sukcesu lub porażki

Typy schematów badań ewaluacyjnych
Schematy eksperymentalne
Schematy quasi-eksperymentalne
Ewaluacja jakościowa

Kontekst społeczny
Problemy logistyczne
Niektóre zagadnienia etyczne
Wykorzystanie wyników badań

Badania wskaźników społecznych
Kara śmierci i odstraszanie
Symulacja komputerowa

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Możesz nie znać słuchowiska *Twende na Wakati* (*Podążajmy z duchem czasu*), choć jest to najbardziej popularny program radiowy w Tanzanii. W tej operze mydlanej głównym bohaterem jest Mkwaju – kierowca ciężarówki. Ma on dość tradycyjne poglądy na temat ról społecznych związanych z płcią oraz z seksem. W programie tym występują również krawiec Fundi Mitindo i jego żona Mama Waridi, prezentujący całkiem odmienne, bardziej nowoczesne poglądy na role kobiece i męskie, zwłaszcza w związku z problemami przeludnienia i planowania rodziny.

Słuchowisko *Twende na Wakati* zostało stworzone przez Population Communications International (PCI) oraz inne organizacje, współpracujące z rządem Tanzanii. Była to odpowiedź na dwa problemy, dotykające ten kraj: (1) dwukrotnie wyższy niż w innych krajach przyrost naturalny oraz (2) epidemię AIDS, występującą przede wszystkim wzdłuż międzynarodowych szlaków komunikacyjnych, gdzie w 1991 roku u ponad jednej czwartej kierowców ciężarówek i u ponad połowy kobiet świadczących usługi seksualne wykryto wirusa HIV. Rozpowszechnienie środków antykoncepcyjnych wynosiło 11% (Rogers i in. 1996, s. 5–6).

Celem tej opery mydlanej było wywołanie zmiany w zakresie wiedzy, postaw i działań związanych z antykoncepcją i planowaniem rodziny. Organizacja PCI sądziła, że poprzez rozrywkę można przekazać to przesłanie skuteczniej niż poprzez przeprowadzenie tradycyjnej kampanii edukacyjnej.

W latach 1993–1995 wyemitowano 108 odcinków *Twende na Wakati*, adresując je do 67% mieszkańców Tanzanii, którzy słuchają radia. Do słuchania programu przyznawało się 84% słuchaczy, co sprawiło, że była to najpopularniejsza audycja radiowa w kraju. Mkwaju, seksistowski kierowca ciężarówki, był rozpoznawany przez 90% odbiorców słuchowiska, a tylko 3% słuchaczy uważało go za przykład pozytywnej roli. Ponad dwie trzecie słuchaczy rozpoznawało w postaci bizneswoman Mama Waridi i w jej mężu krawcu przykłady ról pozytywnych.

Badania sondażowe, przeprowadzone w celu zmierzenia oddziaływania słuchowiska, wskazują na jego wpływ na wiedzę, postawy i zachowania. Na przykład 49% zamężnych kobiet słuchających tego programu zadeklarowało, iż obecnie stosuje zasady planowania rodziny; dla porównania zasady takie stosuje tylko 19% niesłuchających programu. Pojawiły się także inne efekty:

> „W 1994 roku około 72% słuchaczy przyznało, że zastosowało się do zasad profilaktyki dotyczącej HIV/AIDS w wyniku słuchania *Twende na Wakati*; odsetek ten wzrósł do 82% w badaniach sondażowych z roku 1995. Spośród tych osób 77% przyjęło zasady monogamii, 16% zaczęło stosować prezerwatywy, 6% zaś przestało korzystać ze wspólnych maszynek do golenia i igieł".
>
> (Rogers i in. 1996, s. 21)

Możemy ocenić efektywność tej opery mydlanej dzięki szczególnej formie badań społecznych. *Badania ewaluacyjne* – czasem nazywane *ewaluacją programu* – odnoszą się w większym stopniu do celu badań niż do użycia specyficznej metody. Celem tym jest ocena (ewaluacja) wpływu podejmowanych działań (interwencji) społecznych, takich jak nowe metody nauczania, innowacje w zasadach zwolnień warunkowych oraz wiele innych. Do badań ewaluacyjnych mogą być wykorzystane różne metody – między innymi sondaże, eksperymenty.

Ewaluacja jest chyba tak stara, jak same badania społeczne. Zawsze gdy wprowadzano jakąś reformę społeczną dla osiągnięcia określonych celów, zwracano uwagę na jej faktyczne konsekwencje, nawet jeśli nie robiono tego w sposób świadomy, przemyślany czy wyrafinowany. Jednak w ostatnich latach obszar badań ewaluacyjnych cieszy się rosnącą popularnością i szczególną aktywnością badawczą, co znajduje odbicie w podręcznikach, wykładach i projektach. Co więcej, rozwój badań ewaluacyjnych wskazuje na ogólniejszy trend w naukach społecznych. Jako badacz zostaniesz kiedyś prawdopodobnie poproszony o samodzielne przeprowadzenie ewaluacji.

Częściowo rozwój obszaru badań ewaluacyjnych odpowiada rosnącemu pragnieniu badaczy społecznych, dotyczącemu wpływania na świat. Nie możemy jednak lekceważyć wpływu:

(1) rosnących wymagań władz centralnych, domagających się, by ewaluacja towarzyszyła wdrażaniu nowych programów, oraz (2) dostępności funduszy badawczych dla spełnienia tych wymagań. W każdym razie wydaje się jasne, że naukowcy zajmujący się naukami społecznymi będą wnosić swoje umiejętności do realnego świata w stopniu większym niż kiedykolwiek w przeszłości.

Rozdział ten prezentuje niektóre kluczowe elementy tej formy badań społecznych. Po rozważeniu tematów, będących zwykle przedmiotem ewaluacji, przejdziemy ku niektórym głównym aspektom operacjonalizacji: pomiarowi, planowi badań i jego realizacji. Jak się przekonasz, stawianie pytań jest równie ważne, jak udzielanie na nie odpowiedzi. Ponieważ badania ewaluacyjne dotyczą prawdziwego życia, rodzą one swoiste problemy. Niektóre z nich będziemy tu omawiać. Poza problemami logistycznymi, badania ewaluacyjne ogółem oraz stosowane w nich określone procedury stwarzają specyficzne problemy etyczne. Przeglądając raporty z ewaluacji programów, należy zachować szczególną wrażliwość na ten typ zagadnień.

Ewaluacja jest formą badań stosowanych, których zakładanym celem jest wywieranie pewnego wpływu na otaczającą nas rzeczywistość. Warto się zatem zastanowić, czy jej wyniki faktycznie znajdują zastosowanie, a jeśli tak, to jakie. Jak się przekonasz, oczywiste wnioski, wypływające z badań ewaluacyjnych, niekoniecznie wpływają na rzeczywistość. Czasami mogą stać się zarzewiem debaty ideologicznej, a nie naukowej. Mogą zostać kategorycznie odrzucone ze względów politycznych lub innych. Ale chyba najczęściej są one po prostu ignorowane lub zapomniane, skazane na pokrycie się kurzem na regałach w całym kraju.

Rozdział ten kończy się omówieniem szczególnego rodzaju zasobów, przydatnych w szeroko zakrojonej ewaluacji – badań wskaźników społecznych. Ten typ badań również stanowi rozwijającą się szybko specjalizację. Zasadniczo wymaga on stworzenia zagregowanych wskaźników „kondycji" społeczeństwa, na podobieństwo wskaźników ekonomicznych, służących diagnostyce i prognozom gospodarczym.

Tematy odpowiednie dla badań ewaluacyjnych

Badania ewaluacyjne są stosowne zawsze, gdy ma miejsce lub jest planowana jakaś interwencja społeczna. *Interwencja społeczna* jest działaniem podejmowanym w określonym kontekście społecznym w celu osiągnięcia pewnego zakładanego skutku. W najprostszym znaczeniu **badania ewaluacyjne** są procesem ustalania, czy dana interwencja społeczna osiągnęła zakładany skutek.

Tematy właściwe dla badań ewaluacyjnych są nieco ograniczone. Kiedy rząd centralny zniósł system selektywnego poboru do armii (projekt), badacze wojskowi zaczęli zwracać szczególną uwagę na wpływ tej decyzji na rekrutację do wojska. Gdy poszczególne stany liberalizowały prawo, dotyczące używania marihuany, badacze byli zmuszeni do poszukiwania wiedzy, dotyczącej konsekwencji tych zmian, zarówno dla używania marihuany, jak i dla innych form zachowań społecznych. Czy reformy dotyczące rozwodów bez orzekania o winie zwiększają liczbę rozwodów, i czy związane z nimi problemy społeczne zwiększają się, czy zmniejszają? Czy wypłata odszkodowań komunikacyjnych bez orzekania o winie rzeczywiście prowadzi do obniżki składek za polisy?

Kilka lat temu kontrowersje wywołał projekt ewaluacyjny, dotyczący krajowego programu kształcenia kierowców, realizowany przez NHTSA (National Highway and Transportation Safety Administration). Jak stwierdził Philip Hilts we wnioskach z badań:

> „Branża ubezpieczeń komunikacyjnych przez lata przyznawała znaczne zniżki młodzieży uczestniczącej w kursach dla kierowców, ponieważ statystyki wskazują, że grupa ta powoduje mniej wypadków.
>
> Jednak wstępne wyniki najnowszych szeroko zakrojonych badań wskazują na to, że kształcenie kierowców wcale nie zapobiega wypadkom i nie zmniejsza ich liczby".
>
> (1981, s. 4)

Wstępne wyniki badań, oparte na analizie przypadków 17 500 młodych osób z hrabstwa DeKalb w stanie Georgia (włączając w to Atlantę),

wykazały, że uczniowie, którzy uczestniczyli w kursach dla kierowców, spowodowali tyle samo wypadków i zakłóceń komunikacyjnych co ci, którzy w takich kursach nie uczestniczyli. Badania te, jak się wydaje, ujawniły także pewne bardziej subtelne aspekty kształcenia kierowców.

Analiza wykazała przede wszystkim, że pozorny wpływ kształcenia kierowców był w znacznej mierze wynikiem autoselekcji. Ten typ uczniów, który brał udział w kształceniu kierowców, rzadziej powodował wypadki i zakłócenia w ruchu – niezależnie od tego, czy uczestniczył w kursach, czy nie. Na przykład lepsi uczniowie wykazywali większe zainteresowanie uczestnictwem w szkoleniach kierowców; ich także w mniejszym stopniu dotyczyło prawdopodobieństwo uczestnictwa w wypadkach.

Jednak najbardziej zaskakująca była sugestia, że kursy dla kierowców mogą w rzeczywistości powodować zwiększenie liczby wypadków samochodowych! Oferowanie uczestnictwa w takich kursach może zachęcić niektórych uczniów do uzyskiwania prawa jazdy wcześniej, niż gdyby kursy te nie były oferowane. W badaniach przeprowadzonych w dziesięciu miastach w stanie Connecticut, które zaprzestały szkolenia kierowców, wykazano, że około 3/4 uczniów, którzy prawdopodobnie zdobyliby prawo jazdy, uczestnicząc w kursach, wstrzymało się z decyzją dotyczącą uzyskania prawa jazdy aż do momentu ukończenia 18 lat (Hilts 1981, s. 4).

Jak łatwo sobie wyobrazić, wyniki te nie zostały dobrze przyjęte przez tych, którzy byli najbardziej związani ze szkoleniem kierowców. Co więcej, sprawę komplikował fakt, iż badania prowadzone przez NHTSA dotyczyły także ewaluacji nowego, bardziej intensywnego programu szkoleń – a ich wstępne wyniki wykazywały, że nowy program był efektywny.

Kolejny przykład badań ewaluacyjnych dotyczy całkowicie odmiennej problematyki. Przejście do gospodarki rynkowej na Węgrzech było przedmiotem szczególnego zainteresowania Rudolfa Andorki, socjologa z tego kraju. Jeszcze przed dramatycznymi wydarzeniami w Europie Wschodniej w 1989 roku Andorka i jego koledzy monitorowali „drugi obieg" gospodarki narodowej – pracę poza sferą gospodarki socjalistycznej. W swoich sondażach śledzili oni rozwój i upadek szarej sfery oraz badali jej oddziaływanie na społeczeństwo węgierskie. Jedna z konkluzji mówiła, iż „drugi obieg gospodarczy, który uprzednio prawdopodobnie prowadził do zmniejszenia nierówności w dochodach lub przynajmniej poprawiał standard życia najuboższej części ludności, w latach 80. coraz bardziej przyczyniał się do wzrostu tych nierówności" (Andorka 1990, s. 111).

Jak można zauważyć, pytania charakterystyczne dla badań ewaluacyjnych mają duże znaczenie praktyczne: przedmiotem ich zainteresowania są miejsca pracy, programy, inwestycje, jak również przekonania i wartości. Przyjrzyjmy się teraz temu, jak odpowiada się na tego typu pytania – jak prowadzone są badania ewaluacyjne.

Sformułowanie problemu: zagadnienia związane z pomiarem

Kilka lat temu kierowałem instytucjonalnym biurem badań, które realizowało badania, bezpośrednio odnoszące się do działania uniwersytetu. Często proszono nas o ewaluację projektów nowych programów nauczania. Poniższy opis jest dość typowym problemem, który pojawił się w tym kontekście, i wskazuje on na jedną z głównych barier dla dobrych badań ewaluacyjnych.

Profesorowie wydziału pojawiali się w moim biurze, mówiąc, że administracja uniwersytecka poleciła przygotować ewaluację nowego programu, na którego przetestowanie uzyskali zgodę. Pokazuje to pewien powszechny problem: często ludzie, których programy są ewaluowane, nie są podekscytowani tą perspektywą. Dla nich niezależna ewaluacja zagraża bytowi ich programu, a może nawet ich miejscom pracy.

Jednakże zasadniczy problem, który chcę przedstawić, dotyczy celu interwencji, która ma zostać poddana ewaluacji. Pytanie: „co jest zakładanym rezultatem nowego programu?" często prowadzi to takiej mglistej odpowiedzi, jak: „studenci osiągną głębokie i autentyczne rozumienie matematyki zamiast prostych, pamięciowych metod dokonywania obliczeń". Genialne! A jak będziemy mierzyć to „głębokie i autentyczne rozumienie"? Często mówiono

mi, że celem programu są efekty, których nie da się zmierzyć poprzez konwencjonalne testy uzdolnień i osiągnięć. Nie ma problemu; tego się należy spodziewać, kiedy wprowadzamy coś innowacyjnego i wykraczającego poza istniejące konwencje. Co będzie niekonwencjonalnym miernikiem zakładanego rezultatu? Czasem taka dyskusja prowadziła do stwierdzenia, że efekty programu będą „niemierzalne".

Jest to często pojawiająca się trudność badań ewaluacyjnych: mierzenie tego co „niemierzalne". W badaniach ewaluacyjnych chodzi o określenie, czy coś istnieje, czy nie, czy coś się zdarzyło, czy nie. Aby móc przeprowadzić takie badania, musimy być w stanie operacjonalizować, obserwować i rozpoznawać zaistnienie tego, co jest przedmiotem naszych badań, lub jego brak.

Często rezultaty mogą zostać wywiedzione z opublikowanej dokumentacji programu. Było tak, kiedy Edward Howard i Darlene Norman (1981) ewaluowali funkcjonowanie Biblioteki Publicznej Hrabstwa Vigo w stanie Indiana. Rozpoczęli od stwierdzenia celów, uprzednio przyjętych przez członków zarządu biblioteki.

> „Pozyskiwać, drogą zakupu lub darowizny, oraz przez utrwalanie i wytwarzanie, odpowiednie i potencjalnie użyteczne informacje, wytwarzane przez obywateli tej społeczności, o nich lub dla nich;
>
> Porządkować te informacje w celu efektywnego dostarczenia i wygodnego do nich dostępu; wyposażać w sprzęt niezbędny do korzystania z nich; zapewnić pomoc w ich wykorzystaniu; oraz
>
> Doprowadzić do maksymalnego wykorzystania tych informacji dla uczynienia społeczności lepszym miejscem do życia poprzez wspomaganie dążenia do wiedzy jej mieszkańców".
>
> (1981, s. 306)

Jak stwierdzili badacze: „Wszystko, co zrobiła Biblioteka Publiczna Hrabstwa Vigo, może być sprawdzone w odniesieniu do zakładanych celów". Następnie przystąpili do stworzenia mierników operacyjnych dla każdego z celów.

„Oficjalne" cele interwencji są często kluczem do zaprojektowania ewaluacji, chociaż nie zawsze okazują się wystarczające. Na przykład Anna-Marie Madison (1992) przestrzega, że programy, których celem jest niesienie pomocy upośledzonym mniejszościom, nie zawsze odzwierciedlają potrzeby i pragnienia tych, do których pomoc ta jest adresowana.

> „Obciążenia kulturowe, obecne w sposobie, w jaki biali badacze wywodzący się z klasy średniej interpretują doświadczenia mniejszości o niskich dochodach, mogą prowadzić do błędnych założeń i fałszywych twierdzeń, dotyczących przypadkowych zależności, do nietrafnych teorii społecznych i w konsekwencji do nietrafnych założeń teoretycznych programu. Opisowe teorie, wywiedzione z błędnych przesłanek, uprawomocnione w literaturze jako istniejąca wiedza, mogą mieć negatywne konsekwencje dla uczestników programu".
>
> (1992, s. 38)

Tak więc przygotowując ewaluację, badacze muszą zwracać szczególną uwagę na kwestię pomiaru. Przyjrzyjmy się teraz dokładniej typom pomiarów, z którymi ewaluatorzy muszą sobie radzić.

Określenie wyników

Jak już wskazywałem, dla ewaluatorów kluczową mierzoną zmienną jest ta, która mierzy wynik; jest ona także nazywana *zmienną reakcji* (*response variable*). Jeśli program społeczny ma coś osiągnąć, musimy być w stanie to „coś" zmierzyć. Jeśli chcemy zmniejszyć uprzedzenia, musimy być w stanie mierzyć uprzedzenia. Jeśli chcemy zwiększyć harmonię małżeńską, musimy być w stanie ją zmierzyć.

Zasadniczą sprawą jest uprzednie uzgodnienie definicji:

> „Najtrudniejsza sytuacja pojawia się wtedy, gdy nie ma zgody co do standardów. Na przykład wiele zainteresowanych stron może nie zgadzać się co do tego, co jest definiowane jako *poważne* nadużywanie narkotyków – czy lepszą jego definicją jest co najmniej 15% uczniów zażywających narkotyki w ciągu tygodnia, czy też co najmniej 5% zażywających twarde narkotyki, takie jak ko-

kaina lub PCP[1] w ciągu miesiąca; czy też zażywanie narkotyków już w siódmej klasie, a może jakaś kombinacja tych wymiarów: wskaźników procentowych, sposobów zażywania oraz wieku? (...) Badacze zainteresowani zastosowaniem wyników badań powinni już przed przystąpieniem do badań dążyć – na ile to możliwe – do osiągnięcia konsensusu z odbiorcami wyników (na przykład poprzez grupy doradcze) lub przynajmniej powinni upewnić się, że ich badania mogą przynieść dane, odpowiadające standardom przyjmowanym przez wszystkich potencjalnie zainteresowanych".

(Hedrick, Bickman, Rog 1993, s. 27)

W niektórych przypadkach może się okazać, że definicje problemu i wystarczającego rozwiązania są określone prawnie lub poprzez przepisy danej instytucji; jeśli tak, musisz być świadom istnienia takich określeń i zastosować się do nich. Co więcej, niezależnie od uzgodnionych definicji, musisz także osiągnąć porozumienie co do sposobu przeprowadzenia pomiarów. Ze względu na to, że istnieją różne metody, które można zastosować w celu oszacowania na przykład odsetka uczniów „używających narkotyków w ciągu tygodnia", musisz mieć pewność, że wszystkie zainteresowane strony rozumieją i akceptują metodę lub metody wybrane przez ciebie.

W przypadku opery mydlanej z Tanzanii wyniki mierzono za pomocą kilku miar. Po części celem programu było rozszerzenie wiedzy dotyczącej zarówno planowania rodziny, jak i AIDS. Tak więc na przykład jeden z odcinków rozwiewał przekonanie, że wirus HIV jest roznoszony przez komary i można mu przeciwdziałać, używając środków odstraszających te owady. Badania wśród słuchaczy wykazały zmniejszenie się tego przekonania (Rogers i in. 1996, s. 21).

Organizacja Population Communications International (PCI) zamierzała również zmienić postawy mieszkańców Tanzanii dotyczące wielkości rodziny, ról związanych z płcią, AIDS/HIV i innych tematów związanych z tymi zagadnieniami. Badania wykazały, że słuchowisko wpłynęło także na te postawy. Ostatecznie program miał też na celu zmianę zachowań. Wiemy już, że słuchacze relacjonowali zmiany swoich zachowań, dotyczących zabezpieczenia się przed AIDS. Relacjonowali oni również stosowanie w większym stopniu metod planowania rodziny. Jednak ze względu na to, że zawsze może istnieć rozdźwięk między tym, co ludzie mówią, a tym, co faktycznie robią, badacze poszukiwali niezależnych danych, które pozwoliłyby potwierdzić ich wnioski.

Krajowy program kontroli AIDS w Tanzanii oferował bezpłatnie prezerwatywy. Na terenach, gdzie odbierano słuchowisko, liczba rozdanych prezerwatyw wzrosła sześciokrotnie w latach 1992–1994. Było to znacznie więcej niż wzrost 1,4 raza, odnotowany na obszarze kontrolnym, gdzie stacje radiowe nie nadawały tej audycji.

Mierzenie tła eksperymentu

Mierzenie zmiennych zależnych, bezpośrednio związanych z programem eksperymentalnym, to dopiero początek. Jak zaznaczają Henry Riecken i Robert Boruch (1974, s. 120–121), często właściwe i ważne jest mierzenie tych aspektów kontekstu eksperymentu, o których badacze sądzą, iż mogą mieć wpływ na ten eksperyment. Niektóre zmienne, choć zewnętrzne w stosunku do samego eksperymentu, mogą na niego wpływać.

Załóżmy, iż na przykład przeprowadzasz ewaluację programu, mającego za zadanie szkolenie niewykwalifikowanych osób w celu ich zatrudnienia. Podstawowym miernikiem wyników byłoby znalezienie przez nich zatrudnienia po ukończeniu tego programu. Można oczywiście obserwować i obliczać wskaźnik zatrudnienia wśród osób przeszkolonych, lecz podczas ewaluacji należy także określić, co działo się ze wskaźnikami zatrudnienia/bezrobocia w całym społeczeństwie. Ogólne załamanie na rynku pracy powinno być brane pod uwagę podczas oceny tego, co w innych okolicznościach wydawałoby się dość niskim wskaźnikiem zatrudnienia wśród uczestników programu. Lub też, jeśli w następstwie danego programu wszyscy objęci eksperymentem znajdą pracę, należy wziąć pod

[1] Lek przeciwbólowy, powodujący uzależnienie (przyp. red. nauk.).

uwagę każdy ogólny wzrost zatrudnienia. Łączenie pomocniczych mierników z właściwymi schematami z grupą kontrolną powinno stworzyć możliwość dokładnego określenia efektów programu poddanego ewaluacji.

Określenie interwencji

Poza dokonywaniem pomiarów odnoszących się do wyników programu, badacz musi także dokonać pomiaru interwencji programu – bodźca eksperymentalnego. Po części pomiar ten będzie polegał na przypisaniu poszczególnych osób do grupy kontrolnej i eksperymentalnej, jeśli taki jest schemat badawczy. Przypisanie danej jednostki do grupy eksperymentalnej jest tym samym, co wystawienie jej na oddziaływanie bodźca, przypisanie zaś do grupy kontrolnej oznacza brak bodźca. W praktyce nie jest to jednak takie proste.

Pozostańmy przy przykładzie szkoleń zawodowych. Niektóre osoby będą uczestniczyć w tym programie, inne nie. Lecz wyobraźmy sobie przez moment, jak zapewne wyglądają te programy: niektóre osoby będą uczestniczyć w nich w pełni, inne opuszczą wiele zajęć lub będą się obijać, gdy przyjdą. Tak więc może będziesz musiał zmierzyć stopień lub jakość uczestnictwa w programie. Jeśli jest on efektywny, powinno się okazać, że osoby, które uczestniczyły w nim w pełni, mają wyższy wskaźnik zatrudnienia niż osoby, które uczestniczyły w nim mniejszym stopniu.

Zarządzanie bodźcem eksperymentalnym mogą zakłócać jeszcze inne czynniki. Załóżmy, że ewaluujemy nową formę psychoterapii, opracowaną w celu leczenia impotencji seksualnej. Kilku terapeutów stosuje ją wobec osób, które będą tworzyć naszą grupę eksperymentalną. Planujemy porównanie wskaźnika wyzdrowień w grupie eksperymentalnej ze wskaźnikiem uzyskanym w grupie kontrolnej, która uczestniczy w innych formach terapii lub nie leczy się wcale. Użyteczne może być uwzględnienie nazwisk terapeutów, leczących poszczególne osoby w grupie eksperymentalnej, gdyż niektórzy mogą być bardziej efektywni niż inni. Jeśli tak będzie w tym przypadku, musimy odpowiedzieć, dlaczego w przypadku jednych terapeutów terapia przynosiła lepsze efekty niż u innych. To, czego się dowiemy, pozwoli nam lepiej zrozumieć samą terapię.

Określenie populacji

Przy ewaluacji jakiejś interwencji ważne jest zdefiniowanie populacji potencjalnych beneficjentów, dla których przeznaczony jest dany program. Idealnie jest, gdy oni wszyscy lub ich próba zostanie później przypisana do grup eksperymentalnej i kontrolnej, zgodnie ze schematem badawczym. Jednak zdefiniowanie populacji może samo w sobie wymagać określenia pomiarów. Jeśli ewaluujemy na przykład nową formę psychoterapii, prawdopodobnie będzie ona odpowiednia dla osób z problemami psychicznymi. Lecz jak będziemy definiować i mierzyć „problemy psychiczne"? Program szkoleń zawodowych, wspominany uprzednio, jest zapewne adresowany do ludzi mających problemy ze znalezieniem pracy, ale co zaliczymy do „problemów"?

Poza zdefiniowaniem odpowiedniej populacji badacz powinien zatem dokonać dość dokładnego pomiaru zmiennych, uwzględnionych w definicji. Na przykład, chociaż losowy dobór osób w badaniach dotyczących psychoterapii powinien zapewnić równy rozkład osób z nieznacznymi i poważnymi problemami psychicznymi w grupach eksperymentalnej i kontrolnej, musimy mieć informacje o względnej powadze problemów poszczególnych osób, na wypadek gdyby terapia okazała się skuteczna tylko wśród osób z nieznacznymi dolegliwościami. Podobnie powinniśmy zmierzyć takie zmienne demograficzne, jak płeć, wiek, rasa itp., w razie gdyby terapia odnosiła skutek tylko w przypadku kobiet, osób starszych lub jakiejś innej grupy.

Nowe mierniki a mierniki już istniejące

Przewidując pomiary tych różnych rodzajów zmiennych, badacz musi nieustannie wybierać między tworzeniem nowych miar a wykorzystaniem mierników już opracowanych przez innych badaczy. Jeśli badania dotyczą czegoś, co nigdy wcześniej nie było mierzone, wybór jest prosty. Jeśli dotyczą czegoś, co inni próbowali już zmie-

rzyć, badacz powinien oszacować względną wartość różnych już istniejących narzędzi pomiarowych pod kątem własnej sytuacji badawczej oraz celu swoich badań. Przypomnij sobie, że jest to podstawowa zasada badań społecznych, która ma szerokie zastosowanie także poza dziedziną badań ewaluacyjnych. Omówmy pokrótce zalety tworzenia nowych mierników i wykorzystania już istniejących.

Stworzenie pomiarów specjalnie dla danych badań może być bardziej adekwatne i trafne niż wykorzystanie miar już istniejących. Jeśli psychoterapia, którą poddajemy ewaluacji, ma na celu jakiś specyficzny aspekt powrotu do zdrowia, to możemy stworzyć miary, które będą dotyczyć dokładnie tego aspektu. Możemy nie być w stanie znaleźć żadnych standaryzowanych mierników psychologicznych, które trafią w interesujący nas aspekt. Jednak stworzenie naszych własnych miar kosztować nas będzie utratę korzyści, wynikających z użycia miar już istniejących. Stworzenie dobrych miar wymaga czasu i energii; można tego uniknąć przez zaadaptowanie już istniejących technik. Większe znaczenie naukowe mają te pomiary, które były często wykorzystywane przez innych badaczy, gdyż dostarczają materiał do możliwych porównań, co może być istotne dla prowadzonej przez nas ewaluacji. Jeśli eksperymentalna terapia zwiększa wyniki uzyskiwane w standardowym teście średnio o dziesięć punktów, będzie można porównać tę terapię z innymi, które były ewaluowane przy użyciu tego samego miernika. Wreszcie, miary z długą historią ich stosowania mają zwykle znany stopień trafności i rzetelności, natomiast nowo tworzone miary będą wymagały przetestowania lub będą używane ze znaczną niepewnością.

Operacjonalizacja sukcesu lub porażki

Potencjalnie jednym z najbardziej uciążliwych aspektów badań ewaluacyjnych jest ustalenie, czy badany program odniósł sukces, czy poniósł porażkę. Celem programu nauczania języka obcego może być pomoc studentom w lepszym nauczeniu się tego języka, lecz jak duży postęp jest tu wystarczający? Celem programu odwiedzin małżonków w więzieniu może być poprawa nastrojów wśród więźniów, lecz jak bardzo mają być one poprawione, by móc uzasadnić ten program?

Jak można się spodziewać, nie ma ostatecznych odpowiedzi na tego typu pytania. Dylemat ten w oczywisty sposób stał się podstawą *analizy kosztów i korzyści* (*cost-benefit analysis*). Jak wiele dany program kosztuje w stosunku do otrzymanych z niego korzyści? Jeśli korzyści przeważają nad kosztami – utrzymajmy ten program; jeśli jest na odwrót – porzućmy go. Jest to wystarczająco proste i wydaje się stosowne w nieskomplikowanych sytuacjach ekonomicznych: jeśli produkcja kosztuje 20 USD, a możesz to „coś" sprzedać jedynie za 18 USD, nie ma jak zrekompensować kosztów.

Niestety sytuacje, z jakimi zwykle ma się do czynienia podczas badań ewaluacyjnych, rzadko kiedy poddają się prostemu rachunkowi ekonomicznemu. Nauka języka obcego może kosztować okręg szkolny 100 USD w przeliczeniu na jednego ucznia i może podnosić wyniki osiągane w testach średnio o 15 punktów. Ponieważ wyniki testów nie mogą być zamienione na dolary, nie ma oczywistego sposobu wyważenia kosztów i korzyści.

Czasami, z przyczyn praktycznych, kryteria sukcesu i porażki mogą być wypracowane poprzez konkurencję między programami. Jeśli inny program nauczania języka obcego będzie kosztować jedynie 50 USD w przeliczeniu na jednego ucznia i przyniesie wzrost o 20 punktów w wynikach testów, to bez wątpienia będzie uważany za przynoszący większe sukcesy niż ten pierwszy – zakładając, że wyniki testów były odpowiednią miarą celów zakładanych przez oba programy, oraz że mniej kosztowny program nie miał niezamierzonych negatywnych konsekwencji.

Ostatecznie kryteria sukcesu i porażki są często przedmiotem uzgodnień. Osoby odpowiedzialne za program mogą umówić się zawczasu, jaki konkretny wynik będą uznawać za wskaźnik sukcesu danego programu. W takim przypadku należy zrobić wszystko, aby uzyskać całkowitą pewność, że projekt badań będzie mierzyć ten konkretny wynik. Przypominam ten oczywisty wymóg po prostu dlatego,

iż badacze czasami tego nie robią, a nie ma nic bardziej kłopotliwego niż taka sytuacja.

Podsumowując, badacze muszą dokonywać pomiarów w badaniach ewaluacyjnych z całą powagą, starannie określając wszystkie mierzone zmienne i znajdując odpowiednie mierniki dla każdej z nich. Jednak, jak już sugerowałem, takie decyzje nie są na ogół czysto naukowe. Ewaluatorzy często muszą wypracować swoją strategię pomiaru wspólnie z osobami odpowiedzialnymi za dany program. Zwykle nie ma sensu określenie, czy dany program osiągnął rezultat X, skoro jego celem było osiągnięcie rezultatu Y. (Zauważ jednak, że czasem schematy ewaluacyjne mają na celu sprawdzenie niezmierzonych konsekwencji danej interwencji).

Istnieje także polityczny aspekt dokonywania takich wyborów. Ponieważ badania ewaluacyjne często dotyczą zawodowych interesów innych ludzi – ich ulubiony program może zostać wstrzymany, albo też oni sami mogą zostać zwolnieni z pracy lub stracić pozycję zawodową – wyniki badań ewaluacyjnych często są przedmiotem sporów.

Przejdźmy teraz do paru schematów badawczych, powszechnie wykorzystywanych przez ewaluatorów.

Typy schematów badań ewaluacyjnych

Jak zaznaczyłem na początku tego rozdziału, badania ewaluacyjne same w sobie nie są metodą, lecz raczej jednym z zastosowań metod badań społecznych. Jako takie, mogą one pociągać za sobą którykolwiek z kilku schematów badawczych. Tutaj rozważymy trzy główne typy schematów badawczych, które są odpowiednie dla ewaluacji: schematy eksperymentalne, schematy quasi-eksperymentalne oraz ewaluacje jakościowe.

Schematy eksperymentalne

Wiele schematów eksperymentalnych, przedstawionych w rozdziale 8, można użyć w badaniach ewaluacyjnych. Przyjrzyjmy się dla przykładu, jak klasyczny model eksperymentalny może być zastosowany do naszej ewaluacji, dotyczącej nowej psychoterapii, stosowanej do leczenia impotencji seksualnej.

Projektowanie takiej ewaluacji powinniśmy rozpocząć od określenia populacji pacjentów kwalifikujących się do tej terapii. Takiego rozpoznania mogą dokonać badacze eksperymentujący z nową formą terapii. Powiedzmy, że mamy do czynienia z kliniką, w której już się leczy stu pacjentów dotkniętych impotencją. Możemy przyjąć tę grupę oraz definicję impotencji seksualnej, stosowaną w tej klinice, jako punkt wyjścia. Powinniśmy zachować wszelkie istniejące oceny powagi tego problemu dla każdego konkretnego pacjenta.

Jednak dla celów badań ewaluacyjnych powinniśmy opracować dokładniejszy miernik impotencji. Może on dotyczyć tego, czy pacjenci utrzymywali w ogóle stosunki seksualne w określonym czasie, jak często odbywają stosunki seksualne, lub czy w ogóle i jak często osiągają orgazm. Inna miara wyników może bazować na ocenie niezależnych terapeutów, nie zaangażowanych w tę terapię, którzy na koniec przeprowadzą wywiady z pacjentami. W każdym przypadku będziemy musieli uzgodnić mierniki, jakie mają zostać zastosowane.

W najprostszym schemacie przypisalibyśmy losowo 100 pacjentów do grup eksperymentalnej i kontrolnej; pierwsza grupa uczestniczyłaby w nowej terapii, druga grupa nie korzystałaby z żadnej terapii podczas tego eksperymentu. Ponieważ jednak względy etyczne zapewne powstrzymałyby nas przed całkowitym pozbawieniem terapii członków grupy kontrolnej, bardziej prawdopodobne jest, iż grupa kontrolna dalej korzystałaby z konwencjonalnej terapii.

Mając już osoby przypisane do obu grup, powinniśmy uzgodnić czas trwania tego eksperymentu. Być może autorzy nowej terapii uważali, że powinna być ona skuteczna po dwóch miesiącach i da się to uzgodnić. Jednak czas trwania badań nie musi być sztywno określony. Jednym z celów tego eksperymentu i jego ewaluacji może być określenie, jak dużo czasu faktycznie potrzeba, aby nowa terapia okazała się skuteczna. Można więc sobie wy-

obrazić, że uzgodnilibyśmy, iż będziemy mierzyć odsetek wyleczeń, powiedzmy, co tydzień, a określenie ostatecznego czasu trwania eksperymentu uzależnilibyśmy od stałego sprawdzania jego wyników.

Przyjmijmy, że nowa terapia polega na pokazywaniu pacjentom filmów pornograficznych. Powinniśmy uszczegółowić ten bodziec. Jak często pacjenci będą oglądać takie filmy i jak długo będzie trwała każda z takich sesji? Czy pacjenci będą oglądać filmy indywidualnie, czy grupowo? Czy terapeuta będzie przy nich obecny? Być może powinniśmy obserwować pacjentów podczas projekcji i włączyć nasze obserwacje do pomiarów bodźca eksperymentalnego. Czy niektórzy pacjenci oglądają te filmy z entuzjazmem, a inni odwracają głowę od scen pokazywanych na ekranie? Musielibyśmy postawić pytania tego typu i dla każdego z nich opracować odpowiednie pomiary.

Gdy mamy już zaprojektowane takie badania, pozostaje nam je tylko „uruchomić". Badania zostają uruchomione, obserwacje są przeprowadzone i spisane, masa danych przygotowana do analizy. Kiedy badania zakończą się, możemy określić, czy nowa terapia prowadzi do zamierzonych – i być może, jakichś niezamierzonych – rezultatów. Możemy powiedzieć, czy filmy były skuteczniejsze w przypadku osób z nieznacznymi, czy z poważnymi problemami, działały u osób młodych, lecz u starszych nie itp.

Ten prosty przykład pokazuje, jak standardowy schemat eksperymentalny, przedstawiony w rozdziale 8, może być zastosowany w badaniach ewaluacyjnych. Jednak wiele badań, prawdopodobnie większość ewaluacji, opisanych w literaturze naukowej, nie wygląda dokładnie tak, jak w tym przykładzie. Ponieważ badania ewaluacyjne są zakorzenione w prawdziwym życiu, wymagają one często schematów quasi-eksperymentalnych. Przyjrzyjmy się, co to oznacza.

Schematy quasi-eksperymentalne

Quasi-eksperymenty różnią się od „prawdziwych" eksperymentów przede wszystkim brakiem losowego przypisania badanych do grup eksperymentalnej i kontrolnej. W badaniach ewaluacyjnych taki przydział często nie jest możliwy. Zamiast całkowicie porzucić ewaluację, badacze czasami tworzą schematy, które w jakimś stopniu umożliwiają ocenę danego programu. Poniżej opisane są niektóre takie schematy.

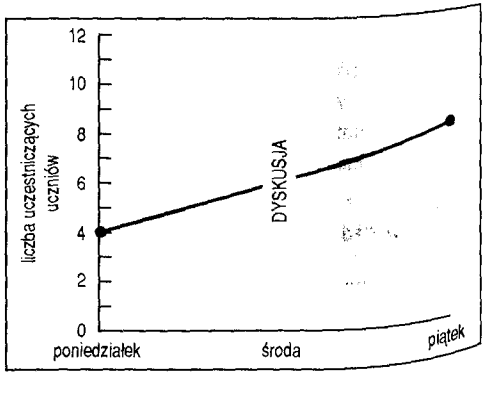
Schematy z uwzględnieniem szeregów czasowych

Dla zilustrowania schematu z uwzględnieniem szeregów **czasowych** (*time-series design*) – badań, które zakładają dokonywanie pomiaru w czasie – rozpocznę od pytania, dotyczącego oceny znaczenia pewnych hipotetycznych danych. Załóżmy, że przedstawię ci coś, co uważam za technikę mobilizowania uczniów do uczestnictwa w zajęciach. By udowodnić to twierdzenie, opowiadam ci, że w poniedziałek tylko czterech uczniów w klasie zabierało głos lub zadawało pytania; w środę poświęciłem zajęcia lekcyjne na otwartą dyskusję o pewnej kontrowersyjnej kwestii, która wzburzyła środowisko szkolnego campusu, a w piątek, kiedy powróciłem do przedmiotu kursowych zajęć, ośmiu uczniów zabierało głos lub zadawało pytania. Innymi słowy, twierdzę, że dyskusja w środę, dotycząca kontrowersyjnego zagadnienia, spowodowała podwojenie aktywnego uczestnictwa w zajęciach. Ten prosty zbiór danych przedstawiony jest graficznie na rysunku 12.1.

Czy przekonałem cię, że otwarta dyskusja w środę miała takie skutki, jak twierdzę? Praw-

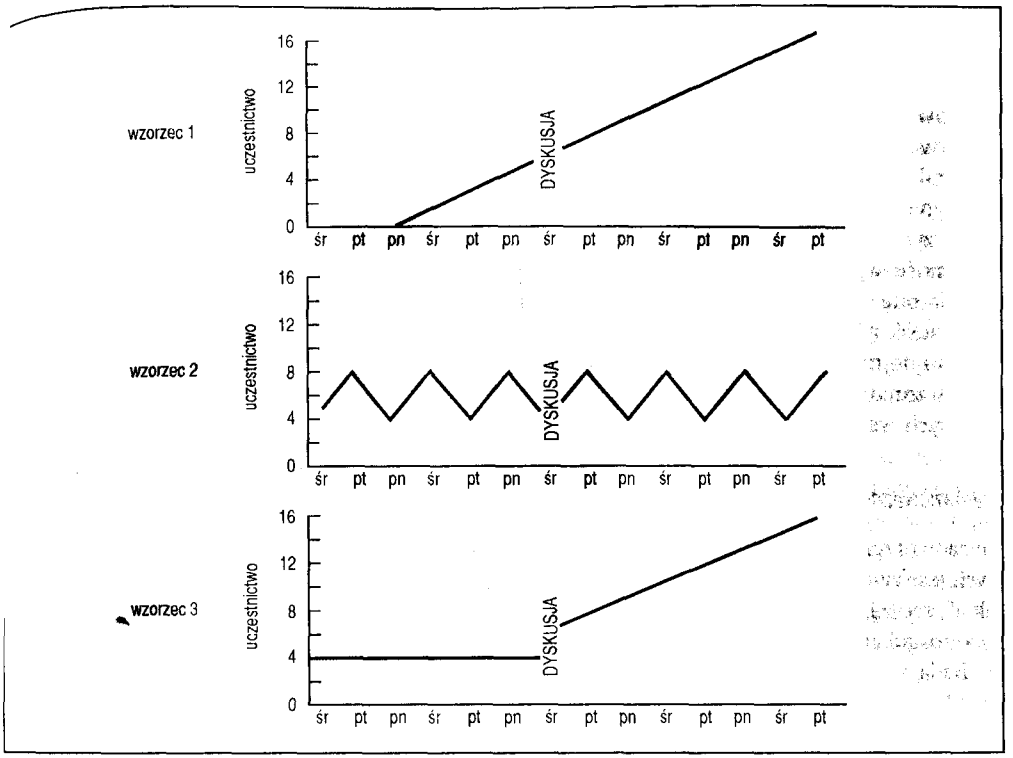
Rysunek 12.1

Dwie obserwacje dotyczące uczestnictwa z zajęciach: przed otwartą dyskusją i po niej



ysunek 12.2

rzy wzorce uczestnictwa w zajęciach w dłuższym przedziale czasu



dopodobnie zaoponowałbyś, mówiąc, że moje dane tego nie potwierdzają. Dwie obserwacje (w poniedziałek i w piątek) nie są naprawdę wystarczającym potwierdzeniem czegokolwiek. Idealnie byłoby, gdybym miał dwie klasy, w obu losowo przydzielonych uczniów i tylko w jednej z nich przeprowadziłbym otwartą dyskusję, tak by w piątek móc je obie porównać. Lecz nie mam dwóch klas z losowo dobranymi uczniami. Mam za to opis aktywnego uczestnictwa w zajęciach w ciągu semestru dla jednej klasy. Ten zapis umożliwia przeprowadzenie ewaluacji szeregów czasowych.

Rysunek 12.2 przedstawia trzy możliwe wzorce aktywnego uczestnictwa w zajęciach w czasie przed otwartą dyskusją w środę i po niej. Który z tych wzorców da ci w jakimś stopniu pewność, że dyskusja miała taki wpływ, jak twierdzę?

Gdyby wyniki szeregów czasowych wyglądały tak, jak w pierwszym wzorcu na rysunku 12.2, prawdopodobnie wyciągnąłbyś wniosek, że proces zwiększenia aktywności na zajęciach rozpoczął się w środę poprzedzającą dyskusję i trwał, niezmiennie, po dniu przeznaczonym na dyskusję. Długookresowe dane wskazują nawet, że taki trend pojawiłby się, nawet gdyby dyskusja w środę nie miała miejsca. Tak więc pierwszy wzorzec przeczy mojemu twierdzeniu, iż specjalna dyskusja zwiększyła poziom zaangażowania uczniów w trakcie zajęć.

Drugi wzorzec zaprzecza mojemu twierdzeniu, gdyż wskazuje, że uczestnictwo w zajęciach waha się w górę i w dół według regularnego wzoru w ciągu całego semestru. Czasami rośnie z jednych zajęć na drugie, czasami maleje; a otwarta dyskusja w środę odbyła się po prostu w takim momencie, gdy poziom uczestnictwa miał wzrosnąć. Co więcej, zauważamy, że aktywność na zajęciach spada ponownie na zajęciach, które odbyły się bezpośrednio po rzekomym wzroście pod wpływem dyskusji.

Tylko trzeci wzorzec zaprezentowany na rysunku 12.2 potwierdza moje twierdzenie o znaczeniu otwartej dyskusji. Jak widać, poziom dyskusji przed ową środą był stały – na kolejnych zajęciach aktywnych było po czterech uczniów. Nie dość, że poziom aktywności uczniów podwoił się na następnych zajęciach po otwartej dyskusji, lecz i później stale wzrastał. Chociaż dane te nie chronią nas przed możliwym wpływem jakichś zewnętrznych czynników (mogłem też wspomnieć, że aktywność wpłynie na stopnie uczniów), pozwalają one wykluczyć możliwość, że rosnąca aktywność była wynikiem stale postępującego procesu dojrzewania (wykazywanego w pierwszym wzorcu) lub regularnych fluktuacji (wykazywanych we wzorcu drugim).

Nierównoważne grupy kontrolne

Opisany przed chwilą schemat szeregów czasowych jest związany tylko z grupą „eksperymentalną"; nie zapewnia takiej wartości, jaka wynika z posiadania grupy kontrolnej. Czasami, kiedy badacze nie mają możliwości utworzenia grup kontrolnej i eksperymentalnej poprzez losowy przydział ze wspólnej próby, mogą znaleźć jakąś istniejącą grupę „kontrolną", która wydaje się podobna do grupy eksperymentalnej. Taka grupa jest nazywana **nierównoważną grupą kontrolną**. Jeśli na przykład próbuje się wprowadzać innowacyjny program nauczania języka obcego w jednej z klas dużej szkoły średniej, możemy znaleźć w niej inną klasę, w której uczy się tego języka, i której uczniowie stanowią bardzo podobną populację: to znaczy klasę niemal identyczną pod względem otrzymywanych stopni, proporcji płci, przynależności etnicznej, ilorazu inteligencji itp. Ta druga klasa będzie więc punktem odniesienia nawet wtedy, jeśli nie będzie stanowić formalnie części badań. Na koniec semestru obie klasy przystąpią do tego samego testu z języka obcego, a my porównamy ich osiągnięcia.

Mamy poniżej przykład, jak dwie szkoły średnie zostały wybrane w celu ewaluacji programu, którego celem było zmniejszenie problemu palenia papierosów, używania alkoholu i narkotyków:

> „Zestawienie w parę dwóch szkół i ich przypisanie do warunków «eksperymentalnych» i «kontrolnych» nie było przypadkowe. Miejscowe Towarzystwo Zapobiegania Chorobom Płuc wskazało szkołę, w której wdrożyliśmy program. Jej administracja poszukiwała rozwiązania problemów palenia, alkoholu i narkotyków, które stwierdzono w szkole. «Kontrolna» szkoła została wybrana jako odpowiednia i zbliżona pod względem cech demograficznych, a administracja zgodziła się na przeprowadzenie przez nas badań sondażowych oraz testów oddechowych. Dyrektor tej szkoły uważał istniejący program edukacji zdrowotnej za skuteczny i był przekonany, że rozpoczynanie palenia nie było powszechne wśród jego uczniów. Środowiska społeczne, z których wywodzili się uczniowie obu szkół, były do siebie podobne. Według wypowiedzi uczniów odsetek palących rodziców nieco przekraczał 40% w obu szkołach".
>
> (McAlister i in. 1980, s. 720)

Początkowe obserwacje prowadzone wśród grup eksperymentalnej i kontrolnej wskazywały na praktycznie taką samą (niską) częstotliwość palenia wśród uczniów. Podczas 21 miesięcy badań liczba palących wzrosła w obu grupach, choć wzrost ten był mniejszy w grupie eksperymentalnej niż kontrolnej, co sugeruje, że dany program miał wpływ na zachowanie uczniów.

Schematy z uwzględnieniem wielu szeregów czasowych

Czasami ewaluację procesów, zachodzących poza „czystą" kontrolą eksperymentalną, może ułatwić zastosowanie analizy więcej niż jednego szeregu czasowego. Schematy z uwzględnieniem **wielu szeregów czasowych** (*multiple time-series designs*) są ulepszoną wersją właśnie opisanego schematu nierównoważnej grupy kontrolnej. Carol Weiss przedstawiła użyteczny przykład:

> „Interesującym przykładem wielokrotnych szeregów czasowych była ewaluacja programu walki z przekraczaniem prędkości na autostradach w Connecticut. Ewaluatorzy zbierali raporty, dotyczące śmiertelnych wypadków drogo-

wych z kilku okresów przed wprowadzeniem nowego programu i po jego wprowadzeniu. Odkryli, że liczba śmiertelnych wypadków spadła, lecz ponieważ szereg ten wykazywał niestabilny wzorzec spadków i wzrostów, występujący na przestrzeni wielu lat, nie było jasne, czy obecny spadek był związany z realizowanym programem. Następnie badacze porównali te statystyki z danymi z szeregów czasowych z czterech sąsiednich stanów, w których nie nastąpiły zmiany w egzekwowaniu przepisów drogowych. W stanach tych nie zanotowano podobnego spadku liczby wypadków śmiertelnych. To porównanie przydało wiarygodności konkluzji, że nowy program przyniósł pewne skutki".

(1972, s. 69)

Mimo że ten schemat badawczy nie jest równie dobry jak ten, w którym badani przydzielani są losowo, jest on jednak pewnym postępem w stosunku do oceniania osiągnięć grupy eksperymentalnej bez jakichkolwiek porównań. To jest to „coś", co czyni te schematy quasi-eksperymentami, a nie bezproduktywną działalnością. Kluczową sprawą tego aspektu badań ewaluacyjnych jest porównywalność, jak to ilustruje poniższy przykład.

Rozwój obszarów wiejskich, przedmiot rosnącej troski w biednych krajach świata, przyciągnął uwagę i wsparcie wielu bogatych państw. Dzięki narodowym programom wsparcia dla zagranicy oraz poprzez takie międzynarodowe instytucje jak Bank Światowy, kraje rozwinięte uczestniczą w procesie dzielenia się swoją wiedzą i umiejętnościami technologicznymi z krajami rozwijającymi się. Jednak wyniki takich programów są zróżnicowane. Często się zdarza, że nowoczesne technologie nie prowadzą do zamierzonych rezultatów, gdy są zastosowane w społeczeństwach tradycyjnych.

Rajesh Tandon i L. Dave Brown (1981) podjęli eksperyment, w którym szkoleniu, dotyczącemu zastosowania technologii, miało towarzyszyć szkolenie dotyczące organizacji wsi. Badacze byli przekonani, że dla ubogich rolników ważne jest nauczenie się, jak zorganizować i wykorzystać zbiorowy wysiłek w ramach wspólnoty wiejskiej – na przykład dla uzyskania potrzebnych działań ze strony urzędników rządowych. Tylko wtedy ich nowe umiejętności technologiczne mogłyby przynieść owoce.

Zarówno interwencja, jak i jej ewaluacja zostały przeprowadzone w ramach programu realizowanego w 25 wsiach wybranych dla szkoleń z zakresu technologii. Z każdej wsi wybrano dwóch ubogich wieśniaków, którzy zostali przeszkoleni w zakresie nowych technologii w rolnictwie. Następnie zostali odesłani do domów, by dzielić się nowo nabytą wiedzą ze swoją wsią i by organizować pozostałych rolników w zespoły, które będą wspomagać upowszechnianie uzyskanej wiedzy. Dwa lata później autorzy wybrali losowo do specjalistycznych szkoleń dwie wsie z owych 25 (nazywane odtąd grupą A i grupą B), oraz 11 innych nie przeszkolonych grup, jako grupy kontrolne. Dokładne porównanie cech demograficznych wykazało, że grupy eksperymentalne i kontrolne są uderzająco podobne, co oznaczało, że były wystarczająco porównywalne do przeprowadzenia badań.

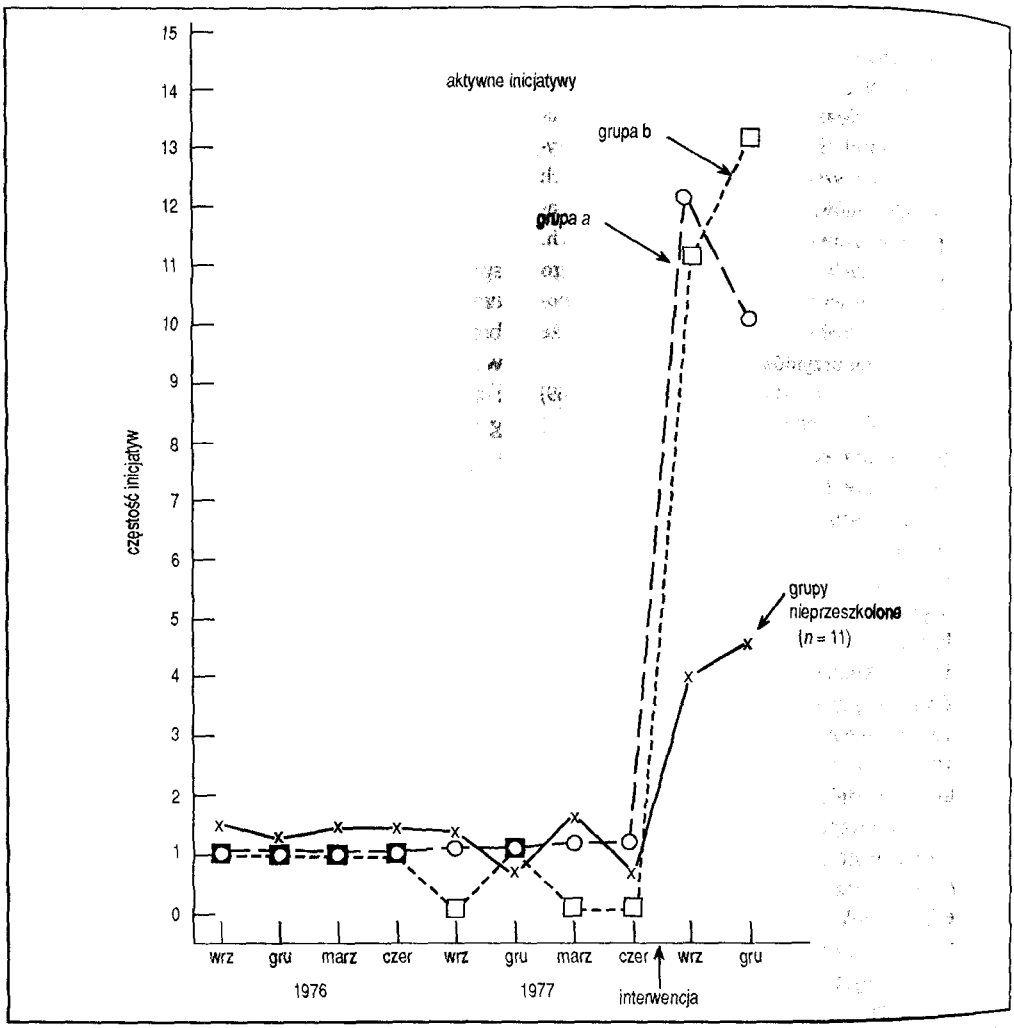
Zespoły rolników z dwóch eksperymentalnych wsi zostały zebrane na specjalnym szkoleniu, dotyczącym zasad tworzenia organizacji. Ich uczestnicy uzyskali pewne informacje o organizowaniu się i zgłaszaniu żądań wobec rządu, a także mieli możliwość odegrania scen podobnych do sytuacji, których doświadczali u siebie na miejscu. Szkolenie trwało trzy dni.

Wszystkie zmienne wynikowe rozważane przez ewaluatorów dotyczyły stopnia, w jakim członkowie zespołów inicjowali działania grupowe, mające na celu poprawę ich sytuacji. Przebadano sześć typów takich działań. Na przykład „aktywna inicjatywa" została zdefiniowana jako „aktywne starania wywarcia wpływu na osoby lub zdarzenia, dotyczące członków grupy, w przeciwieństwie do biernej reakcji lub wycofania" (Tandon, Brown 1981, s. 180). Dane do ewaluacji pochodziły z dzienników, które liderzy zespołów prowadzili od czasu wstępnego szkolenia dotyczącego zastosowania technologii. Badacze czytali dzienniki i liczyli inicjatywy podjęte przez członków zespołów. Dwóch badaczy oddzielnie kodowało informacje z dzienników i porównywało swoją pracę dla sprawdzenia wiarygodności procesu kodowania.

Na rysunku 12.3 porównano liczbę „aktywnych inicjatyw", podejmowanych przez człon-

Rysunek 12.3

Podejmowanie „aktywnych inicjatyw" w czasie



Źródło: Tandon Rajesh, Dave L. Brown, *Organization-Building for Rural Development: An Experiment in India*, „Journal of Applied Behavior Science" kwiecień–czerwiec 1981, s. 182.

ków dwóch grup eksperymentalnych z tymi inicjatywami, które podejmowano w grupach kontrolnych. Podobne wyniki uzyskano dla innych mierników osiągnięć.

Zauważmy dwa wnioski wynikające z tego wykresu. Po pierwsze, istnieje ogromna różnica w liczbie inicjatyw, podejmowanych przez dwie grupy eksperymentalne, w porównaniu z jedenastoma grupami kontrolnymi. Zdaje się to potwierdzać efektywność specjalnego programu szkoleń. Po drugie, liczba inicjatyw wzrosła także wśród grup kontrolnych. Badacze wyjaśniają ten wzorzec jako wynik oddziaływania grup eksperymentalnych. Ponieważ wszystkie wsie były położone blisko siebie, lekcja przyswojona przez członków zespołów z grup eksperymentalnych została częściowo przekazana mieszkańcom wsi z grup kontrolnych.

Ten przykład ilustruje zalety schematów z zastosowaniem wielu szeregów czasowych

w sytuacjach, gdy prawdziwe eksperymenty nie nadają się do ewaluacji danego programu.

Ewaluacja jakościowa

Przedstawiłem kolejne etapy ściśle ustrukturyzowanych, głównie ilościowych, badań ewaluacyjnych, ale ewaluacja może być także mniej ustrukturyzowana i bardziej jakościowa. Na przykład Pauline Bart i Patricia O'Brien (1985) chciały poddać ewaluacji różne sposoby zapobiegania gwałtom. Przeprowadziły więc wywiady pogłębione zarówno z kobietami, które padły ofiarą gwałtu, jak i z takimi, które skutecznie obroniły się przed próbą gwałtu.

Czasami nawet ustrukturyzowana ewaluacja ilościowa może przynieść nieoczekiwane wyniki jakościowe. Paul Steel jest badaczem społecznym, specjalizującym się w ewaluacji programów adresowanych do kobiet w ciąży uzależnionych od narkotyków. Jeden z ewaluowanych przez niego programów obejmował doradztwo, prowadzone przez pielęgniarki publicznej służby zdrowia, które ostrzegały ciężarne kobiety zażywające narkotyki, iż kontynuacja przyjmowania narkotyków może spowodować niedowagę noworodków, których czaszka może być średnio o 10% mniejsza niż normalnie. Jednak podczas wywiadów pogłębionych z uczestniczkami programu Steel odkrył, że w programie tym pominięto jedną istotną informację: że mniejszy wzrost dziecka jest czymś złym. Wiele młodych matek, z którymi przeprowadzono wywiady, sądziło, że mniejszy noworodek będzie oznaczał łatwiejszy poród.

W innym przypadku okręgowy prawnik uruchomił program, który ogólnie mógł być uważany za postępowy i światły. Jeśli zażywająca narkotyki kobieta w ciąży została aresztowana, mogła uniknąć sądu, gdy: (1) zgodziła się zaprzestać używania narkotyków, (2) udało jej się ukończyć program rehabilitacji wychodzenia z uzależnienia. Także i w tym przypadku wywiady pogłębione wykazały, że rzeczywiste funkcjonowanie tego programu niekiedy odbiegało od jego założeń. Przede wszystkim odkryciem Steela było to, iż kiedy jakaś młoda kobieta została aresztowana za używanie narkotyków, współtowarzyszki doradzały jej zajście w ciążę natychmiast, gdy zostanie zwolniona za kaucją. W ten sposób będzie mogła uniknąć sprawy sądowej (przekaz osobisty, 22 listopada 1993 roku).

Najskuteczniejsze badania ewaluacyjne to takie, w których łączą się elementy jakościowe i ilościowe. Przydatne są zarówno porównania statystyczne, jak i pogłębione wyjaśnienia procesów, prowadzących do zaobserwowanych wyników – albo też uniemożliwiających wystąpienie oczekiwanych wyników.

W ewaluacji opery mydlanej w Tanzanii, przedstawionej na początku tego rozdziału, korzystano z kilku technik badawczych. Jak wspomniałem uprzednio, prowadzono tam sondaże wśród słuchaczy i analizowano dane pochodzące z klinik. Dodatkowo badacze przeprowadzili liczne zogniskowane wywiady grupowe (grupy fokusowe) w celu głębszego zbadania wpływu, jaki to słuchowisko miało na odbiorców. Przeprowadzono także analizę treści samych odcinków opery mydlanej oraz licznych listów otrzymanych od słuchaczy. Podjęto zarówno analizy jakościowe, jak i ilościowe (Swalehe i in. 1995).

Kontekst społeczny

Wiele uwag w poprzednich fragmentach zawierało sugestie dotyczące możliwości wystąpienia problemów w czasie faktycznej realizacji ewaluacyjnych projektów badawczych. Oczywiście, wszystkie formy badań mogą prowadzić do wystąpienia problemów, lecz badania ewaluacyjne są na nie szczególnie podatne. W tym podrozdziale przyjrzymy się niektórym problemom logistycznym i specyficznym zagadnieniom etycznym, związanym z badaniami ewaluacyjnymi. Zakończymy go paroma uwagami, dotyczącymi wykorzystania wyników ewaluacji.

Problemy logistyczne

W kontekście wojskowym *logistyka* odnosi się do dostarczania zaopatrzenia – zapewnienia, że ludzie mają żywność, broń i śledzie do namiotów, gdy ich potrzebują. Tutaj używam tego terminu w odniesieniu do nakłaniania ba-

danych, by robili to, czego od nich oczekujemy, do sprawienia, by narzędzia badawcze zostały rozdane i zwrócone z powrotem oraz do innych pozornie mało wymagających zadań. Jednak zadania te wymagają większego wysiłku niż mógłbyś sobie wyobrażać!

Motywowanie marynarzy

Kiedy Kent Crawford, Edmund Thomas i Jeffrey Fink (1980) przystąpili do poszukiwania sposobu motywowania marynarzy o słabych osiągnięciach w Marynarce Wojennej Stanów Zjednoczonych, odkryli, z jak licznymi problemami może przyjść się borykać. Celem tych badań było sprawdzenie potrójnego programu motywowania marynarzy, którzy byli chronicznymi nieudacznikami i często popadali w tarapaty na pokładzie okrętu. Po pierwsze, miały zostać przeprowadzone warsztaty dla pracowników nadzoru, przygotowujące ich do skutecznego dowodzenia marynarzami wykazującymi słabe osiągnięcia. Po drugie, kilku z owych przełożonych miało zostać wybranych i przeszkolonych jako specjalni dowódcy i wzory osobowe – osoby, u których słabsi marynarze mogli poszukiwać porady lub po prostu wsparcia. Na koniec, sami marynarze o słabych osiągnięciach mieliby uczestniczyć w warsztatach, których celem byłoby wzmocnienie ich motywacji i skuteczności w pracy oraz poza nią. Projekt ten miał być realizowany na pokładzie jednego okrętu, z grupą kontrolną dobraną spośród marynarzy z czterech innych jednostek.

Już na wstępie badacze odnotowali, iż personel pomocniczy nie był specjalnie podekscytowany tym programem.

> „Znaczny opór części personelu nadzoru wobec zajmowania się tymi sprawami nie był specjalnym zaskoczeniem. Faktycznie, ich niechęć do przyjęcia własnej odpowiedzialności za problem odzwierciedlała się w przypisywaniu «winy» kilku czynnikom, które mogły przyczyniać się do problemów ich podwładnych. System poboru do wojska, szkolenie rekrutów, rodzice i całe społeczeństwo byli wskazywani jako czynniki wpływające na słabe osiągnięcia – pozostawały one poza kontrolą przełożonych".
>
> (Crawford, Thomas, Fink 1980, s. 488)

Ostatecznie, niechętni przełożeni zostali przekonani, a „owa początkowa niechęć ustępowała ostrożnemu optymizmowi, a później entuzjazmowi" (1980, s. 489).

Jednak z samymi marynarzami był jeszcze większy problem. Plan badań zakładał zarówno pretest, jak i posttest, dotyczący postaw i osobowości, tak aby zmiany powstałe w wyniku programu mogły być zmierzone i ewaluowane.

> „Niestety wszyscy marynarze o słabych osiągnięciach byli stanowczo przeciwni uczestnictwu w tych tak zwanych testach osobowości i z tego powodu doszliśmy do wniosku, że dane zebrane w takich okolicznościach będą wątpliwej wartości. Również etyczne wątpliwości sprawiały, że nie nalegaliśmy na «przetestowanie» marynarzy".
>
> (1980, s. 490)

W konsekwencji badacze musieli zdać się jedynie na wywiady z marynarzami o słabych osiągnięciach i na osąd ich przełożonych jako na mierniki zmiany postaw. Jednak problemy z badanymi trwały dalej.

Początkowo dowództwo statku rozkazało piętnastu słabym marynarzom, aby wzięli udział w tym eksperymencie. Jednak z tych piętnastu jeden trafił do szpitala, innemu przydzielono obowiązki uniemożliwiające mu udział w programie, a trzeci „wyparował" (nieobecny bez zwolnienia). Tak więc gdy eksperyment się rozpoczął, uczestniczyło w nim dwanaście osób, lecz zanim się zakończył trzech kolejnych marynarzy zakończyło służbę w marynarce, a jeszcze jeden został z niej dyscyplinarnie usunięty. Gdy eksperyment dobiegał końca, brało w nim udział osiem osób. Chociaż ewaluacja wskazywała na pozytywne wyniki, bardzo mała liczba przypadków nakazywała ostrożność w formułowaniu jakichkolwiek uogólnień na podstawie tego eksperymentu.

Specyficzne problemy logistyczne badań ewaluacyjnych wyrastają z faktu, iż badania te realizowane są w kontekście realnego życia. Chociaż w badaniach ewaluacyjnych przyjmuje się model eksperymentu, sugerujący kontrolę badacza nad wydarzeniami, ewaluacja zachodzi w toku życia codziennego, często nie dającego się kontrolować. Oczywiście uczestniczący

obserwator w badaniach terenowych również nie ma kontroli nad tym, co jest przedmiotem jego obserwacji, lecz metoda ta nie sili się na uzyskanie takiej kontroli. Wobec zakładanych celów ewaluacji brak kontroli może stwarzać badaczowi rzeczywiste problemy.

Kontrola administracyjna

Jak widać w powyższym przykładzie, logistyczne szczegóły projektu ewaluacyjnego często spadają na administratorów programu. Wyobraź sobie, że poddajesz ewaluacji efekty programu „wizyt małżeńskich" na nastroje wśród żonatych więźniów. Program ten pozwalał współmałżonkom na okresowe odwiedziny u więźniów, podczas których mogli odbywać stosunki seksualne. W czwartym dniu trwania programu jeden z tam osadzonych więźniów przebrał się w strój swojej żony i uciekł. Chociaż możesz przypuszczać, iż jego nastrój zdecydowanie się poprawił w wyniku ucieczki, taki rozwój wypadków na wiele sposobów komplikuje plan twoich badań. Prawdopodobnie dyrektor więzienia zlikwiduje ten program w całości, i co wtedy z jego ewaluacją? Lub, jeśli dyrektor będzie odważny, zażyczy sobie przeglądu akt wszystkich więźniów, których losowo dobrałeś do grupy eksperymentalnej i nie zgodzi się na tych, którzy należą do „grupy wysokiego ryzyka". To uderzy w porównywalność grupy kontrolnej z eksperymentalną. Albo też mogą zostać wprowadzone ostrzejsze środki bezpieczeństwa w celu zapobieżenia kolejnym ucieczkom, a takie środki bezpieczeństwa mogą negatywnie oddziaływać na nastroje wśród więźniów. Tak więc bodziec eksperymentalny zostanie zmieniony w trakcie badań i część danych będzie odzwierciedlać oddziaływanie pierwotnego bodźca, pozostałe zaś odzwierciedlać będą zaistniałą modyfikację. Chociaż zapewne będziesz w stanie oddzielić jedne dane od drugich, twój projekt badawczy stał się logistycznym uściskiem węża.

Lub też wyobraź sobie, że jesteś zaangażowany w ewaluację wpływu wykładów, dotyczących stosunków rasowych, na uprzedzenia w wojsku. Dokładnie badasz dostępnych żołnierzy, losowo dzielisz ich na tych, którzy będą uczestniczyć w wykładach, i na tych, którzy nie będą na nie chodzić. Odpowiednie listy zostały przekazane z kilkutygodniowym wyprzedzeniem i w wyznaczonym dniu i porze wykłady się zaczynają. Wszystko, jak się wydaje, przebiega gładko, aż do momentu rozpoczęcia analizy danych: coś się nie zgadza w zapisanych przez ciebie nazwiskach. Sprawdzając to, dowiadujesz się, że rozpoczynające się manewry, służba w kuchni i liczne nadzwyczajne okoliczności sprawiły, że niektórzy uczestnicy grupy eksperymentalnej są w zupełnie innych miejscach niż na sali wykładowej. Już samo to jest wystarczającym pechem, a tu jeszcze okazuje się, że uczynni oficerowie wysłali innych w zastępstwie brakujących żołnierzy. A jak sądzisz, kogo wybrali na to zastępstwo? Żołnierzy, którzy nie mieli nic innego do roboty, albo którzy cieszą się tak małym zaufaniem, że nie można im powierzyć żadnych istotnych zadań. Może się zdarzyć, że dowiesz się tego wszystkiego mniej więcej tydzień przed ostatecznym terminem złożenia raportu końcowego, dotyczącego wpływu wykładów o stosunkach rasowych.

Takie są niektóre z logistycznych problemów, z którymi mają do czynienia ewaluatorzy. Powinieneś zapoznać się z nimi, aby zrozumieć, dlaczego niektóre procedury badawcze mogą nie dosięgać standardów schematu klasycznego eksperymentu. Jednak czytając raporty z badań ewaluacyjnych, odkryjesz, że – mimo moich wcześniejszych uwag – możliwe jest przeprowadzenie kontrolowanych badań społecznych w połączeniu z eksperymentami, dotyczącymi realnego życia.

Tak jak badania ewaluacyjne mają specyficzne problemy logistyczne, tak również wiążą się one ze specyficznymi problemami etycznymi. Ponieważ problemy etyczne mogą wpływać na naukową jakość badań, powinniśmy pokrótce się im przyjrzeć.

Niektóre zagadnienia etyczne

Etyka i ewaluacja są powiązane na wiele sposobów. Czasami interwencje społeczne poddane ewaluacji powodują powstanie problemów etycznych. Ewaluacja wpływu dowozu dzieci do szkoły autobusami na osiągnięcie integracji edukacyjnej rzuci badacza prosto w wir zagad-

nień politycznych, ideologicznych i etycznych, związanych z samym dowożeniem dzieci autobusami. Nie jest możliwe ewaluowanie programu edukacji seksualnej w szkołach podstawowych bez wdania się w gorącą dyskusję wokół samej edukacji seksualnej. Badaczowi trudno będzie zachować wobec niej bezstronność. Projekt badań ewaluacyjnych będzie *wymagać*, by niektóre dzieci uzyskały edukację seksualną – tak naprawdę to ty możesz być osobą decydującą o tym, które to będą dzieci. (Z naukowego punktu widzenia, *powinieneś* odpowiadać za ten dobór). Oznacza to również, że będziesz osobą bezpośrednio odpowiedzialną przed rodzicami, którzy mogą być oburzeni tym, iż ich dziecko jest uczone o seksie.

Teraz spójrzmy na „jaśniejszą" stronę. Może być tak, że program eksperymentalny ma wielką wartość dla uczestników. Załóżmy, że ewaluowany nowy program, dotyczący bezpieczeństwa pracy, przyczynił się do znaczącego zmniejszenia liczby urazów. Co w takim razie z członkami grupy kontrolnej, których schemat badawczy pozbawił możliwości korzystania z tego programu? Wobec działań ewaluatora mogą zostać wysunięte zarzuty, że są one częściowo przyczyną tego, że członkowie grupy kontrolnej byli narażeni na wypadki.

Czasami pod nazwą badań ewaluacyjnych ukrywa się zachowanie nieetyczne. W rozdziale 9 omawiałem „sondażowe przepychanie" (*push polls*), które rzekomo służyło ocenie wpływu oskarżeń formułowanych w toku kampanii wyborczej, a faktycznie jego celem było rozpowszechnianie złośliwych pomówień. I nie jest to najgorszy z możliwych przykładów.

W 1932 roku badacze w Tuskegee w Alabamie rozpoczęli program zapewniający bez-

Wpływ prawa „trzech razy"

SACRAMENTO (AP) – Autor istniejącego w Kalifornii od pięciu lat prawa „trzech razy" twierdzi, że dzięki temu prawu udało się zapobiec ponad milionowi przestępstw oraz zaoszczędzono 21,7 miliarda dolarów.

W piątek sekretarz stanu Bill Jones przedstawił na konferencji zorganizowanej w Sacramento przez Doris Tate Crimes Victim Bureau swoją interpretację rezultatów prawa „trzech razy".

„BayInsider", 1 marca 1999

W latach 90. byliśmy świadkami przyjęcia na szczeblu federalnym oraz w licznych stanach prawa „trzech razy". Jego intencją było zmniejszenie wskaźnika przestępczości poprzez zamykanie w więzieniach recydywistów. Na przykład na podstawie prawa przyjętego w 1994 roku przez Kalifornię wcześniejszy wyrok skazujący za poważne przestępstwo spowoduje podwojenie wymiaru kary przy kolejnym wyroku za poważne przestępstwo, popełnienie poważnego przestępstwa po raz trzeci spowoduje zaś obligatoryjne orzeczenie kary od 25 lat więzienia do dożywocia. Przez lata tylko Kalifornia stosowała takie prawo z pewnym entuzjazmem.

Zwolennicy przyjęcia prawa „trzech razy", na przykład cytowany powyżej Bill Jones, szybko powiązali gwałtowny spadek wskaźnika przestępczości w latach 90. z nową polityką rozprawiania się z recydywistami. Uznając, że przyjęcie prawa „trzech razy" może nie być jedyną przyczyną spadku przestępczości, Jones dodawał jednak: „gdybyśmy mieli do czynienia z 51-procentowym spadkiem liczby morderstw w ciągu pięciu lat, gwarantowałbym, że prawo «trzech razy» jest tego istotną przyczyną".

Niezależnie od takich gwarancji ze strony polityków, inni obserwatorzy poszukiwali dalszych dowodów na poparcie tezy o wpływie prawa „trzech razy". Niektórzy z jego krytyków zauważyli na przykład, że wskaźnik przestępczości spadał gwałtownie w całym kraju; nie tylko w Kalifornii, ale i w stanach, które nie wprowadziły prawa „trzech razy" bądź też wprowadziły je, lecz nie było ono tam stosowane przez sądy. Faktem jest, że wskaźnik przestępczości spadł w tych hrabstwach stanu Kalifornia, które częściej ignorowały to prawo stanowe. Co więcej, spadek przestępczości w Kalifornii zaczął się, zanim zaczęło obowiązywać prawo „trzech razy".

W 1994 roku Peter Greenwood i jego koledzy z Rand Corporation szacowali, że wprowadzenie tego prawa będzie kosztować kalifornijski wymiar

płatne leczenie syfilisu wśród biednych czarnych mężczyzn cierpiących na tę chorobę. Przez następne lata uczestniczyło w nim kilkuset mężczyzn. Nie wiedzieli jednak, że przez cały ten czas faktycznie nie byli w ogóle leczeni: lekarze przeprowadzający badania chcieli wyłącznie obserwować naturalny rozwój choroby. Nawet gdy okazało się, że penicylina jest skutecznym lekarstwem na tę chorobę, badacze wciąż wstrzymywali się od leczenia. W dzisiejszych czasach panuje powszechna zgoda co do nieetycznej natury tamtych badań, wtedy jednak tak nie było. Nawet gdy rozpoczęto publikację wyników badań, badacze odmawiali uznania, iż zrobili coś złego. Kiedy w 1965 roku środowiska zawodowe złożyły wreszcie skargę do Centrum Kontroli Medycznej (*U.S. Center for Disease Control*), nie było odpowiedzi (J. Jones 1981).

Celem moich uwag nie było rzucenie cienia na badania ewaluacyjne. Chciałem raczej przybliżyć życiowe konsekwencje działań, podejmowanych przez badaczy ewaluatorów. W końcu wszystkie badania społeczne dotykają kwestii etycznych. Wrócimy do tego tematu w rozdziale 18.

Wykorzystanie wyników badań

Jeden z życiowych aspektów badań ewaluacyjnych dotyczy problemu wykorzystania wyników badań. Ponieważ celem ewaluacji jest określenie sukcesu lub porażki interwencji społecznych, można by sądzić, że dany program powinien automatycznie być kontynuowany lub wstrzymany na podstawie wyników badań.

Jednak rzeczywistość nie jest tak prosta i racjonalna. Inne czynniki zakłócają ocenę wyników badań ewaluacyjnych, czasami otwarcie,

sprawiedliwości około 5,5 miliarda dolarów rocznie, zwłaszcza z powodu kosztów osadzenia w więzieniach recydywistów, którym zasądzono długoletnie wyroki. Grupa z Rand nie zaprzeczała, że prawodawstwo „trzech razy" mogłoby mieć *pewien* wpływ na przestępczość – ci, którzy odsiadują wieloletnie wyroki, nie mogą dopuszczać się przestępstw na ulicach miast. Natomiast kolejne badania (Greenwood, Rydell, Model 1996) wskazywały, że było ono nieskuteczne jako sposób zwalczania przestępczości. Badacze oszacowali, że każdy milion dolarów, wydany w związku z ustanowionym prawem „trzech razy", przyczyni się do zapobieżenia 60 przestępstwom, natomiast taka sama suma pieniędzy, wydana na programy zachęcania młodzieży szkół średnich do dalszych studiów, przyczyni się do zapobieżenia 258 przestępstwom.

Kryminolodzy już dawno zauważyli, że sprawcami większości przestępstw są młodzi mężczyźni. Koncentracja uwagi na „starej recydywie" ma mały wpływ na młodocianych przestępców lub nie ma na nich żadnego wpływu. Właściwie wyroki orzekane w ramach prawa „trzech razy" nieproporcjonalnie mocno uderzały w tych, którzy dobiegali końca swojej przestępczej kariery z powodu starzenia się.

John Irwin i James Austin (1997) w swojej bardziej ogólnej krytyce zasugerowali, że wśród mieszkańców Stanów Zjednoczonych można zaobserwować skłonność do nadużywania więzienia jako rozwiązania problemu przestępczości, ignorując inne, bardziej skuteczne rozwiązania. Często osadzenie kogoś w więzieniu stwarza poważniejsze problemy, niż te, na które miało być lekarstwem.

Jednak podobnie jak w przypadku wielu innych interwencji społecznych, znaczna część poparcia dla prawa „trzech razy" w Kalifornii i gdzie indziej dotyczy przede wszystkim emocji społecznych, wywoływanych przez przestępczość, a także politycznych skutków takich emocji. Tak więc badania ewaluacyjne, dotyczące wprowadzenia takiego prawa, mogą ostatecznie spowodować zmiany, lecz wpływ tych badań będzie znacznie wolniejszy, niż można by to logicznie przewidywać.

Źródła: *State Saved $21.7 Billion with Five-Year-Old „Three Strikes" Law*, „BayInsider" 1 marca 1999 [publikacja w wersji elektronicznej, dostępna na stronie: http://www.bayinsider.com/news/1990/03/01/three_strikes.html]; Peter W. Greenwood i in., *Three Strikes and You're Out: Estimated Benefits and Costs of California's New Mandatory Sentencing Law*, Rand Corporation, Santa Monica, Cal. 1994; Peter W. Greenwood, C. Peter Rydell, Karyn Model, *Diverting Children from a Life of Crime: Measuring Costs and Benefits*, Rand Corporation, Santa Monica, Cal. 1996; John Irwin, James Austin, *It's About Time: America's Imprisonment Binge*, Wadsworth, Belmont, Cal. 1997.

a czasami w sposób trudno uchwytny. Richard Nixon, będąc prezydentem, powołał narodową komisję do zbadania skutków pornografii. Po starannej, wielostronnej ewaluacji komisja przedstawiła raport, w którym stwierdzano, że pornografia, jak się wydaje, nie ma żadnych negatywnych konsekwencji, jakie zwyczajowo są jej przypisywane. Na przykład prezentowanie materiałów pornograficznych nie zwiększa prawdopodobieństwa przestępstw seksualnych. W następstwie badań można było oczekiwać liberalizacji prawodawstwa. Jednak prezydent stwierdził, iż komisja była w błędzie.

Mniej dramatyczne przykłady nieuwzględniania wniosków z badań ewaluacyjnych można wymieniać bez końca. Bez wątpienia każdy ewaluator może wskazać badania, które realizował – badania dostarczające jednoznacznych wyników i oczywistych wniosków dla działań społecznych – które zostały zignorowane, tak jak to ilustruje przykład w ramce zatytułowanej „Wpływ prawa «trzech razy»".

Istnieją trzy ważne powody, dla których wnioski z wyników badań ewaluacyjnych nie zawsze są uwzględniane w praktyce. Po pierwsze, wnioski te nie zawsze są prezentowane w sposób zrozumiały dla nie-badaczy. Po drugie, wyniki badań ewaluacyjnych niekiedy przeczą głęboko zakorzenionym przekonaniom. Tak było z pewnością w przypadku komisji do spraw pornografii. Jeśli „wszyscy wiedzą", że pornografia prowadzi do wszelkiego rodzaju zboczeń seksualnych, to bezpośrednie oddziaływanie wyników badań sprzecznych z tym przekonaniem będzie niewielkie. Na tej samej zasadzie ludzie myśleli, że Kopernik zwariował, gdy twierdził, iż Ziemia obraca się wokół Słońca. Każdy mógł przecież stwierdzić, że Ziemia się nie rusza. Trzecią barierą dla wykorzystania wyników ewaluacji jest interes osób zaangażowanych w realizację programu. Gdybym opracował nowy program resocjalizacji, który – o czym jestem przekonany – powstrzymałby byłych skazańców przed powrotem do więzienia, i który ludzie zaczęliby nazywać „planem Babbiego", jak myślisz, jak bym się czuł, gdyby się okazało, że wyniki ewaluacji wskazują, iż program ten nie działa? Powinienem przeprosić za wprowadzenie ludzi w błąd, zwinąć swój namiot i ruszyć na inny front pracy? Bardziej prawdopodobne jest to, że uznam wyniki ewaluacji za bezwartościowe i rozpocznę intensywny lobbing wśród odpowiednich władz na rzecz zapewnienia kontynuacji mojego programu.

We wcześniejszym przykładzie ewaluacji programu kształcenia kierowców Philip Hilts przedstawił kilka reakcji na wstępne rezultaty badań:

> „Ray Burneson, specjalista do spraw bezpieczeństwa ruchu drogowego z Rady Bezpieczeństwa Narodowego, skrytykował te badania, mówiąc, że były one wytworem grupy osób, wywodzących się z NHTSA (National Highway and Transportation Safety Administration), kierowanej przez ludzi, którzy wierzyli, że «nie można nic zrobić, by wykształcić kierowców. Można tylko ulepszyć opiekę medyczną i budować mocniejsze samochody na okoliczność wypadków (...). To uderza w całą filozofię edukacji»".
>
> (1981, s. 4)

Ze swej natury badania ewaluacyjne przeprowadzane są pośród wydarzeń realnego życia codziennego, wpływają na nie i same pozostają pod ich wpływem. Oto kolejny przykład dobrze znany badaczom społecznym.

Reforma prawodawstwa dotyczącego gwałtów

Od lat wielu badaczy społecznych i innych osób zauważa pewien problem ze stawianiem oskarżeń w sprawach o gwałt. Zbyt często, jak się uważa, ofiara cierpi w końcu prawie tak samo, stając przed sądem jako świadek, jak wtedy, gdy zdarzył się gwałt. Często obrońcy oskarżonego przedstawiają ofiarę jako osobę, która zachęcała do aktu seksualnego, oraz jako osobę o podejrzanej moralności. Inne ataki personalne mają na celu zdjęcie z oskarżonego odpowiedzialności za gwałt.

Tego rodzaju krytyka spowodowała przyjęcie w stanowych systemach prawnych różnych rozwiązań w celu zaradzenia tym problemom. Cassie Spohn i Julie Horney (1990) zainteresowały się prześledzeniem oddziaływania tych rozwiązań. Badaczki podsumowują, w jaki sposób nowe prawa miały zmienić sytuację:

„Najczęstsze zmiany to: (1) przedefiniowanie gwałtu i zastąpienie pojedynczego przestępstwa zestawem stopniowalnych napaści, definiowanych poprzez obecność lub niewystąpienie dokuczliwych okoliczności; (2) zmiana standardu zgody na stosunek seksualny poprzez wyeliminowanie wymogu zaistnienia fizycznego oporu ofiary wobec atakującego; (3) eliminacja konieczności potwierdzania zeznań ofiary; (4) nałożenie ograniczeń na załączanie dowodów, dotyczących wcześniejszego życia seksualnego ofiary".

(1990, s. 2)

Powszechnie oczekiwano, że taka legislacja powinna zachęcić kobiety do zeznawania w sprawie gwałtów i powinna zwiększyć liczbę wyroków skazujących w sprawach wniesionych do sądu. Dla sprawdzenia tego ostatniego oczekiwania badaczki skoncentrowały się na okresie od 1970 do 1985 roku w hrabstwie Cook w stanie Illinois: „Nasz zbiór danych obejmuje 4628 spraw o gwałt, 405 spraw o dewiacyjne napaści seksualne, 745 spraw o dokuczliwe kryminalne napaści seksualne, 37 spraw o kryminalne napaści seksualne" (1990, s. 4). Tabela 12.1 przedstawia niektóre odkrycia badaczek.

Spohn i Horney podsumowały te wnioski w następujący sposób:

„Jedynym znaczącym efektem, ujawnionym w trakcie naszych analiz, był wzrost średniej długości maksymalnej kary więzienia: odnotowano wzrost o prawie 48 miesięcy w przypadku gwałtów i prawie o 36 miesięcy w przypadku napastowania seksualnego. Ponieważ rozkład danych wskazuje na wzrost średniej kary, zanim reforma weszła w życie, przemodelowałyśmy szereg, umieszczając interwencję o rok wcześniej w stosunku do faktycznego jej wprowadzenia. Skala tego efektu była nawet większa i nadal znacząca, co wskazuje na to, że efekt zmiany nie powinien być wiązany z reformą prawa".

(1990, s. 10)

Zauważ, że w tabeli 12.1 właściwie nie widać procentowej zmiany dotyczącej spraw zakończonych wyrokiem skazującym za gwałt lub jakieś inne oskarżenia (na przykład napaść). Zatem zmiany w prawie nie oddziałały na prawdopodobieństwo skazania. Jak zauważają badaczki, jedyna zmiana, jaka *jest* ewidentna – zwiększenie długości orzekanych wyroków – nie może być przypisana samej reformie prawa.

Poza analizą istniejących danych statystycznych, Spohn i Horney przeprowadziły wywiady z sędziami i prawnikami w celu określenia, co sądzą oni o oddziaływaniu nowego prawa. Ich odpowiedzi były nieco bardziej zachęcające:

„W Chicago sędziowie, oskarżyciele i obrońcy podkreślali, że sprawy o gwałt są traktowane poważniej, a ofiary gwałtów są traktowane bardziej humanitarnie w wyniku zmian legislacyjnych. Ten efekt edukacyjny jest oczywiście ważny i powinien zadowolić rzeczników reformy prawodawstwa dotyczącego gwałtów".

(1990, s. 17)

Tabela 12.1

Analiza spraw o gwałt przed wprowadzeniem nowego ustawodawstwa i po jego wprowadzeniu

	Gwałty	
	przed (N = 2252)	po (N = 2369)
Wynik sprawy:		
Skazanie na podstawie pierwotnego oskarżenia	45,8%	45,4%
Skazanie na podstawie innego oskarżenia	20,6%	19,4%
Brak wyroku skazującego	33,6%	35,1%
Mediana długości wyroku więzienia w miesiącach:		
Dla skazanych na podstawie pierwotnego oskarżenia	96,0%	144,0%
Dla skazanych na podstawie innego oskarżenia	36,0%	36,0%

Tak więc badania ujawniły inne efekty prawa, poza wynikami jakościowymi, których poszukiwały badaczki. Badania te ukazują znaczenie badania interwencji społecznych w celu określenia, czy, w jaki sposób i w jakim stopniu interwencje takie osiągają zakładane rezultaty.

Zapobieganie przemocy w rodzinie

W badaniach pod pewnymi względami podobnych do poprzednich badacze w Indianapolis skoncentrowali się na problemie przemocy w rodzinie. Szczególne interesowali się tym, czy oskarżenie osoby maltretującej żonę może prowadzić do dalszej przemocy. David Ford i Mary Jean Regoli (1992) zabrali się za przestudiowanie konsekwencji różnych opcji oskarżenia, dopuszczonych w ramach „Eksperymentu dotyczącego powództwa w Indianapolis" (*Indianapolis Prosecution Experiment – IPE*).

Przypadki maltretowania żony mogą wykazywać różne prawidłowości, jak podsumowują Ford i Regoli:

> „Po ataku na kobietę ktoś może wezwać policję na miejsce zdarzenia albo może tego nie robić. Jeśli policja przybędzie, powinna zbadać dowody, stanowiące podstawę tymczasowego aresztowania. Jeśli dowód taki istnieje, policjanci mogą dokonać aresztowania według swojego uznania. Dokonując takiego aresztowania w miejscu zdarzenia, policjanci spisują prawdopodobny przebieg wydarzeń i kierują podejrzanego do sądu na wstępne przesłuchanie. Jeśli policjanci nie zostali wezwani, lub też po wezwaniu nie dokonali aresztowania, ofiara może sama wnieść oskarżenie, zwracając się do biura prokuratora, i pod przysięgą złożyć zeznania przeciw danemu mężczyźnie, dotyczące prawdopodobnego przebiegu zdarzeń. W następstwie decyzji sędziego podejrzany o maltretowanie może zostać albo wezwany do sądu, albo może zostać zatrzymany na podstawie nakazu aresztowania i doprowadzony do sądu na wstępne przesłuchanie".
>
> (1992, s. 184)

Co się dzieje, jeśli żona wnosi oskarżenie przeciwko swojemu mężowi, a później się rozmyśla? W wielu sądach prowadzona jest polityka zakazująca takich działań, oparta na przekonaniu, iż służy to interesowi ofiary poprzez nacisk na doprowadzenie danej sprawy do końca. Jednak podczas eksperymentu IPE niektórym ofiarom zaproponowano możliwość odstąpienia od oskarżenia, jeśli tak zadecydują w dalszym przebiegu procesu. Co więcej, sąd oferował kilka innych opcji. Ponieważ maltretowanie żon jest w dużym stopniu funkcją seksizmu, stresu i niemożności radzenia sobie ze złością, kilka z nowatorskich rozwiązań, zaproponowanych w ramach eksperymentu IPE, dotyczyło zajęć edukacyjnych z poradnictwem w zakresie kontroli agresji.

Jeśli pozwany przyznaje się do winy i wyraża chęć uczestnictwa w programie poradnictwa w zakresie kontrolowania złości, sędzia, mając to na uwadze, może opóźnić rozprawę, a następnie – jeśli pozwany skutecznie ukończy program – może uwolnić go od zarzutów. Lub też pozwany może zostać osądzony – jeśli będzie winny, kara zostanie zawieszona pod warunkiem uczestnictwa w programie kontroli złości. I wreszcie, pozwany może zostać osądzony, a gdy okaże się winny, zastosowana zostanie konwencjonalna kara, na przykład uwięzienie.

Która z tych możliwości najskuteczniej zapobiega kolejnym aktom przemocy w rodzinie? Oto pytanie postawione przez Forda i Regoli. Poniżej przedstawiam kilka sformułowanych przez nich wniosków.

Po pierwsze, badania wykazują, że mężczyźni, którzy zostali doprowadzeni do sądu na przesłuchanie, rzadziej powtarzają akty przemocy wobec swoich żon, bez względu na to, jaki będzie wynik przesłuchania. Po prostu sam kontakt z wymiarem sprawiedliwości wywiera na nich wpływ.

Po drugie, kobiety, które mają prawo do późniejszego wycofania oskarżenia, są w mniejszym stopniu narażone na kolejne akty przemocy niż kobiety, które takiego prawa nie mają. Szczególnie łączna polityka aresztowania pozwanych na podstawie nakazu sądowego i umożliwienia ofiarom wycofania oskarżeń zapewnia im lepsze zabezpieczenie przed kolejnymi aktami przemocy niż jakakolwiek inna procedura sądowa (Ford, Regoli 1992).

Jednak przyznanie ofiarom prawa do wycofania oskarżeń ma pewne zdumiewające konsekwencje. Kobiety, które korzystają z tego prawa, są w większym stopniu narażone na akty przemocy w późniejszym okresie, niż te kobiety, które nalegają na doprowadzenie procedury sądowej do końca. Według interpretacji badaczy wyniki te pokazują, że występowanie kolejnych aktów przemocy spada, gdy poczucie kontroli ze strony ofiar jest wspierane przez jasne i spójne współdziałanie z instytucjami wymiaru sprawiedliwości.

> „Zdecydowana systemowa reakcja na wszelkie naruszenia zasad warunkowego zwolnienia pozwanych, oczywiście włącznie z przypadkami nowych aktów przemocy, powinna służyć pokazaniu, że współpraca między ofiarą a sądem jest ścisła. Przemawia to do oskarżonego, gdyż pokazuje, że ofiara jest zdecydowana ukrócić wszelką przemoc, oraz że system prawny jest niewzruszony we wspieraniu jej dążenia do zapewnienia jej ochrony".
>
> (Ford, Regoli 1992, s. 204)

Skuteczność poradnictwa w zakresie kontrolowania złości nie może być oceniona w sposób prosty. Polityka, mająca na celu skierowanie oskarżonych do specjalistów od kontrolowania złości, wydaje się stosunkowo nieskuteczna w zapobieganiu nowym aktom przemocy. Jednak badacze zauważają, że nie powinno się mylić efektów takiej polityki z faktycznymi rezultatami poradnictwa. Niektórzy pozwani, zapisani na szkolenia, nigdy nie wzięli w nich udziału. Dla poprawnej ewaluacji niezbędne są bardziej szczegółowe informacje na temat wdrażania doradztwa.

Ponadto badacze przestrzegają, iż rezultaty ich badań pokazują tylko ogólne prawidłowości, a maltretowane żony muszą same wybierać kierunek działań odpowiedni do ich konkretnej sytuacji, a nie działać ślepo na podstawie ogólnych wzorców. Badania te są więc prawdopodobnie bardziej użyteczne w części mówiącej o organizacji systemu sądownictwa (na przykład tam gdzie dopuszcza się prawo ofiary do wycofania skargi), niż jako przewodnik dla poszczególnych ofiar przemocy.

Na koniec, program IPE oferuje przykład powszechnego problemu występującego w badaniach ewaluacyjnych. Często rzeczywista praktyka różni się od tego, czego można by oczekiwać na podstawie zasad programu. Na przykład badacze rozważali wpływ różnych, alternatywnych sposobów doprowadzenia podejrzanych do sądu. Mianowicie sąd może albo wezwać męża do stawienia się w sądzie, albo wydać nakaz aresztowania. Badacze zastanawiali się, czy aresztowanie męża może zwiększyć jego złość spowodowaną całą sytuacją. Byli więc nieco zaskoczeni, gdy nie odkryto różnic w poziomie gniewu mężów wzywanych do sądu i aresztowanych.

Rozwiązanie tej łamigłówki leży w rozdźwięku między zasadami a praktyką:

> „Chociaż nakaz aresztowania powinien być z zasady równie dotkliwy, jak aresztowanie w miejscu popełnienia aktu przemocy, w praktyce może on się niewiele różnić od wezwania do sądu. Mężczyzna zazwyczaj wie o nakazie aresztowania i często decyduje sam oddać się w ręce wymiaru sprawiedliwości w dogodnym dla niego momencie, albo też instytucja realizująca nakaz aresztowania kontaktuje się z nim i zachęca, by sam to uczynił. Tak więc może on nie doświadczyć oczywistych przykrych dodatkowych aspektów kary, związanych z aresztowaniem, zakładaniem kajdanek i wyprowadzaniem z miejsca pracy".
>
> (Ford 1989, s. 9–10)

Podsumowując, wiele czynników poza wartością naukową badań ewaluacyjnych wpływa na to, jak wykorzystuje się wyniki ewaluacji. Jak widzieliśmy wcześniej, czynniki pozostające poza kontrolą badacza mogą wpływać na jakość samych badań. Lecz to „zanieczyszczenie" jest równoważone przez potencjalny wkład badań ewaluacyjnych w poprawę życia ludzi.

Badania wskaźników społecznych

Chciałbym zakończyć ten rozdział omówieniem pewnego typu badań, który łączy to, czego do wiedzieliśmy się o badaniach ewaluacyjnych

i o analizie istniejących danych statystycznych. Ten szybko się rozwijający obszar badań społecznych dotyczy opracowywania i monitorowania **wskaźników społecznych**, czyli zagregowanych statystyk, odzwierciedlających kondycję społeczeństwa lub jakichś grup społecznych. Badacze wykorzystują wskaźniki społeczne do monitorowania różnych aspektów życia społecznego w podobny sposób, jak ekonomiści korzystają z indeksów, na przykład z produktu narodowego brutto (PNB) na jednego mieszkańca jako wskaźnika rozwoju gospodarczego danego kraju.

Załóżmy, że zamierzamy porównać względną kondycję zdrowotną różnych społeczeństw. Jedną ze strategii może być porównanie wskaźników śmiertelności (liczby zgonów na 1000 mieszkańców). Lub też konkretniej, możemy przyjrzeć się śmiertelności niemowląt – liczbie niemowląt umierających w pierwszym roku życia, przypadającej na 1000 urodzeń. Zależnie od tego, którym aspektem zdrowia społeczeństwa będziemy się interesować, możemy opracować dowolną liczbę innych mierników: liczba lekarzy na jednego mieszkańca, liczba łóżek szpitalnych na jednego mieszkańca lub liczba dni pobytu w szpitalu na jednego mieszkańca itd. Zauważmy, że porównania między społeczeństwami ułatwione są przez przeliczanie wskaźników na jednego mieszkańca (czyli dzielenie przez wielkość populacji).

Zanim przejdziemy dalej, przypomnijmy sobie problemy, związane z wykorzystaniem istniejących danych statystycznych, omówione w rozdziale 11. Krótko mówiąc, są one często nierzetelne – odzwierciedlają sposób gromadzenia danych, ich przechowywania i obliczania. Mając to w pamięci, przyjrzyjmy się niektórym sposobom możliwego wykorzystania wskaźników społecznych w badaniach ewaluacyjnych na dużą skalę.

Kara śmierci i odstraszanie

Czy kara śmierci odstrasza od popełnienia takich zbrodni, jak morderstwo? To pytanie jest przedmiotem zażartej debaty, pojawiającej się zawsze, gdy jakiś stan rozważa zniesienie kary śmierci lub jej przywrócenie, oraz zawsze, gdy dochodzi do jej wykonania. Zwolennicy kary śmierci często argumentują, że groźba jej wykonania będzie powstrzymywać potencjalnych morderców od zabijania innych ludzi. Przeciwnicy kary śmierci często argumentują, że kara ta nie ma takiego oddziaływania. Wskaźniki społeczne mogą rzucić nieco światła na ten problem.

Jeśli kara śmierci rzeczywiście powstrzymuje ludzi od popełnienia morderstwa, to moglibyśmy oczekiwać niższego wskaźnika morderstw w tych stanach, w których obowiązuje kara śmierci, w stosunku do tych stanów, w których ona nie obowiązuje. Odpowiednie porównania, dotyczące tej kwestii, są nie tylko możliwe, ale są one zestawiane i publikowane. Tabela 12.2 przedstawia dane, zestawione przez Williama Baileya (1975), które bezpośrednio przeczą poglądowi, jakoby kara śmierci powstrzymywała morderców. Zarówno w roku 1967, jak i w 1968 wskaźnik morderstw w stanach, w których obowiązywała kara śmierci, był znacznie wyższy niż w tych, w których nie obowiązywała. Pojawili

Tabela 12.2

Średni wskaźnik morderstw pierwszego i drugiego stopnia na 100 000 mieszkańców w stanach, w których nie obowiązuje kara śmierci i w których ona obowiązuje

	Stany, w których nie obowiązuje kara śmierci		Stany, w których obowiązuje kara śmierci	
	1967	1968	1967	1968
Morderstwa 1 stopnia	0,18	0,21	1,47	1,58
Morderstwa 2 stopnia	0,30	0,43	1,92	1,03
Morderstwa łącznie	0,48	0,64	1,38	1,59

Źródło: Według opracowania **William** C. Baileya, ***Murder*** *and Capital Punishment*, w: *Criminal Law in Action*, William J. Chambliss (red.), **John** Wiley & Sons **Inc.** 1975. Wykorzystano za pozwoleniem.

się krytycy interpretacji dokonanej przez Baileya, mówiący, że choć niektóre stany mają zapisaną w swoim prawie karę śmierci, to nie była ona od lat wykonywana. Może to wyjaśniać, dlaczego, jak się wydaje, nie działa ona odstraszająco. Jednak dalsza analiza przeczy takiej interpretacji. Kiedy Bailey porównał te stany, które nie korzystały z istniejącej możliwości wykonywania kary śmierci, z tymi, w których ją wykonywano, nie znalazł żadnej istotnej różnicy we wskaźnikach morderstw.

Możliwa jest jednak kolejna kontrinterpretacja. W przypadku interpretacji danych, dokonanej przez Baileya, mogło się zdarzyć, że była „odczytana od końca". Może istnienie kary śmierci, jako jednej z opcji, było konsekwencją wysokiego wskaźnika morderstw: stany z wysokim wskaźnikiem wprowadziły ją do swojego prawodawstwa; tam gdzie wskaźnik był niski, nie wprowadzano kary śmierci lub ją wycofano, jeśli była przewidziana w prawodawstwie. Mogło być zatem tak, że wprowadzenie kary śmierci spowodowało obniżenie wskaźnika morderstw, jej zniesienie zaś spowodowało wzrost liczby morderstw, prowadząc w efekcie – przy dużym stopniu agregacji – do uzyskania takich danych, jak przedstawione w tabeli 12.2. Jednak tak nie jest. Analizy w czasie nie wskazują ani na wzrost wskaźnika morderstw w stanach, które zlikwidowały karę śmierci, ani na jego spadek tam, gdzie karę śmierci wprowadzono. Nowsze badania Baileya i Ruth D. Peterson (1994) potwierdzają wcześniejsze wyniki i wskazują także, że urzędnicy wymiaru sprawiedliwości wątpią w istnienie efektu odstraszania.

Zauważ na podstawie powyższego omówienia, że można wykorzystywać dane, dotyczące wskaźników społecznych, dla porównań różnych grup w danym momencie lub też w różnych okresach. Niejednokrotnie przeprowadzenie obu rodzajów porównań rzuca więcej światła na badany problem.

W chwili obecnej praca nad wykorzystaniem wskaźników społecznych przebiega na dwóch frontach. Badacze opracowują coraz bardziej wyrafinowane wskaźniki, badając, które wskaźniki danej ogólnej zmiennej są najbardziej użyteczne dla monitorowania życia społecznego. Jednocześnie prowadzone są badania, mające doprowadzić do odkrycia związków między zmiennymi dotyczącymi całych społeczeństw.

Podobnie jak z wieloma innymi aspektami życia społecznego, tak i w tym przypadku Internet stał się wartościowym źródłem informacji. Aby lepiej poznać możliwości wykorzystania wskaźników społecznych, sprawdź na przykład strony internetowe Sociometrics Corporation (http://www.socio.com/). Lub też, wykorzystując przeglądarki internetowe, wyszukaj hasło: „social indicators" (wskaźniki społeczne)[2].

Symulacja komputerowa

Jedną z bardziej ekscytujących perspektyw badań z wykorzystaniem wskaźników społecznych jest dziedzina symulacji komputerowych. Gdy badacze przystępują do opracowywania równań matematycznych, opisujących związki łączące między sobą zmienne społeczne (na przykład związki między wzrostem liczby ludności a liczbą samochodów), równania te mogą być zapisywane i łączone ze sobą za pomocą komputera. Dysponując wystarczającą liczbą odpowiednio trafnych równań, badacze będą kiedyś w stanie przetestować implikacje określonych zmian społecznych za pomocą komputera zamiast w realnym życiu.

Załóżmy na przykład, że jakiś stan rozważa możliwość podwojenia wielkości przemysłu turystycznego. Możemy wprowadzić tę propozycję do modelu symulacji komputerowej i w ciągu kilku sekund lub minut otrzymać opis wszystkich bezpośrednich i pośrednich konsekwencji rozwoju turystyki. Możemy dowiedzieć się, jakie inwestycje publiczne będą niezbędne; które służby publiczne, takie jak policja i straż pożarna wymagać będą rozbudowania i w jakim stopniu; jak będzie wyglądać rynek pracy; jakie szkolenia trzeba będzie przeprowadzić dla przyszłych pracowników; jak zwiększą się podatki i dochody – itd., poprzez wszelkie zamierzone i niezamierzone konsekwencje planowanych działań. Osoby planujące te działania, zależnie od otrzymanych wyników,

[2] W Polsce liczne wskaźniki społeczne można znaleźć w wydawnictwach Głównego Urzędu Statystycznego, a także na jego stronach internetowych (www.stat.gov.pl). Inne adresy polskich stron internetowych, zawierających tego typu wskaźniki, podane są w załączniku B (przyp. red. nauk.).

mogą powiedzieć: „załóżmy wzrost rozwoju przemysłu turystycznego tylko o połowę", i natychmiast będzie można otrzymać nowy zestaw skutków.

Doskonałą ilustrację komputerowej symulacji, łączącej zmienne społeczne i fizyczne, można znaleźć w badaniach Donelli i Denisa Meadowsów oraz ich kolegów z Dartmouth i z Massachusetts Institute of Technology. Wzięli oni jako dane wejściowe znane i szacowane rezerwy różnych nieodnawialnych surowców naturalnych (takich jak ropa naftowa, węgiel, żelazo) oraz występujące w przeszłości wzorce wzrostu liczby ludności i wzrostu gospodarczego, a także zależności między rozwojem a zużyciem zasobów. Stosując złożony model symulacji komputerowej, byli oni w stanie oszacować między innymi, na ile lat wystarczą różne zasoby, z uwzględnieniem odmiennych wzorców korzystania z nich w przyszłości. Wychodząc poza początkowo ponure przewidywania, modele takie dają możliwość zaplanowania mniej ponurej przyszłości poprzez określenie działań niezbędnych do osiągnięcia zakładanych celów. Oczywiście wartość symulacji komputerowych nie ogranicza się do badań ewaluacyjnych, choć mogą spełniać w nich istotną funkcję.

Ta możliwość wskazuje na szczególną wartość badań ewaluacyjnych w ogóle. W ciągu całej historii ludzkości eksperymentowaliśmy z naszymi urządzeniami społecznymi, dążąc do coraz lepszych rezultatów. Badania ewaluacyjne dostarczają nam narzędzi do natychmiastowej odpowiedzi na pytanie, czy konkretny eksperyment rzeczywiście sprawia, że świat staje się lepszy. Wskaźniki społeczne pozwalają nam na ustalenie tego w wielkiej skali, a łączenie ich z symulacjami komputerowymi otwiera przed nami możliwość sprawdzenia, jak będzie wyglądać konkretna interwencja bez konieczności doświadczania ryzyka jej realizacji.

Główne zagadnienia

- Badania ewaluacyjne są formą stosowanych badań społecznych, w których przedmiotem zainteresowania są skutki interwencji społecznych.
- Staranne sformułowanie problemu, wraz z odpowiednimi pomiarami i kryteriami sukcesu lub porażki, to podstawa badań ewaluacyjnych. W szczególności ewaluator musi starannie określić wyniki, dokonać pomiaru kontekstu eksperymentu, określić badaną interwencję i grupę docelową, do której jest skierowana, oraz zadecydować, czy korzystać z istniejących mierników, czy opracować nowe.
- Ewaluatorzy korzystają na ogół ze schematów eksperymentalnych lub quasi-eksperymentalnych. Przykłady schematów quasi-eksperymentalnych obejmują badania szeregów czasowych oraz wykorzystanie nierównoważnych grup kontrolnych.
- Ewaluatorzy mogą także korzystać z jakościowych metod zbierania danych. Zarówno analiza danych ilościowych, jak i analiza danych jakościowych może być odpowiednia dla badań ewaluacyjnych, niejednokrotnie w ramach jednego projektu badawczego.
- Realizacja badań ewaluacyjnych pociąga za sobą specyficzne problemy logistyczne i etyczne, ponieważ dotyczy codziennych wydarzeń realnego życia.
- Wnioski z badań ewaluacyjnych niekoniecznie muszą być zastosowane w praktyce, zwłaszcza gdy są sprzeczne z oficjalnymi poglądami.
- Wskaźniki społeczne umożliwiają zrozumienie szerszych procesów społecznych.
- Modele stworzone w symulacjach komputerowych dają badaczom szansę studiowania możliwych rezultatów interwencji społecznych bez ryzyka ponoszenia ich konsekwencji w realnym życiu.

Kluczowe pojęcia

badania ewaluacyjne
quasi-eksperymenty
schematy z uwzględnieniem szeregów czasowych
nierównoważna grupa kontrolna
schematy z uwzględnieniem wielu szeregów czasowych
wskaźniki społeczne

Pytania i ćwiczenia

1. Załóżmy, że społeczność lokalna organizuje młodzieżowe centrum, w którym obowiązuje abstynencja od alkoholu i narkotyków, jako sposób ograniczenia sięgania po narkotyki i alkohol przez nastolatków. Opisz, jak można by przystąpić do oceny skuteczności takiego centrum. Określ, czy twój projekt będzie eksperymentalny, quasi-eksperymentalny, czy jakościowy (lub będzie pewną kombinacją tych schematów).
2. Wróć do ewaluacji programu motywowania marynarzy w Marynarce Wojennej Stanów Zjednoczonycj opisanej w tym rozdziale. Zaplanuj od nowa ten program i jego ewaluację tak, by poradzić sobie z problemami, jakie pojawiły się podczas tamtych badań.
3. Omów niektóre problemy polityczne i etyczne, jakie mogą być związane z badaniami opisanymi w ćwiczeniu 1.
4. Zastanów się przez moment, jak bardzo społeczeństwo, w którym żyjesz, zmieniło się podczas twojego życia. Określ trzy lub cztery wskaźniki społeczne, które mogłyby być wykorzystane do monitorowania skutków co najmniej jednej z tych zmian dla jakości życia w twoim społeczeństwie.
5. Biuro do Spraw Więziennictwa (*U.S. Bureau of Prisons*) zajmuje się badaniami ewaluacyjnymi wielu aspektów więziennictwa. Zlokalizuj jedno z ich badań i napisz krótkie podsumowanie, dotyczące projektu badań oraz wniosków z nich wyprowadzonych. Skorzystaj ze strony internetowej: http://www.bop.gov/index.html

Dodatkowe lektury

Berg Richard, Rossi Peter H., *Thinking about Program Ewaluation*, Sage, Thousand Oaks, Cal. 1998. Wspaniała książka, jeżeli poszukujesz solidnych podstaw badań ewaluacyjnych, ubarwionych obszernym zbiorem przykładów[3].

[3] Drugie, nieznacznie zmienione wydanie: Berg Richard, Rossi Peter H., *Thinking about Program Evaluation*, Sage, Thousand Oaks, Cal. 1999 (przyp. tłum.).

Chen Huey-Tsych, *Theory-Driven Evaluations*, Sage, Newbury Park, Cal. 1990. Chen dowodzi, że badania ewaluacyjne, jeśli mają być znaczące i użyteczne, muszą mieć solidne podstawy teoretyczne.

Cunningham J. Barton, *Action Research and Organizational Development*, Praeger, Westport, Conn. 1993. Książka ta ponagla badaczy do przerzucenia pomostu między teorią a działaniem, tak by stali się oni zaangażowanymi uczestnikami ewolucji organizacji społecznych oraz wykorzystywali badania społeczne do monitorowania rodzących się tam problemów i sposobów ich rozwiązywania.

***Handbook** of Applied Social Research Methods*, Leonard Bickman, Debra J. Rog (red.), Sage, Thousand Oaks, Cal. 1998. Oboje redaktorzy tej książki dostarczają przykładów ilustrujących wszystkie etapy badań ewaluacji; od planowania, poprzez gromadzenie danych i ich analizę. Ponadto omawiają etyczne problemy w szczególnym kontekście badań ewaluacyjnych.

Hedrick Terry E., Bickman Leonard, Rog Debra J., *Applied Research Design. A Practical Guide*, Sage, Newbury Park, Cal. 1993. Jest to wprowadzenie do badań ewaluacyjnych, stanowiące – jak głosi jego podtytuł – praktyczny przewodnik omawiający w sposób bezpośredni kompromisy, jakich na ogół trzeba dokonywać w trakcie projektowania badań i ich realizacji.

Rossi Peter H., Freeman Howard E., *Evaluation: A Systematic Approach*, Sage, Newbury Park, Cal. 1996. To obszerne omówienie badań ewaluacyjnych jest świetnym materiałem źródłowym. Poza omówieniem kluczowych pojęć badań ewaluacyjnych autorzy podają liczne przykłady, które mogą pomóc w kierowaniu własnymi projektami.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Część
4

Analiza danych

13. Analiza danych jakościowych

14. Kwantyfikacja danych

15. Podstawy analizy ilościowej

16. Model analizy rozbudowanej

17. Statystyka w naukach społecznych

W tej części książki będziemy omawiać analizę danych, zgromadzonych podczas badań społecznych, a także przyjrzymy się kolejnym krokom, które oddzielają obserwację od końcowej prezentacji wniosków z pracy badawczej.

W rozdziale 1 wprowadziłem zasadnicze rozróżnienie danych jakościowych i ilościowych. W następującym po nim omówieniu mogliśmy się przekonać, że wiele z podstawowych założeń badań społecznych odnosi się do obu typów danych. Jednak sposób analizy danych jakościowych i ilościowych jest całkowicie odmienny. Z tego też powodu omawiam je oddzielnie.

W rozdziale 13 przyjrzymy się analizie danych jakościowych. Na początku omówimy niektóre z podstaw teoretycznych tego podejścia badawczego. Następnie spojrzymy na niektóre procedury, które mogą być przydatne dla poszukiwania znaczeń wśród zgromadzonych danych jakościowych. Na koniec zaprezentuję niektóre spośród programów komputerowych, stworzonych specjalnie dla analizy danych jakościowych.

Rozdział 14 dotyczy sposobów kwantyfikacji danych, które mogły być gromadzone w ramach każdego z modeli obserwacji omówionych w części 3. Zobaczymy w nim, jak znaczący może być wpływ komputerów na tego typu czynności.

Wstępne zagadnienia dotyczące logiki analizy danych ilościowych przedstawiam w rozdziale 15. Rozpoczniemy go od przyjrzenia się metodom analizy i prezentacji danych dotyczących pojedynczej zmiennej. Następnie zajmiemy się związkiem między dwiema zmiennymi oraz nauczymy się, jak tworzyć i odczytywać proste tabele rozkładów procentowych. Rozdział ten zakończy wprowadzenie do analizy wielozmiennowej.

W rozdziale 16 opisany jest elaboracyjny model rozbudowanej analizy danych opracowany przez Paula Lazarsfelda z Uniwersytetu Columbia. W rozdziale tym dalej będziemy się zajmować analizą wielozmiennową. Tam też zaprezentowana zostanie analiza przyczynowa za pomocą tabel rozkładów procentowych. Zastosujemy tę samą logikę analizy w następnych rozdziałach, gdy będziemy omawiać użycie innych technik statystycznych. Ten model logiczny został opracowany dla potrzeb analizy danych ilościowych, sądzę jednak, że zauważysz jego przydatność również dla celów analizy danych jakościowych.

Rozdział 17 zawiera wprowadzenie do niektórych metod analizy statystycznej najczęściej stosowanych w naukach społecznych, wraz z wybranymi, bardziej zaawansowanymi metodami analizy wielozmiennowej. W miejsce prostej prezentacji, jak należy obliczać statystyki korzystając z tych metod (to potrafią robić komputery), starałem się pokazać, jak statystyki te mieszczą się w kontekście wcześniejszych rozważań teoretycznych. Pokażę także ich związek z logiką analizy danych. Kończąc ten rozdział, powinieneś zatem wiedzieć, kiedy korzystać z poszczególnych miar statystycznych pomiarów, a także, jak je obliczać. ■

Rozdział
13

Analiza danych jakościowych

Przegląd holograficzny

Analiza danych jakościowych jest nieliczbowym szacowaniem wartości obserwacji dokonanych dzięki zastosowaniu obserwacji uczestniczącej, analizy treści, wywiadów pogłębionych i innych technik badań jakościowych. Chociaż analiza jakościowa jest w takim samym stopniu sztuką jak nauką, ma swoją własną logikę i charakterystyczne dla siebie techniki. Niektóre z nich są wspomagane specjalnymi programami komputerowymi.

Wprowadzenie

Powiązanie teorii z analizą
Odkrywanie prawidłowości
Metoda teorii ugruntowanej
Semiotyka
Analiza konwersacyjna

Przetwarzanie danych jakościowych
Kodowanie
Zapisywanie notatek
Tworzenie mapy pojęciowej

Programy komputerowe dla danych jakościowych
Księga Kapłańska widziana oczami programu NUD*IST
Sandrine Zerbib: zrozumieć kobietę reżysera

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Dalsze rozdziały w części 4 tej książki będą dotyczyć *ilościowej* analizy danych w badaniach społecznych, zwanej niekiedy *analizą statystyczną*. Badania społeczne w ostatnich dziesięcioleciach zmierzają do koncentrowania się na technikach ilościowej analizy danych. Praktyka taka zdominowała inne podejście do nadawania znaczeń obserwacjom społecznym: **analizę jakościową** – metodę analizy danych, pochodzących z badań społecznych, bez sprowadzania ich do formatu liczbowego. Podejście takie było jednak wcześniejsze niż analiza ilościowa. Pozostaje ono wciąż użytecznym narzędziem analizy danych, ciesząc się ponownym zainteresowaniem wśród badaczy społecznych.

Analizy statystyczne mogą wprowadzić w zakłopotanie niektórych studentów. Niekiedy możliwe jest jednak wyuczenie się kolejnych kroków procedury badawczej przez powtarzanie. Chcę powiedzieć, że w miarę nabierania wprawy, ćwiczenie umiejętności badań ilościowych może prowadzić do bardziej pogłębionego zrozumienia logiki leżącej u podstaw wykorzystania tych technik badawczych.

Uczenie analizy jakościowej jako serii powtarzalnych procedur jest znacznie trudniejsze. W tym przypadku zrozumienie musi poprzedzać praktykę. W tym rozdziale rozpoczniemy od powiązań między badaniami a teorią w jakościowej analizie danych. Następnie sprawdzimy kilka procedur badawczych, przydatnych dla osiągnięcia celów teoretycznych. Po rozważeniu niektórych prostych technik manualnych, zabierzemy się za programy komputerowe.

Powiązanie teorii z analizą

Jak sugerowałem w rozdziale 10 i w innych fragmentach tej książki, metody jakościowe zakładają stały związek między gromadzeniem danych a teorią. Dlatego też wspominałem już o jakościowej analizie danych przy okazji wcześniejszych rozważań, dotyczących badań terenowych i analizy treści. W badaniach ilościowych łatwo jest czasami pogrążyć się w procedurach gromadzenia danych i ich statystycznym opracowaniu, tracąc przez to na pewien czas z pola widzenia związek z teorią. W przypadku badań jakościowych jest to mniej prawdopodobne, gdyż proces gromadzenia danych, analiza i teoria są wyraźnie ze sobą powiązane.

W dalszych rozważaniach będziemy korzystać z wersji teorii zaproponowanej przez Anselma Straussa i Juliet Corbin (1994, s. 278), jako składającej się z: *„wiarygodnie* wysnutych relacji między pojęciami i zbiorami pojęć". Podkreślają oni „wiarygodność" dla zasygnalizowania, że teorie najlepiej odzwierciedlają nasze rozumienie tego, jak funkcjonuje życie. Im bardziej nasze badania potwierdzają szczególny rodzaj związków między poszczególnymi pojęciami, tym większej nabieramy pewności, że nasz sposób rozumienia świata odpowiada rzeczywistości społecznej.

Podczas gdy badania jakościowe czasem są realizowane w celach czysto opisowych – jak na przykład czyni to antropologiczna etnografia, opisująca style życia plemion wcześniej nieznanych – w dalszej części rozdziału będziemy się koncentrować przede wszystkim na poszukiwaniu schematów wyjaśniających. Jak zobaczymy, niekiedy takie prawidłowości pojawiają się z czasem, niekiedy zaś przybierają formę związku przyczynowego między zmiennymi. Spójrzmy na parę sposobów, dzięki którym badacze o orientacji jakościowej odkrywają takie prawidłowości.

Odkrywanie prawidłowości

John i Lyn Loflandowie (1995, s. 127–145) zasugerowali sześć różnych sposobów poszukiwania prawidłowości w pewnym konkretnym temacie badań. Załóżmy, że przedmiotem twojego zainteresowania jest przemoc w stosunku do dzieci w danym środowisku. Masz możliwość zadania paru pytań, które pomogą ci zrozumieć zgromadzone przez ciebie dane.

1. Częstotliwość: Jak często przemoc w stosunku do dzieci pojawia się w rodzinach w badanym środowisku? (Pamiętaj, iż może istnieć różnica między tą częstotliwością a tym, co ludzie zechcą ci o niej powiedzieć).

2. Nasilenie: Jaki jest poziom przemocy? Jaki jest stopień jej brutalności?
3. Struktura: Jakie są różne typy przemocy: fizyczna, psychiczna, seksualna? Czy są one ze sobą w jakikolwiek sposób powiązane?
4. Procesy: Czy występuje jakieś uporządkowanie wśród elementów tej struktury? Czy przemoc rozpoczyna się od znęcania psychicznego, po którym następuje przemoc fizyczna i molestowanie seksualne, czy też kolejność elementów zmienia się?
5. Przyczyny: Jakie są przyczyny przemocy wobec dzieci? Czy pojawia się częściej wśród określonej klasy społecznej lub poszczególnych grup wyznaniowych lub etnicznych? Czy przypadki przemocy zdarzają się częściej w dobrych, czy w złych okresach?
6. Konsekwencje: Jak przemoc wobec dzieci wpływa na jej ofiary, zarówno w krótkim, jak i w długim czasie? Jakie wywołuje ona zmiany w osobach stosujących przemoc?

W większości przypadków podczas analizy danych będziesz poszukiwać prawidłowości ujawniających się w kilku obserwacjach reprezentujących na ogół różne przypadki będące przedmiotem badań. Michael A. Huberman i Matthew B. Miles (1994, s. 435f) proponują dwie strategie przekrojowej analizy przypadków (*cross-case-analysis*): analizę zorientowaną na zmienną (*variable-oriented analysis*) oraz analizę zorientowaną na przypadek (*case-oriented analysis*). **Analiza zorientowana na zmienną** jest podobna do modelu, który omawialiśmy już kilkakrotnie w tej książce. Jeśli próbujemy określić, jaka będzie decyzja co do pójścia do college'u, Huberman i Miles sugerują rozważenie takich zmiennych, jak: „płeć, status społeczny, oczekiwania rodziców, wyniki nauczania i decyzja związana z pójściem do college'u" (1994, s. 435). Tak więc moglibyśmy określić, czy kobiety, czy też mężczyźni częściej idą do college'u. W centrum zainteresowania naszej analizy będą relacje między zmiennymi, a obserwowani ludzie będą przede wszystkim nośnikami tych zmiennych.

Analiza zorientowana na zmienne może przypominać rozważania z rozdziału 1, w którym została wprowadzona idea wyjaśniania nomotetycznego. Jej celem jest osiągnięcie cząstkowego ogólnego wyjaśnienia przy wykorzystaniu stosunkowo niewielkiej liczby zmiennych. Sondaże polityczne, których celem jest wyjaśnienie preferencji wyborczych na podstawie dwóch lub trzech kluczowych zmiennych, stosują takie podejście badawcze. Badacz nie rości sobie pretensji do przewidywania zachowania każdej jednostki ani do dogłębnego wyjaśniania motywacji jakichś pojedynczych osób. Czasami przydatne jest jednak chociaż cząstkowe wyjaśnienie ogólnych tendencji i podejmowanych działań.

Możesz także przypomnieć sobie wprowadzenie do wyjaśniania idiograficznego z rozdziału 1, kiedy to staralibyśmy się zrozumieć dogłębnie pojedynczy przypadek. W przykładzie z głosowaniem staraliśmy się dowiedzieć wszystkiego, co może być brane pod uwagę jako czynniki determinujące decyzje pojedynczych osób, dotyczące odpowiedzi na pytanie „jak głosować?". Taka orientacja leży u podstaw tego, co Huberman i Miles nazywają **analizą zorientowaną na przypadek.**

> „W analizie zorientowanej na przypadek przyglądamy się dokładniej konkretnemu przypadkowi, powiedzmy: Przypadkowi 005. Jest nim kobieta z klasy średniej, której rodzice mają wysokie oczekiwania itp. Jednak mierniki te są «mizerne». Aby dokonać autentycznej analizy przypadku, musimy przyjrzeć się pełnej historii Przypadku 005: Nynke van der Molen, której matka z wykształcenia jest pracownikiem społecznym i jest rozgoryczona faktem, że nigdy nie pracowała poza domem; której ojciec pragnął, aby Nynke pracowała w rodzinnej kwiaciarni. Chronologia badanego przypadku jest również ważna: dwa lata temu najbliższa przyjaciółka Nynke postanowiła podjąć studia w college'u, na krótko przed tym, jak Nynke zaczęła pracować w stajni, oraz zanim matka Nynke pokazała jej album ze zdjęciami z czasów nauki w szkole dla pracowników społecznych. Po tym Nynke zdecydowała, że zapisze się na studia weterynaryjne". (1994, s. 436)

Ten skrócony komentarz powinien dać nam pewne wyobrażenie o szczegółowości analizy tego typu. Oczywiście cała analiza powinna być

bardziej obszerna i pogłębiona. Jednak takie pełne, idiograficzne badanie nie mówi nam nic o ludziach w ogóle. W żaden sposób nie dostarcza nam teorii dotyczącej tego, dlaczego ludzie w ogóle idą na studia.

Mimo to badacz, poza dogłębnym zrozumieniem pojedynczej osoby, traktuje decydujące elementy, związane z doświadczeniem jednostki, jako przykłady bardziej ogólnych pojęć lub zmiennych społecznych. Na przykład zawód pracownika społecznego matki Nynke może być postrzegany jako „edukacja matki"; decyzja przyjaciółki może być postrzegana jako „wpływ bliskich osób". Bardziej precyzyjnie – mogą być one postrzegane jako zmienne niezależne, mające wpływ na zmienną zależną, jaką jest rozpoczęcie studiów.

Oczywiście pojedynczy przypadek nie tworzy jeszcze teorii – z tego powodu Huberman i Miles odnoszą się do **przekrojowej analizy przypadków** (*cross-case analysis*), w której badacz zwraca się ku innym badanym, również analizując drobiazgowo szczegóły z ich życia, lecz poświęcając szczególnie wiele uwagi tym zmiennym, które wydawały się ważne w pierwszym przypadku. Jakiego rodzaju edukację odebrały matki pozostałych osób i jak długo ona trwała? Czy są jakieś świadectwa o bliskich przyjaciołach, którzy poszli na studia?

Niektóre kolejne przypadki będą podobne do pierwszego, jeśli chodzi o ujawniony wpływ poszczególnych zmiennych. Inne przypadki nie będą w ogóle jednak przypominać pierwszego. Te ostatnie mogą wymagać zidentyfikowania nowych ważnych zmiennych, które skłonią badacza do poszukiwania odpowiedzi na pytanie, dlaczego niektóre przypadki układają się, jak się wydaje, w jeden schemat, schemat kolejnych zaś jest inny.

Metoda teorii ugruntowanej

Opisana powyżej metoda przekrojowej analizy przypadków wygląda jakby znajomo. W rozważaniach dotyczących teorii ugruntowanej, zawartych w rozdziale 10, widzieliśmy, jak badacze zorientowani jakościowo próbują czasami stworzyć teorię na czysto indukcyjnych fundamentach. Takie podejście rozpoczyna się raczej od obserwacji niż od formułowania hipotez i zmierza do odkrycia prawidłowości oraz stworzenia teorii od podstaw, bez wstępnych założeń, choć zdarza się i tak, że badacze mogą konstruować i opracowywać ją na zrębach wcześniejszych teorii ugruntowanych.

Teoria ugruntowana po raz pierwszy została sformułowana przez socjologów Barneya Glasera i Anselma Straussa (1967), jako próba poradzenia sobie z badaniami klinicznymi w ramach socjologii medycyny. (Możesz wysłuchać Glasera, omawiającego teorię ugruntowaną w Internecie na stronie: http://www.grounded-theory.com/vidseries1.html). Od tego czasu rozwinęła się jako metoda, jej twórcy zaś podążyli różnymi ścieżkami. W poniższym omówieniu zaprezentowano podstawowe koncepcje i procedury znane jako **metoda teorii ugruntowanej (MTU)**.

Oprócz fundamentalnej, indukcyjnej zasady tworzenia teorii z danych, MTU wykorzystuje **metodę stałego porównywania** (*constant comparative method*). Glaser i Strauss opisali pierwotnie tę metodę jako składającą się z czterech etapów (1967, s. 105–113):

1. „Porównywanie zdarzeń odnoszących się do każdej kategorii". Gdy Glaser i Strauss badali reakcje pielęgniarek na zbliżającą się śmierć pacjentów, pozostających pod ich opieką, zauważyli, iż pielęgniarki szacowały „stratę społeczną" następującą w wyniku śmierci pacjenta. Gdy to pojęcie pojawiło się w wyniku analizy jednego przypadku, badacze zaczęli poszukiwać świadectw tego zjawiska w innych przypadkach. Kiedy odkryli, że pojęcie to pojawia się w przypadku kilku pielęgniarek, zaczęli porównywać różne zdarzenia. Taki proces jest podobny do konceptualizacji opisanej w rozdziale 5, w którym szczegółowo omówiono naturę i wymiary wielu pojęć wypływających z danych.
2. „Integrowanie kategorii i ich własności". Na tym etapie badacz zaczyna zwracać uwagę na relacje między pojęciami. Na przykład w szacowaniu straty społecznej Glaser i Strauss odkryli, że pielęgniarki zwracały szczególną uwagę na wiek pacjenta, jego wykształcenie i obowiązki rodzinne. Jed-

nak dla uchwycenia tych właśnie relacji badacze musieli zwrócić uwagę na wszystkie te pojęcia.

3. „Określanie zakresu teorii". W końcu, gdy wzorce relacji między pojęciami stawały się wyraźne, badacze mogli zignorować niektóre z uprzednio odnotowanych pojęć; takie, które w sposób jednoznaczny nie wiązały się z problematyką badań. Oprócz redukcji liczby kategorii sama teoria może również ulec uproszczeniu. Na przykład w badaniach straty społecznej Glaser i Strauss odkryli, że proces szacowania straty społecznej może być uogólniony poza pielęgniarki i umierających pacjentów. Proces ten, jak się wydaje, odnosi się do sposobów, w jakie cały personel traktuje wszystkich pacjentów (umierających lub nie).

4. „Pisanie teorii". Na koniec badacz musi przełożyć swoje wnioski na słowa, by podzielić się nimi z innymi ludźmi. Jak może wiesz z własnego doświadczenia, akt komunikowania własnego rozumienia czegoś powoduje modyfikację, czy nawet poprawę własnego rozumienia danego tematu. Etap pisania jest więc postrzegany jako część procesu badawczego w ramach MTU. Kwestia ta rozwinięta zostanie w dalszej części tego rozdziału, dotyczącej zapisywania notatek.

Ten krótki przegląd powinien dać ci pewne wyobrażenie o poszczególnych etapach zastosowania teorii ugruntowanej. O technikach związanych z MTU napisano wiele. Liczne informacje na ten temat można znaleźć w Internecie. Jedną z kluczowych pozycji jest *Basics of Qualitative Research* (1990) Anselma Straussa i Juliet Corbin, w której zawarto omówienie i rozwinięcie wielu koncepcji i technik, pojawiających się już w pierwszej książce Glasera i Straussa. W Internecie można korzystać z opracowania Gaelle T. Morin *Grounded Theory Methodology on the Web* na stronie: http://www.geocities.com/ResearchTriangle/Lab/1491/gtm-i-gtmweb.html.

Metoda teorii ugruntowanej jest jednym z analitycznych ujęć danych jakościowych. Z innymi wyspecjalizowanymi technikami zapoznamy się na kolejnych stronach.

Semiotyka

Semiotyka jest powszechnie definiowana jako „nauka o znakach" i zajmuje się symbolami i znaczeniami. Często kojarzy się ją z analizą treści, która została omówiona w rozdziale 11, chociaż może mieć zastosowanie w różnych kontekstach badawczych.

Peter K. Manning i Betsy Cullum-Swan (1994, s. 466) zaproponowali następujący sposób rozumienia zastosowania semiotyki: „Chociaż semiotyka osadzona jest w języku, język jest tylko jednym spośród systemów znaków, różniących się między sobą stopniem całościowości, możliwościami zastosowania i złożonością. Alfabet Morse'a, savoir vivre, matematyka, muzyka, a nawet znaki drogowe są przykładami systemów semiotycznych".

Jednak żaden znak nie ma przyrodzonego mu znaczenia. Wszystkie znaczenia znajdują się w umyśle, co sprawia, że dany znak znaczy coś dla danej osoby. To, co czyni z semiotyki naukę *społeczną*, to właśnie nasze porozumienie, dotyczące znaczeń poszczególnych znaków. Jak zaznaczają Manning i Cullum-Swan:

> „Na przykład lilia konwencjonalnie kojarzy się ze śmiercią, Wielkanocą i zmartwychwstaniem. Dym jest łączony z papierosami i rakiem, a Marilyn Monroe z seksem. Każdy z tych związków jest konstruowany społecznie i arbitralnie, co sprawia, że istnieje wiele rodzajów powiązań między wyrażeniem a treścią".
>
> (1994, s. 466)

Aby odkryć takie powiązania, sprawdź, czy potrafisz połączyć znaki przedstawione na rysunku 13.1 z ich znaczeniem. Jestem pewny, że znasz wszystkie „poprawne" skojarzenia, tak że nie ma potrzeby, abym podawał ci właściwe odpowiedzi. (W porządku – powinieneś powiedzieć: 1c, 2a, 3b, 4e, 5d). Rzecz w tym, co sprawia, że każdy z tych znaków ma swoje własne znaczenie? Napisz i wyślij pocztą elektroniczną wiadomość do jakiegoś marsjańskiego socjologa, wyjaśniając logikę, która rządzi tworzeniem takich związków. (Może będziesz chciał dołączyć do listu jakieś emotikony, takie jak :–) – kolejny przykład semiotyki).

Rysunek 13.1

Łączenie znaków i ich znaczeń

Znak	Znaczenie
1. kwiat gwiazdy betlejemskiej	a. szczęście
2. podkowa	b. główna nagroda
3. niebieska wstążka	c. Boże Narodzenie
4. „powiedz: «marmolada»"	d. działanie
5. „połam nogi"	e. uśmiech do zdjęcia

Choć nie ma wątpliwości dotyczących każdego ze związków przedstawionych na rysunku 13.1, wszystkie znaczenia, które ty i ja współcześnie „znamy", są konstruowane społecznie. Analiza semiotyczna polega na poszukiwaniu znaczeń intencjonalnie i nieintencjonalnie przypisywanych do znaków.

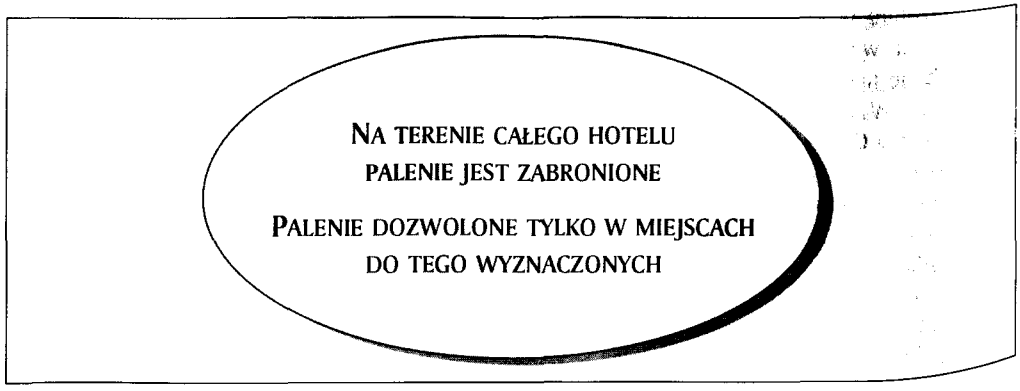
Rozważmy znak, znajdujący się w recepcji hotelu w Portland w stanie Oregon, który przedstawiono na rysunku 13.2. Co komunikuje ten dość dwuznaczny znak? Pierwsze zdanie zdaje się mówić, iż hotel nadąża za powszechną w całych Stanach Zjednoczonych tendencją do odchodzenia od używania tytoniu. Gość chcący zatrzymać się w wolnym od dymu środowisku, nie musi szukać dalej. W całym hotelu powietrze jest bowiem zdrowe i czyste. Jednak, jak głosi drugie zdanie, hotel nie chce być postrzegany jako miejsce niegościnne dla palaczy. Jest w nim miejsce dla każdego. Nikt nie musi się czuć odrzucony. Znak ten łatwiej zrozumieć w ramach paradygmatu marketingowego niż paradygmatu logiki.

Oczywiście „znaki", badane przez semiotykę, nie ograniczają się to tego typu znaków. Większość z nich jest tak naprawdę całkiem inna. *Znakami* są wszystkie rzeczy, którym przypisano specjalne znaczenia. Mogą nimi być takie rzeczy, jak: znaki logo, zwierzęta, ludzie, dobra konsumpcyjne. Czasami symbolizm jest nieco bardziej subtelny. Klasyczną analizę można znaleźć w pracy Ervinga Goffmana *Gender Advertisements* (1979). Goffman skupił się na ilustracjach reklamowych w gazetach i czasopismach. Nadrzędnym celem tych reklam było oczywiście sprzedanie określonych produktów. Ale co jeszcze zostało tam zakomunikowane? – zapytał Goffman. Co w szczególności reklamy mówiły o kobietach i mężczyznach?

Podczas analizy zdjęć, przedstawiających zarówno kobiety, jak i mężczyzn, Goffmana zaintrygował fakt, iż wizerunki mężczyzn były zawsze większe i wyższe niż wizerunki towarzyszących im kobiet. (Faktycznie, w wielu przypadkach układ zdjęcia wyraźnie dawał do zrozumienia, że kobieta jedynie towarzyszy mężczyźnie). Choć najbardziej narzucającym się wyjaśnieniem jest to, że przeciętnie mężczyźni są silniejsi i wyżsi niż kobiety, Goffman zasugerował, iż prawidłowość ta ma inne znaczenie: wielkość i umiejscowienie sugeruje *status* – ci potężniejsi i wyżsi mają prawdopodobnie wyższą pozycję społeczną – większą władzę i autorytet (1979, s. 28). Goffman zasugerował, iż reklamy komunikują, że mężczyźni są ważniejsi od kobiet.

Rysunek 13.2

Sprzeczne komunikaty?



W duchu wyjaśnień Freuda „czasami cygaro jest tylko cygarem" (sam Freud był palaczem): czy omawiane reklamy po prostu odzwierciedlają różnice biologiczne przeciętnych wymiarów mężczyzn i kobiet, czy też niosą przesłanie dotyczące statusu społecznego? Po części wnioski Goffmana opierały się na analizie przypadków wyjątkowych: takich reklam, na których pojawiające się kobiety wydawały się wyższe niż mężczyźni. W tych przypadkach mężczyzna miał na ogół niższy status społeczny – na przykład był to kucharz obok statecznej damy. Potwierdzało to zasadnicze spostrzeżenie Goffmana: wielkość i wysokość wskazują na status społeczny.

Taki sam wniosek został wywiedziony z ilustracji przedstawiających mężczyzn różnego wzrostu. Mężczyźni o wyższym statusie byli przedstawiani jako wyżsi, czy był to dżentelmen mówiący do kelnera, czy też kierownik pouczający swoich młodszych podwładnych. Gdy faktyczny wzrost prezentowanej postaci był niejasny, Goffman zwracał uwagę na ulokowanie głów osób przedstawionych na zdjęciu. Podwładni kucali, ich szef zaś pochylał się nad nimi. Głowa służącego była pochylona, tak że znajdowała się poniżej głowy osoby wydającej polecenie.

Tak więc ukryte przesłanie, przekazywane poprzez reklamy, było takie: im wyżej znajdowała się głowa osoby pojawiającej się na zdjęciu, tym ta osoba była ważniejsza. A na większości reklam, przedstawiających mężczyzn i kobiety, ci pierwsi wyraźnie byli ukazani jako ważniejsi. Niezależnie od tego, czy było to zamierzone, czy też nie, ukryte przesłanie płynące z reklam było takie, że mężczyźni mają większą władzę i cieszą się wyższym statusem społecznym niż kobiety.

Poza fizycznymi rozmiarami Goffman przeanalizował kilka innych różnic w sposobach przedstawiania mężczyzn i kobiet. Kolejnym przykładem było to, że mężczyzn prezentowano w rolach aktywnych, kobiety zaś – w pasywnych. Lekarz (mężczyzna) badał dziecko, natomiast pielęgniarka lub matka dziecka (kobiety) spoglądały na niego często z podziwem. Mężczyzna uczył kobietę grać w tenisa (jego głowa była wyżej od głowy kobiety). Mężczyzna ściskał wodze galopującego konia, kobieta jadąca za nim oplatała jego talię ramionami. Kobieta trzymała piłkę, mężczyzna zaś piłkę kopał. Mężczyzna robił zdjęcie przedstawiające wyłącznie kobiety.

Goffman zasugerował, że takie obrazowe wzorce subtelnie utrwalają liczne stereotypy związane z płcią. Nawet jeśli ludzie publicznie głoszą równouprawnienie, takie zdjęcia reklamowe po cichu sprowadzają mężczyzn i kobiety do „właściwych" ról.

Analiza konwersacyjna

Etnometodologia, jak sobie przypominasz, zmierza do ujawnienia ukrytych założeń i struktur życia społecznego. **Analiza konwersacyjna** (*conversation analysis – CA*) poszukuje metody osiągnięcia tego celu poprzez drobiazgowe badanie sposobów prowadzenia rozmowy. Podczas omówienia etnometodologii, zawartego w rozdziale 11, mogliśmy przyjrzeć się niektórym przykładom analizy konwersacyjnej. W tym miejscu omówimy tę technikę nieco bardziej dogłębnie.

David Silverman, dokonując przeglądu dorobku innych teoretyków i badaczy, reprezentujących nurt analizy konwersacyjnej, wymienił trzy fundamentalne założenia tej techniki. Po pierwsze, konwersacja jest czynnością społecznie ustrukturyzowaną. Tak jak inne struktury społeczne, ustanowiła swoje reguły zachowania. Na przykład oczekujemy, że tylko jedna osoba będzie mówić w danym momencie. Podczas rozmowy telefonicznej od osoby, która podniesie słuchawkę, oczekujemy, że pierwsza się odezwie (na przykład powie „halo"). Możesz od czasu do czasu zweryfikować funkcjonowanie tej zasady, gdy odbierzesz telefon i nie odezwiesz się do słuchawki. Przypomnij sobie, że tego typu zachowania są typowe dla etnometodologów.

Po drugie, Silverman zaznacza, że konwersacja musi być rozumiana kontekstualnie. Identyczna wypowiedź będzie miała całkowicie odmienne znaczenie w różnych kontekstach. Wyobraź sobie na przykład, jak różne będzie znaczenie słowa „nawzajem", jeśli będzie poprzedzone stwierdzeniem: „nie podoba mi się

twój wygląd" lub gdy poprzedzać go będzie „życzę ci miłego dnia".

Po trzecie, analiza konwersacyjna zmierza do zrozumienia struktury i znaczenia konwersacji poprzez niezwykle drobiazgową transkrypcję rozmowy. Nie tylko dokładne brzmienie słów jest tak zapisywane, lecz także wszystkie „oh", „hmm", błędy gramatyczne i pauzy. Przerwy w rozmowie notuje się z dokładnością co do dziesiętnej części sekundy.

Istnieje wiele praktycznych zastosowań analizy tego typu. Na przykład Ann Marie Kinnel i Douglas Maynard (1996) analizowali rozmowy między personelem kliniki wykonującej testy na obecność wirusa HIV i osobami poszukującymi tam porady. Przedmiotem badań było to, jak przekazywano informacje o bezpiecznym seksie. Badacze odkryli między innymi, że członkowie personelu raczej skłaniali się do udzielania standardowych informacji niż do rozmowy uwzględniającej specyficzną sytuację każdej osoby poszukującej porady. Ponadto wydawało się, że personel niechętnie udzielał bezpośrednich porad odnoszących się do bezpiecznego seksu, i ograniczał się wyłącznie do informacji.

Powyższe omówienie powinno dać ci pewne wyobrażenie o różnorodności jakościowych metod analitycznych możliwych do wykorzystania przez badaczy. Teraz przyjrzyjmy się wybranym technikom przetwarzania i analizy danych powszechnie używanym w badaniach jakościowych.

Przetwarzanie danych jakościowych

Chciałbym rozpocząć ten podrozdział od przestrogi. Czynność, którą będziemy teraz omawiać, należy zarówno do dziedziny sztuki, jak i do nauki. W wielkim uproszczeniu, nie istnieją w jej ramach wyraźnie określone i zdefiniowane etapy działań gwarantujące osiągnięcie sukcesu.

Przypomina to bardziej uczenie się, jak akwarelą namalować obraz lub jak skomponować symfonię. Nauczanie takich czynności jest oczywiście możliwe i uniwersytety prowadzą takie zajęcia. Nauczanie malarstwa i kompozycji ma swoiste dla nich konwencje i techniki, jak również podpowiedzi, które mogą być użyteczne dla adeptów każdej z nich. Jednak możliwy instruktaż dociera tylko po pewnego miejsca. Efekt finalny wynika z indywidualnej pracy artysty. To, co można powiedzieć o przetwarzaniu danych jakościowych, jest bardzo podobne do tego schematu.

W tym podrozdziale przedstawione zostaną niektóre idee związane z kodowaniem danych jakościowych, zapisywaniem notatek oraz graficznym tworzeniem map pojęciowych. Choć jest to odległe od instrukcji obsługi typu „jak-to-zrobić", prezentowane idee są dobrym punktem wyjścia do poszukiwania porządku w danych jakościowych.

Kodowanie

Bez względu na to, czy prowadzisz obserwację uczestniczącą, wywiady pogłębione, czy zbierasz narracyjne dane biograficzne, czy prowadzisz analizę treści lub jakieś inne badania jakościowe, staniesz przed stale rosnącą górą danych – najczęściej przyjmującą formę materiałów tekstowych. Co robić dalej?

Kluczowym procesem analitycznym podczas badań jakościowych w naukach społecznych jest *kodowanie* – klasyfikowanie lub kategoryzowanie pojedynczych fragmentów danych, połączone z jakimś systemem ich odszukiwania. Procedury te pozwolą ci odszukać później interesujące cię fragmenty materiału.

Załóżmy, że relacjonujesz rozwój jakiegoś ruchu społecznego. Przypominasz sobie, że sporządziłeś jakieś notatki, dotyczące szczegółów formowania się tego ruchu. Teraz potrzebujesz tych informacji. Jeśli twoje notatki zostały skatalogowane według tematu, wydobycie tych, których potrzebujesz, powinno być proste. Możesz utworzyć zestaw segregatorów, opatrzonych różnymi nazwami, takimi jak „Historia". W tym przypadku wydobycie danych oznacza wyciągnięcie segregatora zatytułowanego „Historia" i przetrząsanie notatek w nim zawartych do momentu, aż znajdziesz to, czego potrzebujesz.

Jak zobaczysz w dalszej części tego rozdziału, obecnie istnieją niezwykle skomplikowane programy komputerowe, sprawiające, że proces porządkowania i wyszukiwania danych staje się szybszy, bardziej czytelny i bardziej precyzyjny. Zamiast przeglądać cały plik „Historia", możesz bezpośrednio przejść do notatek, dotyczących „wczesnej historii" lub „powstania" danego ruchu.

Kodowanie ma inny, jeszcze ważniejszy cel. Jak to omówiono wcześniej, celem analizy danych jest odkrywanie w nich prawidłowości, które będą wskazywać na teoretyczne wyjaśnienie życia społecznego. Kodowanie i poszukiwanie związków między pojęciami jest kluczowym elementem tego procesu i wymaga bardziej wyrafinowanego systemu niż zestaw tekturowych segregatorów. W tym podrozdziale zakładamy, iż kodowanie swoich danych będziesz przeprowadzał ręcznie. W końcowej części tego rozdziału zostanie przedstawione zastosowanie programów komputerowych do analizy danych jakościowych.

Jednostki kodowania

Jak pamiętasz z wcześniejszego omówienia analizy treści, w analizie statystycznej ważne jest określenie standaryzowanych jednostek analizy, poprzedzające kodowanie. Gdybyś na przykład porównywał amerykańskie i francuskie powieści, możesz zaznaczyć i zakodować zdania, akapity, rozdziały lub całą książkę. Jednak ważne będzie kodowanie tych samych jednostek w każdej z analizowanych powieści. Taka unifikacja jest konieczna w analizie ilościowej, gdyż pozwala nam formułować stwierdzenia w rodzaju: „23% akapitów zawiera metafory". Jest to możliwe tylko wtedy, jeśli będziemy kodować te same jednostki – akapity – w każdej z analizowanych powieści.

Jednak kodowanie danych w analizie jakościowej znaczy coś zupełnie innego. *Pojęcie* jest główną zasadą kodowania danych jakościowych. W tym przypadku jednostka tekstu odpowiednia dla zakodowania będzie się zmieniać w ramach danego dokumentu. Tak więc gdy przedmiotem badania będą organizacje, ich „wielkość" może wymagać tylko kilku słów do zakodowania, „misja" zaś może zabrać kilka stron. Długi opis gorącej dyskusji podczas zebrania akcjonariuszy może zostać zakodowany na przykład jako „wewnętrzne różnice poglądów".

Musisz też zdawać sobie sprawę, że dana kategoria kodowa może być zastosowana do materiałów tekstowych o całkowicie odmiennych długościach. Na przykład niektóre odniesienia do misji danej organizacji mogą być krótkie, inne zaś – długie. Standaryzacja jest główną zasadą w analizach ilościowych, inaczej natomiast jest w przypadku analiz jakościowych.

Kodowanie jako czynność

Zanim dalej będziemy omawiać logikę kodowania, zatrzymajmy się na moment, by zobaczyć, jak faktycznie wygląda sama czynność kodowania. John i Lyn Loflandowie (1995, s. 188) zaproponowali następujący opis ręcznego fiszkowania:

> „Zanim upowszechnił się dostęp do komputerów osobistych w drugiej połowie lat 80., kodowanie przybierało niejednokrotnie specyficzną formę fiszkowania. Badacz tworzył rozbudowany zestaw segregatorów z nazwami kodów na zakładkach, i fizycznie wkładał do nich albo poszczególne karty zawierające dane, albo notatki, dotyczące umiejscowienia danych w odpowiednich plikach skoroszytów. (...) Zanim kserowanie stało się dostępne i tanie, niektórzy badacze sporządzali swoje notatki na maszynie do pisania przez kalkę, dopisując kody na marginesach poszczególnych kopii swoich notatek, a potem cięli je nożyczkami. Następnie wkładali pocięte świstki papieru do odpowiednich segregatorów".

Loflandowie zwracają uwagę, iż komputery osobiste zdecydowanie uprościły to zadanie. Jednak wyobrażenie sobie kawałków papieru, zawierających fragmenty tekstu, wkładanych do segregatorów reprezentujących poszczególne kategorie kodowe, jest przydatne do zrozumienia procesu kodowania. W następ-

nym podrozdziale, gdy zaproponuję, byśmy zakodowali fragmenty tekstu, wyobraź sobie, że na kartach papieru mamy fragmenty tekstu przepisane na maszynie, które będziemy umieszczać w segregatorach zawierających nazwy odpowiednich kodów. Jeśli kiedykolwiek przypiszemy dwa kody do jakiegoś fragmentu tekstu, wyobraź sobie, że stworzymy dwie kopie danego fragmentu i włożymy je do dwóch różnych segregatorów reprezentujących dwa różne kody.

Tworzenie kodów

Czym więc powinny być kategorie kodowe? Glaser i Strauss (1967, s. 101f) dopuszczają możliwość kodowania danych dla przetestowania hipotez, które wynikają z wcześniej przyjętej teorii. W takim przypadku teoria powinna zasugerować kody, przybierające formę zmiennych.

Jednak w tym miejscu skoncentrujemy się na bardziej powszechnie stosowanym procesie **kodowania otwartego**. Strauss i Corbin (1990, s. 62) definiują je w następujący sposób:

> „Kodowanie otwarte jest tą częścią analizy, która szczególnie dotyczy nazywania i kategoryzowania zjawisk poprzez dokładne badanie danych. Bez tego pierwszego, podstawowego etapu analizy jej dalsza część wraz z prezentacją wyników nie mogłaby mieć miejsca. Podczas kodowania otwartego dane zostają podzielone na rozłączne części, dokładnie zbadane i porównane pod względem ich podobieństw i różnic. Następnie stawia się pytania, dotyczące zjawiska wyłaniającego się z danych. Podczas tego procesu sprawdzane i eksplorowane są kolejne założenia dotyczące tego zjawiska, co prowadzi do nowych odkryć".

Przejdźmy do konkretnego przykładu ilustrującego możliwość wykorzystania tej procedury. Przypuśćmy, że interesujesz się religijnym podłożem uprzedzeń wobec homoseksualistów. Przeprowadziłeś parę wywiadów z ludźmi, będącymi przeciwnikami homoseksualizmu, którzy odwoływali się do religii jako podstawy swoich odczuć. Szczególnie często odsyłali cię oni do następujących fragmentów z Księgi Kapłańskiej[1]:

„18:22 Nie będziesz cieleśnie obcował z mężczyzną jak z kobietą. Jest to obrzydliwością.

20:13 Mężczyzna, który obcuje cieleśnie z mężczyzną tak jak z kobietą, popełnia obrzydliwość; obaj poniosą śmierć; krew ich spadnie na nich".

Chociaż punkt widzenia wyrażony w tych cytatach zdaje się być jednoznaczny, możesz zdecydować się przeanalizować go dokładniej. Być może jakościowa analiza Księgi Kapłańskiej może przynieść pełniejsze rozumienie tego, w którym miejscu nakaz sprzeciwu wobec homoseksualizmu wpisuje się w szerszy kontekst moralności judeochrześcijańskiej.

Zacznijmy naszą analizę od zbadania dwóch właśnie zacytowanych fragmentów. Możesz rozpocząć kodowanie każdego z nich od przypisania im kodu „homoseksualizm". W sposób oczywisty jest to kluczowe pojęcie w twojej analizie. Zawsze gdy będziesz skupiał się na zagadnieniu homoseksualizmu podczas analizy Księgi Kapłańskiej, będziesz rozważał te dwa fragmenty.

Ponieważ homoseksualizm jest w tym przypadku kluczowym pojęciem, przyjrzyjmy się dokładniej, co on oznacza, biorąc pod uwagę dane będące przedmiotem studiów. Zauważmy na początek, jak *homoseksualizm* jest określany: mężczyzna cieleśnie obcujący z mężczyzną „jak z kobietą". Możemy sobie wprawdzie wyobrazić prawnika stającego u bram nieba i mówiącego: „z mojego punktu widzenia, jeśli cieleśnie nie obcowaliśmy...", jednak można spokojnie zakładać, że ten fragment dotyczy uprawiania seksu, choć nie jest do końca jasne, jakie specyficzne czynności mogą być tu zaliczone, a jakie nie.

Jednak zauważmy, że zakaz, jak się wydaje, dotyczy tylko *męskiego* homoseksualizmu; kobiecy homoseksualizm nie jest wymieniony. Tak

[1] Wszystkie cytaty na podstawie Biblii, zawarte w tym rozdziale, pochodzą z przekładu dokonanego przez Brytyjskie i Zagraniczne Towarzystwo Biblijne, Warszawa 1975 (przyp. tłum.).

więc w naszej analizie każdy z tych fragmentów może być również zakodowany jako „męski homoseksualizm". Ilustruje to dwa kolejne aspekty kodowania: (1) każda jednostka może być kodowana więcej niż jeden raz, (2) można zastosować kody hierarchiczne (jeden zawierający się w drugim). Obecnie do każdego fragmentu mamy przypisane dwa kody.

W tym miejscu można wprowadzić kod jeszcze bardziej ogólny: „Zakazane zachowanie". Jest to ważne z dwóch powodów. Po pierwsze, z analitycznego punktu widzenia homoseksualizm nie jest zły sam z siebie. Naszym celem było zbadanie sposobu, w jaki homoseksualizm staje się czymś złym według tekstu religijnego, będącego przedmiotem analizy. Po drugie, nasze studia nad Księgą Kapłańską mogą odkryć inne zakazane zachowania.

W cytowanych fragmentach pojawiają się co najmniej dwa inne, istotne pojęcia: „obrzydliwości" oraz „poniesienia śmierci". Zauważmy, że są wyraźnie związane z „zakazanym zachowaniem", lecz oznaczają coś zupełnie innego. Parkowanie samochodu bez wrzucenia monet do parkometru jest zakazane, lecz tylko niektórzy nazwaliby to obrzydliwością, a jeszcze mniej osób domagałoby się kary śmierci za takie wykroczenie. Przypiszmy te dwa nowe kody do naszych dwóch pierwszych fragmentów.

W tym miejscu chciałbym zostawić owe dwa kluczowe fragmenty i przeanalizować pozostałą część Księgi Kapłańskiej. Dlatego zbadamy i zakodujemy każdy z pozostałych rozdziałów i wersów. W naszej dalszej analizie będziemy korzystać z kodów już stworzonych oraz będziemy dodawać nowe tam, gdzie będzie to potrzebne. Gdy będziemy dodawać nowe kody, ważne będzie przeglądanie dotychczas zakodowanych fragmentów, aby sprawdzić, czy nowe kody nie dotyczą tych stworzonych wcześniej.

Oto fragmenty, co do których zadecydowaliśmy, że będziemy kodować je jako „obrzydliwość". (Zaznaczyłem w nich, co jest obrzydliwe).

„7:18 A jeżeli z **mięsa ofiary pojednania zje się coś trzeciego dnia**, to ten, który ją składał, nie znajdzie upodobania. Nie będzie ona mu zaliczona. Jest to rzecz nieczysta[2], a ten, kto z niej będzie jadł, ściągnie na siebie winę.

7:21 A kto **dotknie się czegokolwiek nieczystego**: nieczystości pochodzącej od człowieka lub nieczystego zwierzęcia, lub w ogóle jakiejś nieczystej obrzydliwości, a **zje coś z mięsa ofiary pojednania**, która należy do Pana, będzie wytracony spośród swego ludu.

11:10 **Wszystko zaś, co nie ma płetw i łusek w morzach i w rzekach**, wszystko, co się roi w wodzie, wszystkie istoty żywe, które są w wodzie, będą dla was obrzydliwością.

11:11 I pozostaną dla was obrzydliwością: **mięsa ich jeść nie będziecie i będziecie się brzydzić ich padliną.**

11:12 **Wszystko w wodzie, co nie ma płetw i łusek**, będzie dla was obrzydliwością.

11:13 Z ptactwa zaś tymi brzydzić się będziecie i **nie będziecie ich spożywali**, bo są obrzydliwością: **orzeł, sęp górski, orlik.**

11:14 Wszelkie gatunki **kani i sokołów.**

11:15 Wszelkie gatunki **kruków.**

11:16 **Struś, sowa, mewa** i wszelkie gatunki **jastrzębi.**

11:17 **Puszczyk, kormoran i puchacz.**

11:18 **Sójka, pelikan, ścierwojad.**

11:19 **Bocian**, wszelkie gatunki **czapli, dudek i nietoperz.**

11:20 **Wszelkie uskrzydlone czworonożne owady** będą dla was obrzydliwością.

11:41 **Wszelkie małe zwierzęta, które biegają po ziemi**, są obrzydliwością, nie wolno ich jeść.

11:42 Nie będziecie też jedli niczego, co czołga się na brzuchu, i niczego, co chodzi na czterech i więcej nogach spośród wszystkich **małych zwierząt, które biegają po ziemi**, gdyż są obrzydliwością.

11:43 Nie kalajcie siebie samych przez spożywanie któregokolwiek **płaza, który**

[2] W angielskim tekście Biblii zarówno „nieczystość", jak i „obrzydliwość" określane są wspólnym terminem „abomination", dlatego w zestawieniu znalazły się fragmenty, które w języku polskim byłyby raczej kodowane jako „nieczystość" (przyp. red. nauk.).

pełza i nie zanieczyszczajcie się nim, bo przez nie staniecie się nieczystymi.

18:22 Nie będziesz **cieleśnie obcował z mężczyzną jak z kobietą.** Jest to obrzydliwością.

19:6 Będzie ona spożywana w dniu waszej ofiary i nazajutrz, a co pozostanie do trzeciego dnia, będzie spalone w ogniu.

19:7 Jeżeli by **spożywano ją trzeciego dnia,** będzie nieczysta, nie zyska upodobania.

19:8 Ten, kto ją spożyje, stanie się winny, gdyż zbezcześcił świętość Pana i wytracony będzie ze swego ludu.

20:13 **Mężczyzna, który obcuje cieleśnie z mężczyzną tak jak z kobietą,** popełnia obrzydliwość; obaj poniosą śmierć; krew ich spadnie na nich.

20:25 Będziecie więc rozróżniali między zwierzętami czystymi i nieczystymi i między ptactwem nieczystym a czystym. **Nie kalajcie samych siebie przez zwierzęta i ptactwo i przez wszystko, co pełza po ziemi,** a co oddzieliłem, aby było dla was nieczyste".

Tak więc męski homoseksualizm nie jest jedyną obrzydliwością, określoną w Księdze Kapłańskiej. Gdy będziesz porównywać te fragmenty, poszukując podobieństw i różnic, stanie się widoczne, że większość obrzydliwości ma charakter reguł żywieniowych – zwłaszcza wtedy, gdy potencjalne jedzenie jest uważane za „nieczyste". Inne obrzydliwości wypływają z niewłaściwego obchodzenia się z ofiarami rytualnymi. „Reguły żywieniowe" i „Ofiary rytualne" reprezentują zatem dodatkowe koncepcje i kody, które mogą być użyte w naszej analizie.

Wspominałem wcześniej, że kara śmierci, jako kolejne pojęcie, może być także zastosowana w naszej analizie. Kiedy pójdziemy tym śladem, odkryjemy, iż wiele innych zachowań poza męskim homoseksualizmem prowadzi do kary śmierci. Wśród nich są takie zachowania, jak:

„20:2 Oddanie kogoś ze swojego potomstwa Molochowi (ofiara z człowieka).

20:9 Złorzeczenie swojemu ojcu i swojej matce.

20:10 Cudzołóstwo z żoną bliźniego swego.

20:11 Obcowanie cielesne z żoną swego ojca.

20:12 Obcowanie cielesne ze swoją synową.

20:14 Pojęcie za żonę kobiety oraz jej matki.

20:15 Obcowanie cielesne mężczyzny ze zwierzęciem (zwierzę także zostanie uśmiercone).

20:16 Obcowanie cielesne kobiety ze zwierzęciem.

20:27 Wywoływanie duchów lub wróżenie.

24:16 Bluźnierstwo imieniem Pana.

24:17 Zabicie człowieka".

Jak można zauważyć, kara śmierci jest szeroko stosowana w Księdze Kapłańskiej w stosunku do wszystkiego, począwszy od przekleństwa, a kończąc na morderstwie, z męskim homoseksualizmem gdzieś pośrodku.

Szersza analiza zakazanych zachowań, nie uwzględniająca obrzydliwości i śmierci, będzie także prowadzić do stworzenia obszernej listy. Między innymi będzie na niej oszczerstwo, zemsta, zawiść, klątwa sprawiająca głuchotę oraz rzucanie zawadzających głazów na drodze ślepców. W rozdziale 19, wers 19, Księga Kapłańska cytuje Boga nakazującego: „Będziecie przestrzegać ustaw moich. Nie będziesz twego bydła parzył z odrębnym gatunkiem. Twego pola nie będziesz obsiewał dwojakim gatunkiem ziarna i nie wdziewaj na siebie szaty zrobionej z dwóch gatunków przędzy". Niewiele dalej dodaje: „Nie będziecie jedli niczego z krwią; nie będziecie wróżyć ani czarować. Nie będziecie strzygli włosów dookoła waszej głowy, ani nie podcinaj końca swej brody". Tatuaże także zostały zakazane, aczkolwiek Księga Kapłańska nie wspomina nic o piercingu. Odniesienia do tych wszystkich praktyk powinny zostać zakodowane jako „zakazane zachowania" i być może powinny zostać opatrzone dodatkowymi kodami (pamiętaj o „regułach żywieniowych").

Mam nadzieję, że ten krótki rzut oka na możliwą analizę da ci pewne wyobrażenie o procesie tworzenia kodów oraz ich stosowania. Powinieneś także zacząć dostrzegać, w jaki sposób takie kodowanie umożliwi ci lepsze zrozumienie przesłania, leżącego u podstaw analizowanego tekstu i pozwoli na wybieranie takich danych, jakich będziesz potrzebował.

Zapisywanie notatek

W metodzie teorii ugruntowanej (MTU) proces kodowania zakłada coś więcej niż prostą kategoryzację kawałków tekstu. Kiedy kodujesz dane, powinieneś także korzystać z techniki **sporządzania notatek** (*memoing*) – zapisywania notatek (*memos*) dla siebie lub innych osób, zaangażowanych w realizację danego projektu badawczego. Część z tego, co zapiszesz podczas analizy, może znaleźć się w twoim końcowym raporcie; większość tych notatek będzie co najmniej inspiracją do tego, co w nim napiszesz.

W MTU takie notatki mają szczególne znaczenie. Strauss i Corbin (1990, s. 197f) rozróżniają trzy ich rodzaje: *notatki kodowe*, *notatki teoretyczne* i *notatki operacyjne*.

Notatki kodowe określają nazwy kodów i ich znaczenie. Jest to szczególnie ważne dlatego, że większość pojęć, stosowanych we wszystkich naukach społecznych, poza swoim technicznym znaczeniem ma także znaczenie w języku potocznym. Zatem zapisanie jednoznacznie tego, co rozumiemy pod danym kodem, używanym w naszej analizie, jest niezbędne. Na przykład w analizie Księgi Kapłańskiej chcielibyśmy sporządzić notatki kodowe związane ze znaczeniem pojęcia „obrzydliwość" oraz tego, jak będziemy wykorzystywać ten kod w naszej analizie tekstu.

Notatki teoretyczne dotyczą wielu różnych zagadnień: refleksji nad wymiarami i głębszym znaczeniem pojęć, związków między pojęciami, tez teoretycznych itp. Każdy z nas poświęca jakiś czas na rozmyślania o naturze rzeczy, stara się coś przemyśleć, zrozumieć. W analizie danych jakościowych ważne jest, aby spisywać te myśli, nawet jeśli je później odrzucimy jako nieprzydatne. Notatki takie będą miały bardzo różną długość, aczkolwiek powinieneś starać się zawierać w nich pojedyncze główne myśli, tak aby później można było je sortować i organizować. W analizie Księgi Kapłańskiej jedna z teoretycznych notatek może omawiać sposób, w jaki większość zakazów dotyczy *implicite* zachowania *mężczyzn*, wspominając o kobietach tylko okazjonalnie.

Notatki operacyjne są to notatki, które dotyczą przede wszystkim zagadnień metodologicznych. Niektóre z nich będą zwracać uwagę na okoliczności zbierania danych, które mogą okazać się istotne dla zrozumienia tych danych w przyszłości. Inne zawierać będą uwagi, dotyczące przyszłego gromadzenia danych.

Sporządzanie tego typu notatek następuje podczas całego procesu zbierania danych i ich analizy. Myśli wymagające zanotowania będą pojawiać się podczas powtórnego odczytywania zapisanych już notatek i transkrypcji, kodowania kawałków tekstu lub też omawiania projektu badawczego z innymi osobami. Spisanie notatek, gdy tylko coś przyjdzie ci na myśl, tak szybko jak to jest możliwe – to wspaniała idea i bardzo dobry zwyczaj.

John i Lyn Loflandowie (1995, s. 193f) mówią o zasadach sporządzania notatek nieco inaczej. Opisują notatki, które zbliżają się do końcowej fazy pisania raportu z badań. *Podstawowa notatka* jest:

> „uszczegółowionym, analitycznym odzwierciedleniem pewnej względnie konkretnej kwestii. Zależnie od skali projektu badacz może napisać tylko jedną notatkę lub też może sporządzić ich kilka tuzinów, a nawet więcej. Zbudowane z wybranych kodów i kodowań, są one najbardziej podstawową, prozaiczną częścią projektu".

Notatka porządkująca opiera na kilku notatkach podstawowych. Takie notatki prezentują kluczowe wątki w analizie danych. Podstawowe notatki tworzymy tak, jak pojawiają się one w naszym umyśle – ani bez jakiegoś określonego z góry sensu, ani też bez szczególnego porządku, notatki porządkujące natomiast sporządzamy jako swego rodzaju próbę odkrycia lub nadania sensu analizowanym danym. Notatka porządkująca wiąże w jedną całość grupę bliskich sobie notatek podstawowych. Dany projekt badawczy może skłaniać badacza do stworzenia wielu notatek porządkujących uwzględniających różne aspekty projektu.

Na koniec, *notatka integrująca* łączy razem kilka notatek porządkujących, porządkując cały projekt badawczy. Prezentuje nam spójną i całościową historię, osadzając ją w kontekście teoretycznym. Jednak w ramach każdego

projektu badawczego istnieje wiele sposobów na osiągnięcie tego rodzaju zintegrowania. Dlatego też rezultatem analizy danych może być kilka notatek integrujących.

Zauważmy, że jeśli myślimy o pisaniu jako o procesie linearnym – mającym swój początek i prowadzącym do konkluzji – to sporządzanie notatek jest czymś całkowicie odmiennym. Czynność ta może być scharakteryzowana jako tworzenie chaosu, a następnie jako znajdowanie w nim porządku.

Jeśli chcesz dalej zagłębiać się w ten proces, zajrzyj do prac cytowanych w trakcie powyższego omówienia oraz wymienionych na końcu tego rozdziału. Liczne informacje na ten temat możesz znaleźć w Internecie. Na przykład, by poznać reguły sporządzania notatek według Barneya Glasera, wejdź na stronę http://www.vlsm.org/gnm/gnm-gtm3.en.html. Ostatecznie praktyka będzie najlepszą nauką tego procesu. Nawet jeśli nie realizujesz żadnego projektu badawczego w chwili obecnej, możesz spróbować kodować swoje notatki z wykładów. Możesz też zacząć prowadzić dziennik i zakodować go.

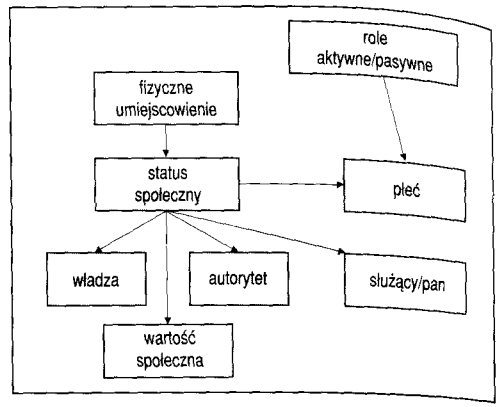
Tworzenie mapy pojęciowej

Teraz już powinno być jasne, iż analizując dane jakościowe spędzisz wiele czasu przelewając swoje myśli na papier (lub też do plików komputerowych), choć proces ten nie ogranicza się wyłącznie do tekstu. Niejednokrotnie możemy lepiej przemyśleć związki między pojęciami, przedstawiając je w formie graficznej. Proces ten nazywany jest tworzeniem **mapy pojęciowej**. Niektórzy badacze uważają, że przedstawienie wszystkich zasadniczych pojęć na jednej kartce papieru może być przydatne, choć są i tacy, którzy wykorzystują do tego liczne kartki, tablice, tablice magnetyczne, ekrany komputerów i inne media. Rysunek 13.3 pokazuje, jak możesz myśleć o niektórych pojęciach wynikających z badań Goffmana dotyczących płci i reklamy. (Obrazek został stworzony przy zastosowaniu programu komputerowego „Inspiration", tworzącego mapy pojęciowe).

Okazjonalnie wiele z omawianych tutaj tematów ma zastosowanie zarówno w badaniach ilościowych, jak i jakościowych. Oczywiście

Rysunek 13.3

Przykład mapy pojęciowej



tworzenie map pojęciowych jest odpowiednie dla obu typów analizy. Różne typy notatek także będą przydatne w obu typach analizy, a omówienie kodowania doskonale odnosi się do kodowania pytań otwartych w kwestionariuszu, którego celem jest analiza statystyczna. (Ponownie przyjrzymy się kodowaniu w następnym rozdziale, dotyczącym kwantyfikacji danych).

Po odnotowaniu przenikania się technik ilościowych i jakościowych, stosowne wydaje się w tym momencie zwrócenie uwagi na narzędzie, które pierwotnie związane było z badaniami ilościowymi, ale obecnie sprawdza się również w analizach jakościowych. Jest to komputer osobisty.

Programy komputerowe dla danych jakościowych

Nastanie ery komputerów – osobistych i stacji roboczych – stało się dobrodziejstwem dla badań ilościowych, gdyż umożliwiło szybkie obliczanie bardzo złożonych operacji statystycznych. Znaczenie komputerów dla badań jakościowych dostrzeżono nieco później. Niektórzy badacze jakościowi szybko zaadaptowali podstawowe funkcje komputerów dla celów nieobliczeniowych, lecz programiści potrzebowali nieco więcej czasu, by zaspokoić szczególne potrzeby badań jakościowych jako takich. Jednak mamy już

w chwili obecnej kilka dobrych programów komputerowych odpowiednich do tych celów.

Rozpocznijmy ten podrozdział od krótkiego przeglądu kilku sposobów wykorzystania podstawowych funkcji komputerów w badaniach jakościowych. W rozdziale 10 omawialiśmy już znaczenie komputera jako urządzenia służącego sporządzaniu notatek. Prawdopodobnie tylko ten, kto spędził wiele godzin, pracując z kalką maszynową i korektorem, może w pełni docenić wielkość komputerów w tym względzie. Określenia takie jak: „łatwa edycja" i „łatwe powielanie" nie oddają w pełni zakresu tego usprawnienia.

Wychodząc poza podstawowe funkcje zapisywania i przechowywania danych, proste edytory tekstu mogą być wykorzystane podczas niektórych analiz. Komendy „znajdź" i „szukaj" dają możliwość odnalezienia fragmentów tekstu, zawierających słowa kluczowe. Lub, idąc krok dalej, możesz dopisać słowa-kody obok wybranych fragmentów notatek, tak abyś później mógł wyszukać te hasła.

Bazy danych i arkusze kalkulacyjne także mogą być wykorzystywane podczas analizy danych jakościowych. Rysunek 13.4 jest prostą ilustracją tego, jak można opracować w tabeli kilka wersów z Księgi Kapłańskiej. Trzy kolumny z lewej strony reprezentują trzy wcześniej omawiane pojęcia. Znak „x" oznacza, że dany fragment z prawej strony zawiera to pojęcie. Jak widać, fragmenty te są uporządkowane w taki sposób, że wybrane zostały te, które dotyczą kary poprzez uśmiercenie. Inna prosta komenda „sortuj" pozwoli nam zebrać wszystkie te fragmenty, które dotyczą seksu, homoseksualizmu lub dowolnego z pozostałych zakodowanych pojęć.

Ta krótka ilustracja powinna dać ci pewne wyobrażenie o możliwościach zastosowania łatwo dostępnych programów jako narzędzi jakościowej analizy danych. Na szczęście w chwili obecnej istnieje wiele programów specjalnie stworzonych do tego celu. Poniżej zamieszczona jest doskonała lista, przygotowana przez socjologów z Uniwersytetu w Surrey w Anglii (http://www.soc.surrey.ac.uk/sru/SRU1.html):

The Etnograph
HyperQual
HyperResearch
HyperSoft
NUD*IST
QUALPRO
QUALOG
Textbase Alpha
SONAR
Atlas-ti

Rysunek 13.4

Wykorzystanie arkusza kalkulacyjnego w analizie jakościowej

Seks	Homoseks	Śmierć	Wers	Fragment
X	X	X	20:13	„Mężczyzna, który obcuje cieleśnie z mężczyzną tak jak z kobietą, popełnia obrzydliwość; obaj poniosą śmierć; krew ich spadnie na nich".
X		X	20:12	„Mężczyzna, który obcuje cieleśnie ze swoją synową, razem z nią poniesie śmierć; dopuścili się kazirodztwa, krew ich spadnie na nich".
X		X	20:15	„Mężczyzna, który obcuje cieleśnie ze zwierzęciem, poniesie śmierć; zwierzę także zabijcie".
		X	20:09	„Ktokolwiek złorzeczy swemu ojcu i swojej matce, poniesie śmierć. Złorzeczył swemu ojcu i swojej matce, przeto krew jego spadnie na niego".
		X	20:02	„Ktokolwiek z synów izraelskich i z obcych przybyszów mieszkających w Izraelu odda kogoś ze swego potomstwa Molochowi, poniesie śmierć; miejscowa ludność ukamienuje go".
X	X		18:22	„Nie będziesz cieleśnie obcował z mężczyzną jak z kobietą. Jest to obrzydliwością".

Wskazana powyżej strona internetowa dostarcza również krótkich opisów każdego z wymienionych programów, łącznie z ceną i adresem kontaktowym, pod którym można je nabyć.

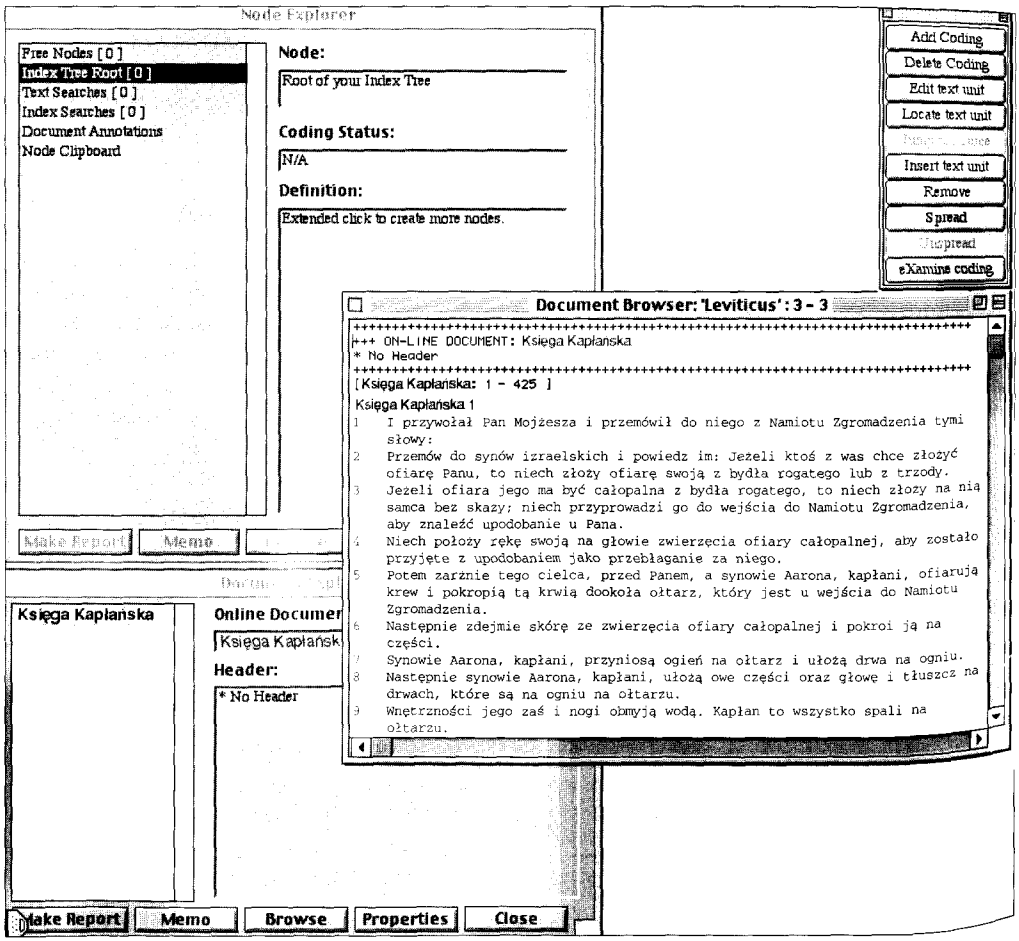
*Księga Kapłańska widziana oczami programu NUD*IST*

Przyjrzyjmy się bliżej, jak radzi sobie z analizą danych jakościowych NUD*IST (Nonnumeric Unstructured Data, Index Searching and Theorizing) – jeden ze wspomnianych programów. Chociaż każdy z nich ma pewne cechy odmienne i różne podejście do analizy, NUD*IST jest jednym z najbardziej popularnych programów. Oferuje reprezentatywną ilustrację możliwości wykorzystania programów tego typu. Zaczniemy od krótkiego zbadania Księgi Kapłańskiej, a następnie przejdziemy do zbadania projektu badawczego, koncentrującego się na doświadczeniach kobiet reżyserów filmowych.

Choć jest możliwe bezpośrednie wpisywanie danych tekstowych do programu NUD*IST w celu ich zakodowania, najczęściej mamy do czynienia z sytuacją, w której materiały już istniejące – jako notatki terenowe lub też, w naszym przypadku, wersy z Księgi Kapłańskiej – są *importowane* do programu. Polecenia zawarte w menu programu sprawiają, że jest to czynność łatwa, aczkolwiek tekst musi być w uproszczonym formacie *tekst*

Rysunek 13.5

Sposób prezentacji materiałów tekstowych w programie NUD*IST



wym (tzn. niesformatowany przez edytor tekstu lub inny rodzaj formatowania).

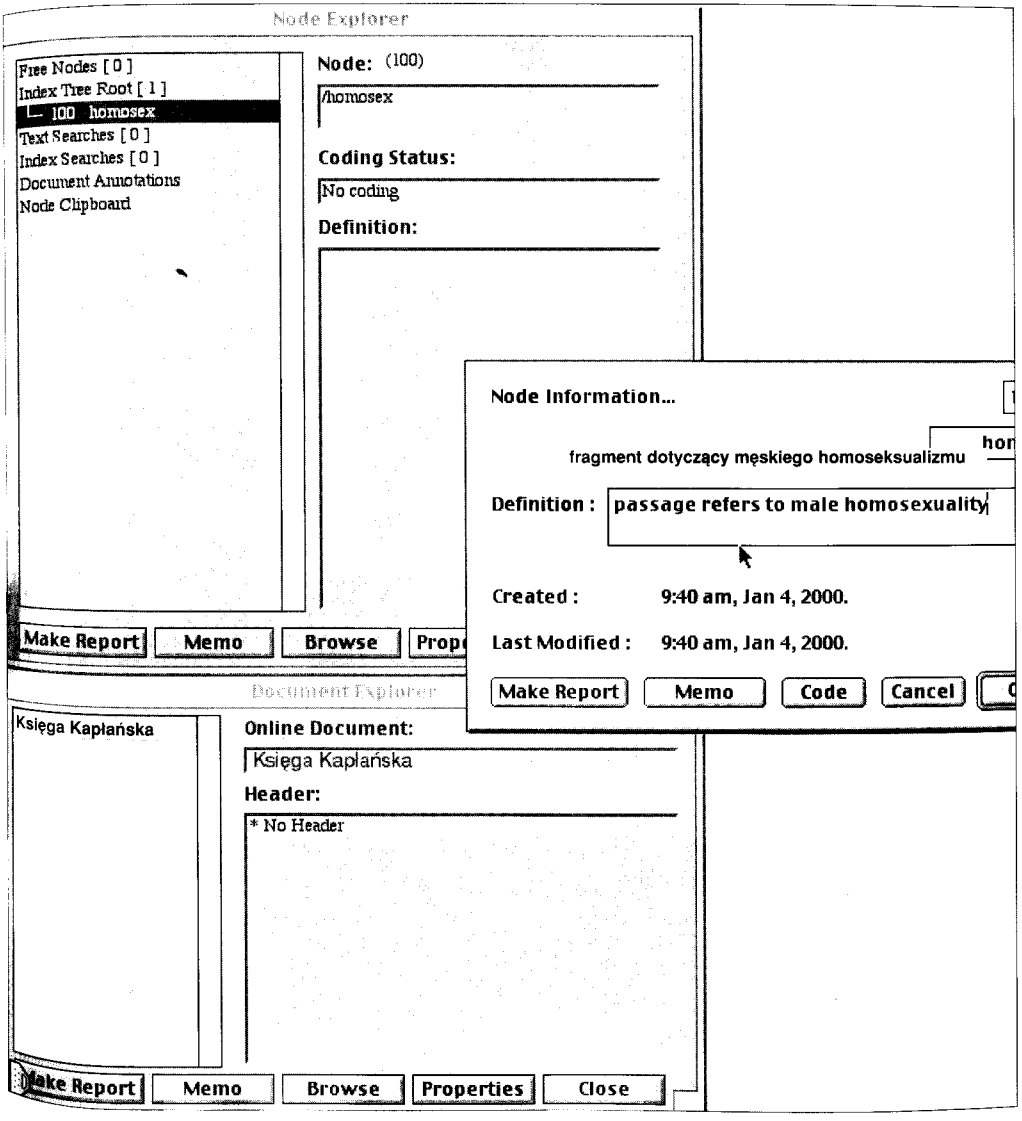
Rysunek 13.5 pokazuje, jak wygląda tekst w programie NUD*IST. Dla zilustrowania tego fragmentu wykorzystałem wersję programu dla Macintosha. Wersję dla Windows wykorzystam dla zilustrowania fragmentów dotyczących reżyserów filmowych. Dzięki temu będziesz mógł dostrzec różnice i podobieństwa między obydwiema platformami.

Aby zobaczyć dokument, wybierz jego nazwę w oknie „Document Explorer" i kliknij „Browse". Można zmienić rozmiar okna tekstowego oraz przesuwać je po ekranie monitora, w zależności od twojego gustu. Zauważ zestaw przycisków w prawym górnym rogu ilustracji. Pozwalają one wybrać fragment tekstu w celu jego edycji, kodowania lub dla innych operacji.

Teraz stwórzmy kod pojęcia: „homosex". Będzie to skrót odnoszący się do homoseksua-

Rysunek 13.6

Tworzenie kodu „homosex"



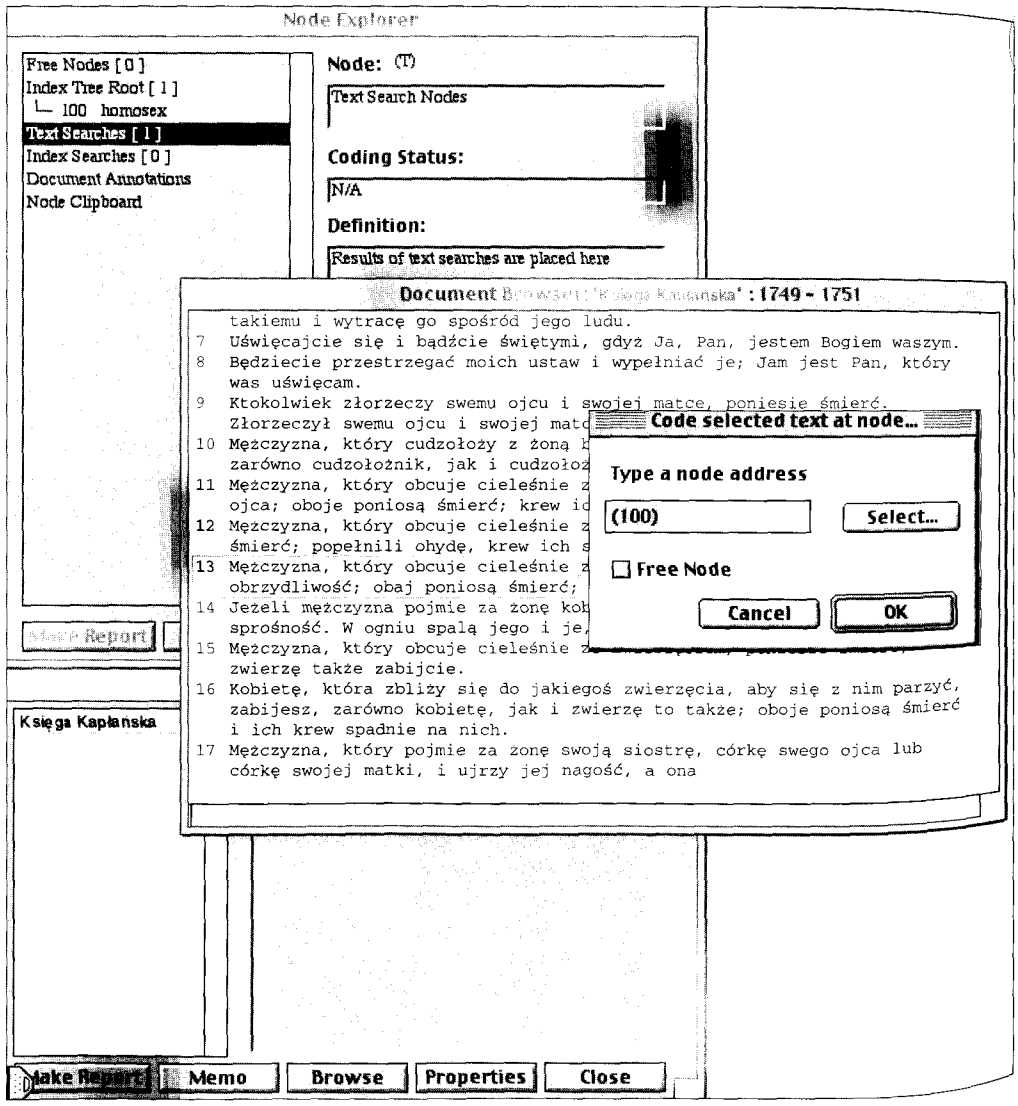
lizmu mężczyzn. Rysunek 13.6 pokazuje, jak wygląda tworzenie takiego kodu pojęciowego.

Jeśli stworzyliśmy kody dla naszych pojęć, możemy wykorzystać je do kodowania materiałów tekstowych. Rysunek 13.7 ilustruje, jak się to robi. W wyszukiwarce tekstowej możesz zobaczyć, że wybrany został wers 22:13 (wers ten znajduje się w ramce). Po zrobieniu tego, kliknij przycisk „Add Coding" (nie pokazany na tej ilustracji). Sprawia to, że komputer prosi nas o określenie odpowiedniego kodu. Najprostszym sposobem udzielenia odpowiedzi jest kliknięcie przycisku „Browse", za pomocą którego zaprezentowana zostanie lista obecnie używanych kodów. W tym przykładzie wybrałem „homosex" i wprowadziłem identyfikator tego kodu (100).

Gdy materiał tekstowy zostanie zakodowany, program ten może być wykorzystany do jego analizy. W prostym przykładzie możemy

Rysunek 13.7

Kodowanie fragmentu tekstu



zebrać razem wszystkie fragmenty zakodowane jako „homosex". Pozwoli to nam zobaczyć je wszystkie razem oraz na poszukać podobieństw i różnic między nimi.

Rysunek 13.8 pokazuje, jak NUD*IST zbiera razem fragmenty dotyczące męskiego homoseksualizmu. Aby to zrobić, jedyne co powinieneś uczynić, to wybrać nazwę kodu w oknie „Node Explorer" i kliknąć przycisk „Make Report".

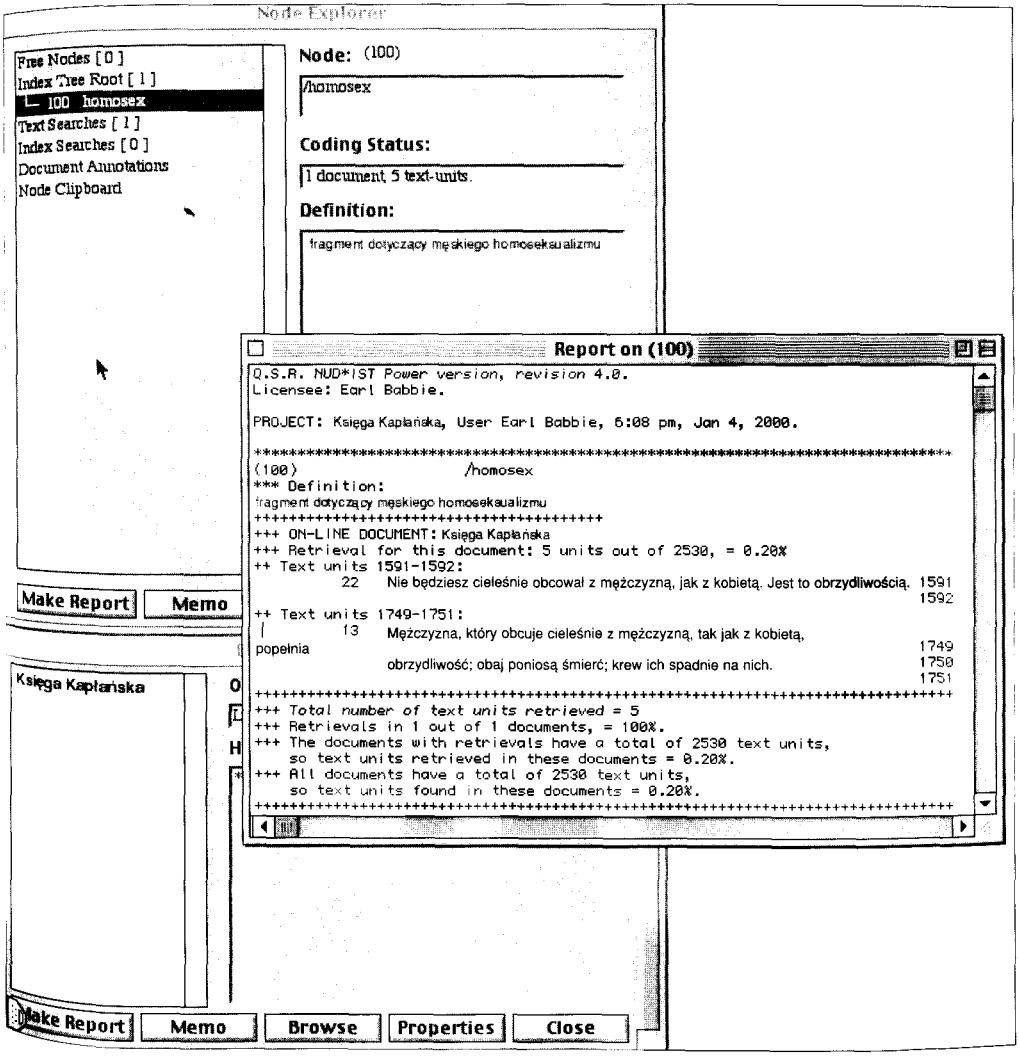
Ten prosty przykład pokazuje otwierające się możliwości, stworzone przez program specjalnie przeznaczony dla potrzeb analizy danych jakościowych. Aby zobaczyć jego bardziej zaawansowane zastosowanie, przejdźmy do innego przykładu badań.

Sandrine Zerbib: zrozumieć kobietę reżysera

Sandrine Zerbib jest Francuzką, socjologiem zainteresowanym zrozumieniem specyficznych trudności, jakich doświadczają kobiety, wkraczające w zdominowany przez mężczyzn świat reżyserii. Dla zbadania tego problemu prze-

Rysunek 13.8

Prezentacja raportu dotyczącego kodu „homosex"

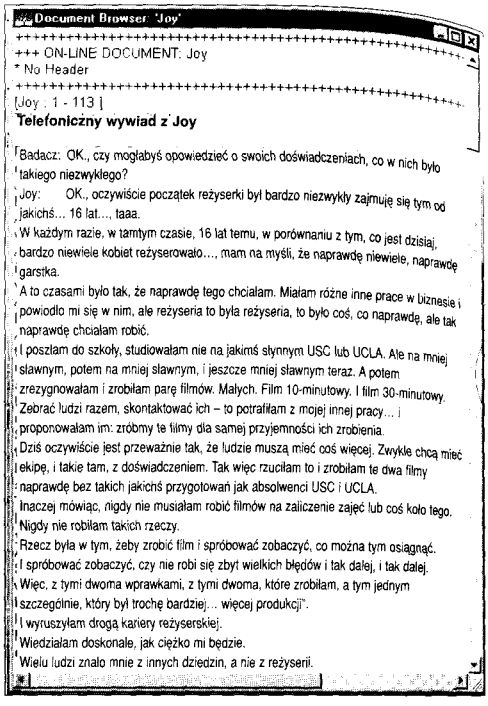


prowadziła 30 wywiadów pogłębionych z kobietami reżyserami. Zgromadziwszy wiele godzin zapisanych wywiadów, postanowiła wykorzystać program NUD*IST jako narzędzie analizy. Spójrzmy na przebieg tego procesu, opisany jej własnymi słowami.

„Większość oprogramowania, przeznaczonego dla analizy jakościowej, pozwala badaczom na jednoczesną analizę kilku wywiadów, zebranych przez różne osoby je przeprowadzające. Jednak dla mnie bardziej efektywne było rozpoczęcie pracy od importu tylko jednego wywiadu do NUD*IST. Ponieważ trzeba dokonać transkrypcji wywiadów lub co najmniej przeczytać je na wstępie, można wybrać od razu taki wywiad, który wydaje się najbardziej obiecujący. Powinno się sobie wierzyć, ponieważ stajemy się ekspertami w tym, co w danym momencie badamy, a także dlatego, że porównywanie i zestawianie wywiadów powinno nam pomóc uchwycić istotę tego, jak dokładna jest nasza analiza.

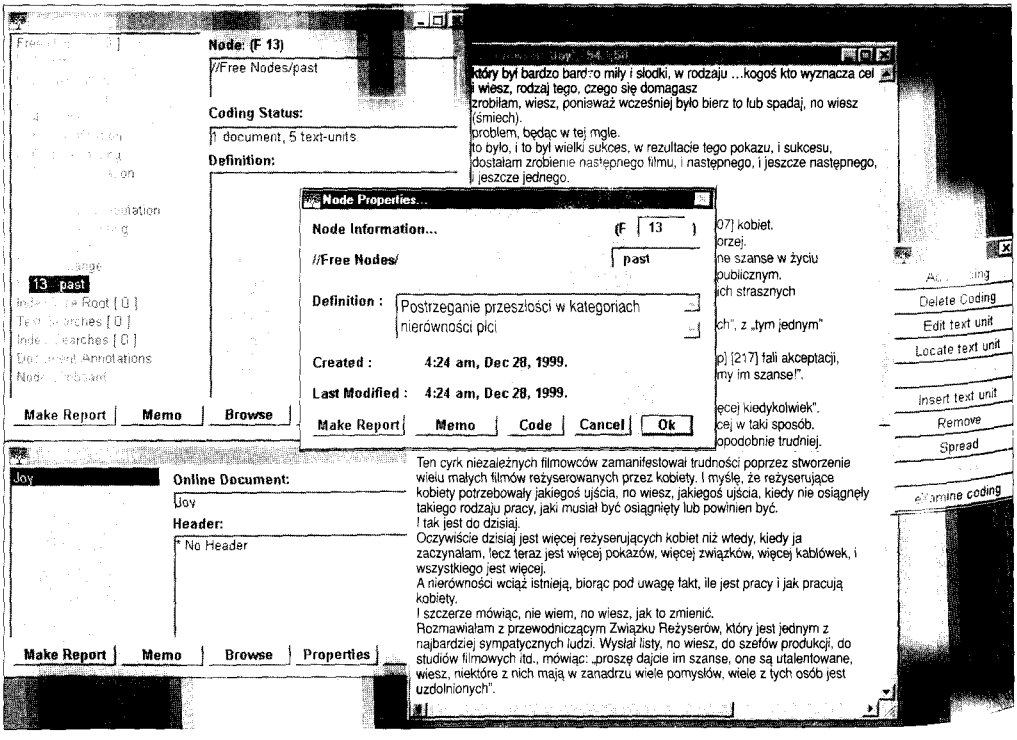
Rysunek 13.9

Tekst wywiadu przeprowadzonego z Joy



Rysunek 13.10

Tworzenie kodu „przeszłość"



Po zebraniu około 30 wywiadów z kobietami twórcami filmowymi zorientowałam się, jakie są główne wątki, ponieważ pojawiły się one w każdym z przeprowadzonych wywiadów. Niemniej jednak potrzebowałam narzędzia do syntezy tych wszystkich stron zapisanych wywiadów. Na początek wybrałam wywiad, który sama przeprowadziłam z Joy. Gdzieś tam sobie zanotowałam, aby wykorzystać ten właśnie wywiad jako punkt wyjścia. Wydawało się, że jako starszy reżyser filmowy ma ona jasno sprecyzowane poglądy.

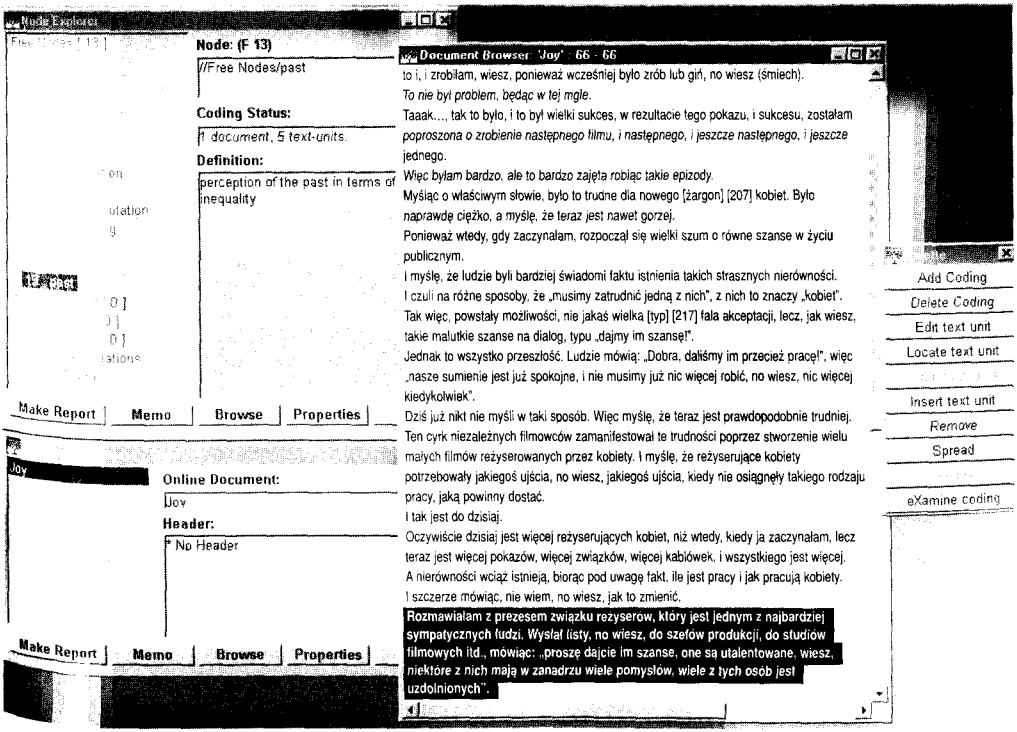
Rysunek 13.9 przedstawia mój wywiad z Joy, który został zaimportowany jako plik tekstowy. (W prezentowanym oknie widoczna jest tylko część tego pliku).

Podczas kodowania pamiętaj, że NUD*IST czyta tylko dokumenty tekstowe. Nie ma potrzeby, abyś przesadzał z transkrypcją swoich wywiadów; wszystkie formaty zostaną wykasowane. W tym momencie jesteś już gotowy, by cieszyć się procesem kodowania. Możesz sobie po prostu zaznaczać słowa, zdania, akapity i dodawać do nich oznaczenia (lub kody). Pierwszym krokiem jest stworzenie «wolnych oznaczeń» (*free nodes*), czyli takich oznaczeń, które będą od siebie niezależne. To, jak obszerny fragment tekstu przypiszesz do poszczególnego oznaczenia kodowego, zależy tylko od ciebie. Jednak miej na uwadze, że będziesz korzystać z tych cytatów podczas pisania raportu z badań. Powinien on być przekonujący. Pamiętaj, że chcesz zdekomponować cały wywiad. Staraj się niczego nie utracić. Łatwiej jest zrezygnować z użycia jakiegoś cytatu, gdy znajdzie się inny, lepszy, niż nie mieć nic, ponieważ nie było się wystarczająco konsekwentnym w rozbiorze wywiadu.

Kiedy tworzy się oznaczenie lub kody, początkowo ma się chęć, by korzystać z bardziej ogólnych kategorii, obejmujących inne możliwe cytaty. Lecz chce się także zachować dość szczegółowości, by system kodowy był trafny. Na przykład na rysunku 13.10 wprowadziłam wolne oznaczenie «przeszłość» (*past*), gdyż moja roz-

Rysunek 13.11

Kodowanie fragmentu wywiadu

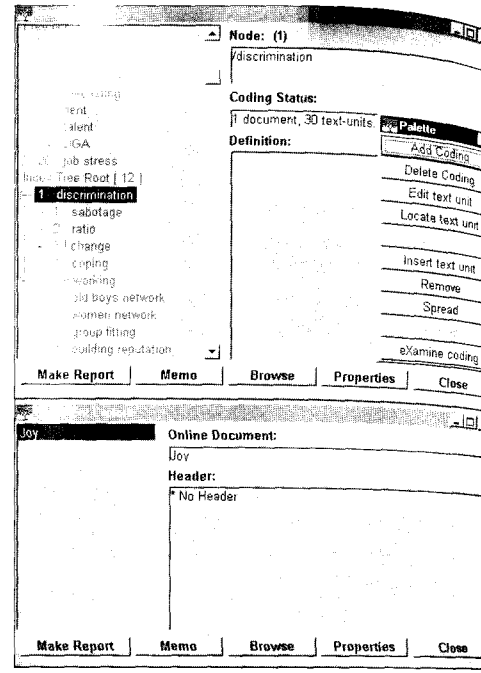


mówczyni odnosiła się do przeszłości jako do czasów, wymagających od kobiet pragnących zostać reżyserami podejmowania wielkich wyzwań. Bardzo niewiele kobiet było wtedy reżyserami, znacznie mniej niż obecnie. Zdecydowałam się dodać definicję do tego oznaczenia, tak bym mogła pamiętać, dlaczego użyłam słowa «przeszłość» jako oznaczenia. Przewidywałam także stworzenie innego wolnego oznaczenia, które nazwałabym «dzisiaj». Dzięki temu mogłam przesunąć oznaczenie «zmiana» do pnia drzewa indeksu i umieścić «przeszłość» i «dzisiaj» jako podkategorie «zmiany».

Na rysunku 13.11 mam zaznaczony fragment, poruszający kilka spraw. Joy mówi o Amerykańskim Zrzeszeniu Reżyserów (*Directors Guild of America – DGA*), czyli inaczej związku reżyserów, oraz bardziej szczegółowo o staraniach podejmowanych przez jego prezesa. Ponadto wyraża swoje odczucia dotyczące nierówności płci. Według niej samo posiadanie talentu w Hollywood nie wystarcza, jeśli są tam uprze-

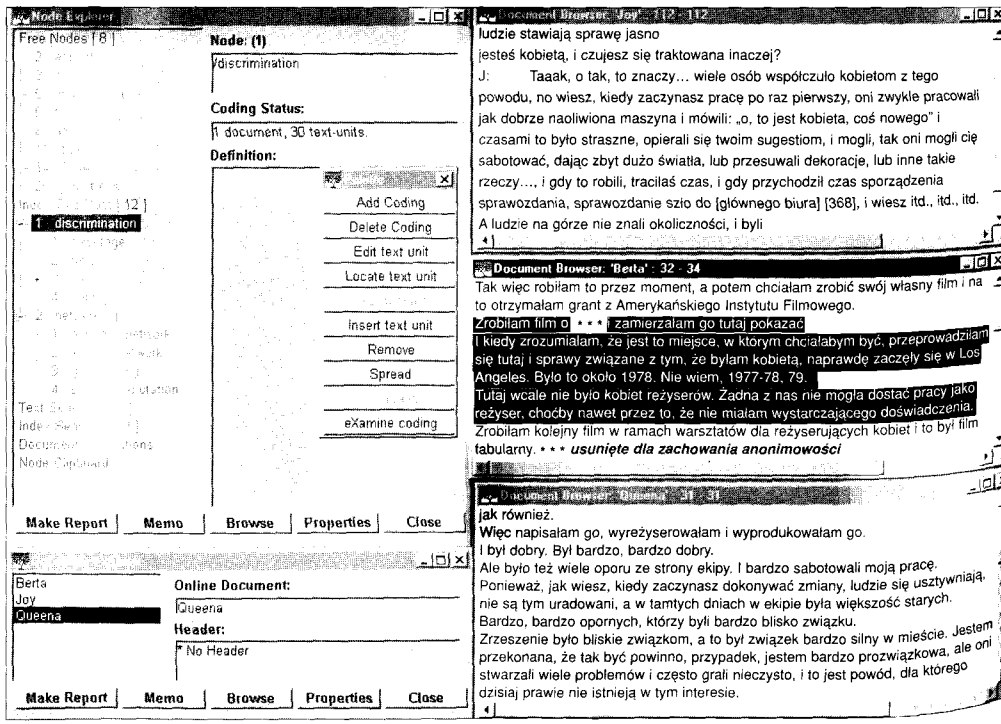
Rysunek 13.12

Tworzenie indeksu w formie „drzewa"



Rysunek 13.13

Włączenie dwóch dodatkowych przypadków do analizy



dzenia wobec kobiet. Zdecydowałam się dodać dwa oznaczenia do tego cytatu: «DGA», które musiałam wtedy stworzyć, oraz «dyskryminacja», które miałam już wcześniej.

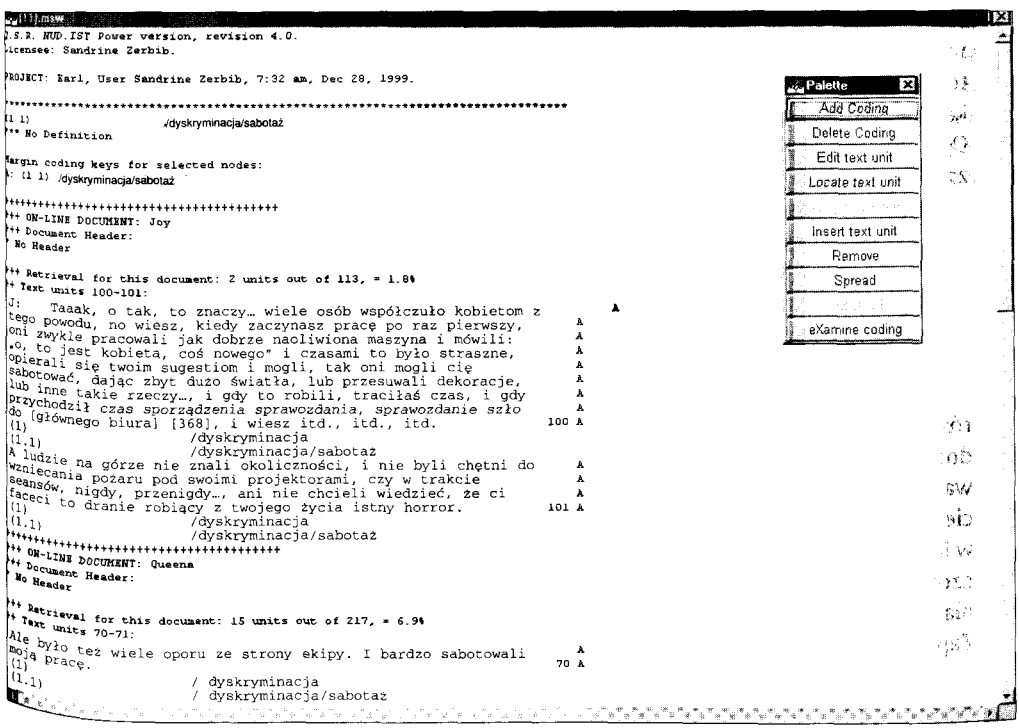
Na rysunku 13.12 starałam się przekształcić niektóre wolne oznaczenia w drzewa indeksowe (*index trees*). Oprogramowanie jest na tyle elastyczne, że mogłam przesuwać oznaczenia, zmieniać ich nazwy lub oglądać, które cytaty kryją się pod każdym z oznaczeń. Można przypisać odmienne oznaczenie do cytatu, który się wcześniej błędnie zakodowało. Lepiej jest zacząć od wolnych oznaczeń, zanim zacznie się budować hierarchię kodów (czy inaczej drzewo), ponieważ zrozumienie tego, jak kategorie są ze sobą powiązane, wymaga czasu i cierpliwości. Kodowanie kolejnych wywiadów powinno pomóc uporządkować system kodowy.

Rysunek 13.13 ilustruje moją decyzję, dotyczącą zaimportowania dwóch kolejnych wywiadów, z Bertą i z Queeną. Mogłam przeglądać wszystkie trzy wywiady na jednym ekranie. Ponieważ był to wciąż początek procesu analizy, zdecydowałam się analizować te dwa nowe wywiady jeden po drugim. To był moment, w którym zaczęłam poszukiwać znaczeń, zaczęłam poszukiwać wzorów. NUD*IST pozwolił mi zanotować liczbę wystąpień każdego z oznaczeń, które były przypisane do cytatu, nie tylko w wywiadzie z Joy, ale także w wywiadach z Bertą i Queeną. Gdy stwierdziłam, że kilka oznaczeń jest często przypisywanych do cytatu, mogłam za pomocą analizy jakościowej stwierdzić, które oznaczenia częściej nakładają się na inne.

Jedną z moich pierwszych obserwacji było to, że Joy i Queena dość często używały terminu «sabotaż». Zdecydowałam się sporządzić raport, który miał dokonać syntezy wszystkich cytatów, do których dopisałam oznaczenie «sabotaż». Rysunek 13.14 prezentuje pierwszą stronę raportu stworzonego przez program NUD*IST. Program wyszukał wszystkie cytaty opatrzone oznaczeniem «sabotaż», będącym podkategorią «dyskryminacji», we wszystkich dokumentach wprowa-

Rysunek 13.14

Analiza oznaczenia „sabotaż"



dzonych do komputera. Do każdej jednostki tekstu został przypisany numer, co pozwoliło mi zrobić krok wstecz i zobaczyć poszczególne cytaty w kontekście całego dokumentu.

Taka procedura jest tylko jedną z wielu możliwości oferowanych przez program. Możesz spędzić więcej czasu, ucząc się funkcjonowania tego oprogramowania, zanim się do niego przekonasz. To, co dla mnie było użytecznym narzędziem, wcale nie musi być takie dla ciebie. Na rynku istnieje wiele rodzajów oprogramowania dla analizy badań jakościowych; staraj się dobrać takie, na którym będzie ci się dobrze pracować".

Główne zagadnienia

- Analiza jakościowa jest nieliczbowym badaniem i interpretacją obserwacji.
- Analiza jakościowa zakłada stałe poruszanie się między teorią a analizą. Podczas analizy danych staramy się odkryć prawidłowości, takie jak zmiany w czasie lub możliwe związki przyczynowe między zmiennymi.
- Przykładowe podejścia do odkrywania i wyjaśniania takich prawidłowości to metoda teorii ugruntowanej (MTU), semiotyka i analiza konwersacyjna.
- Opracowywanie danych jakościowych jest zarówno sztuką, jak i nauką. Trzy główne narzędzia przygotowywania danych do analizy to kodowanie, zapisywanie notatek, tworzenie mapy pojęciowej.
- W przeciwieństwie do standaryzowanych jednostek, używanych w kodowaniu dla potrzeb analizy statystycznej, w analizie jakościowej jednostki kodowania mogą różnić się między sobą w ramach jednego dokumentu. Choć kody mogą być wyprowadzone z badanej teorii, badacze częściej korzystają z kodowania otwartego, w którym kody same „narzucają się" badaczowi w wyniku rozpatrywania i sprawdzania danych.
- Zapisywanie notatek jest stosowane na różnych etapach opracowywania danych w celu uchwycenia znaczenia kodów, pomysłów teoretycznych, wstępnych wniosków oraz innych myśli, które mogą być przydatne podczas analizy danych.
- Tworzenie mapy pojęciowej polega na wykorzystaniu diagramów do graficznego badania związków między danymi.
- Parę programów komputerowych, takich jak NUD*IST, zostało specjalnie stworzonych dla wsparcia badaczy podczas analizy danych jakościowych. Dodatkowo badacze mogą usprawniać swoją pracę, wykorzystując możliwości popularnego oprogramowania, takiego jak edytory tekstu, bazy danych, arkusze kalkulacyjne.

Kluczowe pojęcia

analiza jakościowa
analiza zorientowana na zmienną
analiza zorientowana na przypadek
przekrojowa analiza przypadków
metoda teorii ugruntowanej – MTU
metoda stałego porównywania
semiotyka
analiza konwersacyjna
kodowanie otwarte
zapisywanie notatek
mapa pojęciowa

Pytania i ćwiczenia

1. Przypomnij sobie badania Goffmana dotyczące reklam związanych z płcią; zbierz i przeanalizuj zestaw zdjęć reklamowych z czasopism lub gazet, tak aby zbadać występujące tam relacje między płcią a statusem.
2. Przejrzyj omówienie homoseksualizmu w Księdze Kapłańskiej i przedstaw sposoby możliwego ustrukturyzowania tych badań w przekrojowej analizie przypadków.
3. Wyobraź sobie, że prowadzisz przekrojową analizę przypadków na podstawie rewolucyjnych dokumentów, na przykład takich jak Deklaracja Niepodległości i Deklaracja Praw Człowieka i Obywatela (z czasów rewolucji francuskiej). Określ kluczowe pojęcia, które można zakodować w następującym zdaniu:

„Ilekroć wskutek biegu wypadków koniecznym się staje dla jakiegoś narodu, by zerwał więzy polityczne łączące go z innym narodem i zajął wśród potęg ziemskich oddzielne i równorzędne stanowisko, do którego upoważniają go prawa natury i jej Bóg, to właściwy respekt wymaga, aby naród ten podał powody, które zmusiły go do oderwania się"[3].

4. Napisz jedną notatkę kodową i jedną notatkę teoretyczną dotyczącą punktu 3.
5. Korzystając z biblioteki, Internetu lub strony InfoTrac, odszukaj raporty z badań, w których korzystano z analizy konwersacyjnej. Przedstaw podsumowanie głównych wniosków własnymi słowami.

Dodatkowe lektury

Denzin Norman K., Lincoln Yvonna S., *Handbook of Qualitative Research*, Sage, Thousand Oaks, Cal. 1994. Jest to bogate źródło informacji, dotyczących wielu aspektów badań jakościowych, zarówno w teorii, jak i w praktyce.

Glaser Barney G., Strauss Anselm L., *The Discovery of Grounded Theory for Qualitative Research*, Aldine, Chicago 1967. Klasyczne opracowanie teorii ugruntowanej, z praktycznymi wskazówkami, które są wciąż użyteczne.

Hutchby Ian, Wollfitt Robin, *Conversation Analysis: Principles, Practice and Applications*, Polity Press, Cambridge 1998. Wspaniały przegląd metody analizy konwersacyjnej. Książka ta omawia teorię stojącą za tą techniką, metody jej wykorzystania i możliwości jej zastosowań.

Strauss Anselm, Corbin Juliet, *Basics of Qualitative Research: Grounded Theory. Procedures and Techniques*, Sage, Newbury Park, Cal. 1990. Zaktualizowane omówienie teorii ugruntowanej stanowiące wyjątkowy przewodnik po kodowaniu i sporządzaniu notatek.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

[3] *Encyklopedia Historii Stanów Zjednoczonych Ameryki*, Andrzej Bartnicki, Krzysztof Michałek, Izabela Rusinowa (red.), Warszawa 1992, s. 67, 68 (przyp. red.).

14

Kwantyfikacja danych

Przegląd holograficzny

Wiele badań społecznych wymaga ilościowej (statystycznej) analizy danych, którą współcześnie wykonuje się zwykle za pomocą komputerów. Przygotowanie masy obserwacji do analizy komputerowej wymaga przełożenia danych na terminy ilościowe, które program komputerowy jest w stanie odczytać i przetworzyć. Przede wszystkim dane muszą być zakodowane, a błędy w kodowaniu usunięte w toku procedury zwanej czyszczeniem danych.

Wprowadzenie

Komputery w badaniach społecznych

Kodowanie
Tworzenie kategorii kodowych
Budowa książki kodowej

Sposoby kodowania i wprowadzania danych
Arkusze kodowe
Kodowanie brzegowe
Bezpośrednie wprowadzanie danych
Wprowadzanie danych przez ankieterów
Kodowanie na karty do skanera optycznego
Bezpośrednie użycie kart do skanera optycznego
Powiązanie z programem do analizy danych

Czyszczenie danych
Czyszczenie dotyczące dopuszczalnych kodów
Czyszczenie logiczne

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Komputer w ilościowych badaniach społecznych odgrywa podobną rolę, jak mikroskop w biologii czy teleskop w astronomii. Komputery pozwalają nam uzyskać obraz życia społecznego nieosiągalny w żaden inny sposób. Co więcej, wciąż odkrywamy możliwe metody wykorzystania komputerów w naukach społecznych. Główna ich zaleta polega na tym, że umożliwiają badaczom przetwarzanie i analizę olbrzymich ilości kwantyfikowalnych danych.

W poniższym rozdziale opisuję metody przekształcania danych, pozyskiwanych w naukach społecznych, na *format czytelny dla urządzeń elektronicznych* – czyli taki, który może zostać odczytany i przetworzony przez komputery i inne podobne urządzenia używane w **analizie ilościowej**. Gdybyś czytając tę książkę prowadził równolegle ilościowy projekt badawczy, twoje dane miałyby w tej chwili postać wypełnionych kwestionariuszy, arkuszy kodowych używanych w analizie treści itp. W stadium omawianym w tym rozdziale dane te zostałyby zapisane na twardy dysk, dyskietkę czy inny nośnik informacji w formacie czytelnym dla komputera.

Zarówno sprzęt komputerowy, jak i oprogramowanie zmieniają się nieustannie. Jednak zasada użytkowania komputera jest zasadniczo taka sama, niezależnie od tego, jakich urządzeń i programów akurat używamy. Z tego powodu rozdział ten zaczyna się od krótkiego przeglądu rozwoju zastosowania komputerów w badaniach społecznych. Wiedza o dawnych urządzeniach i technikach może się okazać użyteczna, ponieważ logika analizy danych daje się na ich przykładzie wyjaśnić łatwiej niż na przykładzie zaawansowanych technologii, podobnie jak działanie samochodu łatwiej wyjaśnić na przykładzie syrenki niż mercedesa. Wiedza ta pozwoli nam rozważyć podstawowe etapy kwantyfikacji danych, co umożliwi ci wykorzystanie narzędzi, które stawia do twojej dyspozycji komputer. Przekonasz się, że nagły rozwój i rozpowszechnienie się technik komputerowych sprawiły, że badania społeczne są dziś bardziej fascynujące niż kiedykolwiek wcześniej.

Komputery w badaniach społecznych

Z naszego punktu widzenia wykorzystanie technik komputerowych w naukach społecznych rozpoczęło się we Francji w 1801 roku. Mniej więcej w tym czasie pojawiły się też pierwsze zalążki nowoczesnych nauk społecznych. W 1801 roku Joseph-Marie Jacquard zapoczątkował w przemyśle tekstylnym rewolucję, która miała wywrzeć wpływ na najmniej nawet prawdopodobne zakamarki życia.

Aby ułatwić tkanie drobnych wzorów, Jacquard wynalazł automatyczne krosna, które pobierały instrukcje z perforowanych kart. W miarę przesuwania serii kart przez „czytnik" krosien, drewniane kołki przechodziły przez otwory w karcie, krosna zaś przekładały tę informację na splot tkaniny. Aby stworzyć nowy wzór, Jaquard musiał już tylko wykonać odpowiednie otwory w kolejnych kartach, aby wywołać odpowiednią reakcję krosien.

Rewolucyjny charakter tego wynalazku polegał na tym, że *informacja* (na przykład pożądany splot tkaniny) mogła być *zakodowana* i *przechowywana* w formie materialnej (otwory w kartach), a następnie odzyskiwana przez maszynę, która „odczytywała" informację i wykonywała oparte na niej czynności. W zasadzie właśnie to komputery robią po dziś dzień.

Kolejny etap w naszej wyrywkowej historii komputeryzacji wyznacza spis powszechny ludności Stanów Zjednoczonych z 1890 roku. Zgodnie z Konstytucją Stanów Zjednoczonych co dziesięć lat przeprowadzano powszechny spis ludności, począwszy od rejestru blisko czterech milionów obywateli amerykańskich w 1790 roku. W miarę wzrostu liczby ludności nowego narodu zwiększyło się też zadanie pomiaru jej liczebności. Spis z 1880 roku wykazał, że populacja liczy już ponad 62 miliony – ale stworzenie tabel zajęło pracownikom Biura Spisowego ponad dziewięć lat! Tymczasem zbliżał się kolejny spis. Stało się oczywiste, że przed przeprowadzeniem spisu w 1890 roku niezbędny był przełom technologiczny. Biuro poszukiwało rozwiązań.

Herman Hollerith, zatrudniony wówczas jako młody instruktor inżynierii w MIT (*Massa-*

chussets Institute of Technology), były pracownik Biura, który pracował przy spisie z 1880 roku, wpadł na pewien pomysł. Zaproponował mianowicie, by zaadaptować karty Jacquarda do liczenia ludności kraju. Po sporządzeniu lokalnych zestawień, miały one zostać naniesione na karty. Następnie maszyna licząca pomysłu Holleritha miała odczytać karty i określić liczbę ludności w całym kraju.

System Holleritha został przetestowany i uznany za najszybszy w porównaniu z innymi ofertami. W wyniku konkursu Biuro Spisowe kupiło od nowo powstałej Wytwórni Maszyn Liczących Holleritha sprzęt o wartości 750 000 dolarów, a raport o liczbie ludności w 1890 roku został sporządzony w ciągu sześciu tygodni. Firma Holleritha tworzyła nowe urządzenia, łączyła się z innymi pionierskimi firmami, aż w końcu zmieniła nazwę na International Business Machines Corporation (IBM). W latach 50. perforowane karty – zwane powszechnie *kartami* IBM – były stosowane do przechowywania i odczytywania danych z badań społecznych. Do dziś bywają – bardzo rzadko – wykorzystywane w tym celu.

Obecnie większość analiz danych przeprowadzana jest za pomocą komputerów, które przechowują dane w formacie elektronicznym, zamiast w formie kart perforowanych, papierowych taśm itp. Elektroniczny komputer potrafi nie tylko liczyć i segregować, lecz również prowadzić złożone obliczenia i tworzyć skomplikowane formy prezentacji wyników. Popularne programy badają jednocześnie kilkanaście zmiennych i obliczają ich rozmaite własności statystyczne. Dane przechowywane na dyskietkach, twardych dyskach czy CD-ROM-ach mogą być odczytywane znacznie szybciej niż dane z kart perforowanych, komputery zaś dokonują złożonych obliczeń statystycznych znacznie szybciej i dokładniej niż ludzie. Jednak, podobnie jak operacje dokonywane za pomocą kalkulatora można także przeprowadzić nieco bardziej pracochłonną metodą ręczną, również komputery wykonują tylko polecenie obliczenia tego, co statystycy liczyli dotąd na papierze. Pierwsze „komputery" były tak naprawdę ludźmi – były to pomieszczenia pełne matematyków, obliczających tory lotu pocisków. Kiedy maszyny przejęły tę funkcję, nazwano je także komputerami.

Dzisiejsze komputery przetwarzają naturalnie informacje na wiele innych sposobów poza prowadzeniem operacji matematycznych. Podobnie jak inni użytkownicy komputerów, także badacze społeczni korzystają z wielu „standardowych" programów, jak na przykład edytory tekstu. Pozwalają one opracowywać tekst i grafikę, sprawiają, że gromadzenie, poprawianie, redakcja i włączanie do tekstu informacji z rozmaitych źródeł są obecnie łatwiejsze niż niegdyś. Arkusze kalkulacyjne i bazy danych mogą być wykorzystywane do przechowywania ogromnych ilości danych słownych i liczbowych w formie rekordów, które można następnie dowolnie przeglądać, porządkować i analizować.

Oprócz tych i innych programów o ogólnym zastosowaniu, wiele istniejących obecnie programów zostało zaprojektowanych specjalnie do analizy danych z zakresu nauk społecznych. Załącznik I tej książki podaje instrukcje i nieco informacji na temat metod posługiwania się jednym z najbardziej popularnych programów tego rodzaju, a mianowicie SPSS (*Statistical Package for the Social Sciences*). Kiedy zapoznasz się w ogólnych zarysach z działaniem SPSS, zrozumiesz też ogólną logikę posługiwania się komputerem w analizie danych. Być może masz dostęp do innych programów komputerowych: ABtab, AIDA, A.STAT, BMDP, CRISP, DAISY, Data Desk, DATA-X, Dynacomp, INTER-STAT, MASS, MicroCase, Microquest, Microstat, Micro-SURVEY, Minitab, POINT FIVE, P-STAT, SAM, SAS, SNAP, STATA, STAT80, Statgraf, Statpak, StatPro, STATS PLUS, Statview, Survey Mate, SURVTAB, SYSTAT i TECPACS, by wymienić tylko niektóre z nich. Niezależnie od tego, jakiego programu używasz, reguły przechowywania i analizy danych są zasadniczo takie same.

Jeszcze w latach 70. wszystkie analizy komputerowe prowadzone były przez ogromne, niezwykle kosztowne komputery – zwane niekiedy dużymi komputerami (*mainframes*) – utrzymywane przez zcentralizowane ośrodki obliczeniowe. Do dziś niektóre analizy prowadzi się właśnie tam. Aby skorzystać z tego typu

urządzeń, badacz musiał dostarczyć dane wraz z instrukcjami co do typu pożądanej analizy. Instrukcje mogły być przygotowane w formacie SPSS, dane zaś badacz dostarczał do ośrodka obliczeniowego w formie kart lub taśm magnetycznych.

W opisanym powyżej procesie wprowadzono stopniowo dwa udogodnienia. Systemy wprowadzania danych roboczych na odległość umieściły czytniki kart i drukarki w zasięgu badaczy, co umożliwiło przysyłanie prac i otrzymywanie wyników w miejscach oddalonych od ośrodków obliczeniowych. *Współużytkowanie komputera* (*time sharing*) było kolejnym osiągnięciem. W pewnym sensie praca komputerów zawsze opierała się na tej właśnie zasadzie – wielu użytkowników korzystających z tego samego komputera. Początkowo jednak urządzenia obliczeniowe były wykorzystywane metodą seryjną: komputer przeliczyłby najpierw twoje dane, potem moje, potem kogoś innego. Bardziej zaawansowane komputery były jednak w stanie wykonywać kilka różnych zadań jednocześnie. Potrafiły one zapoznawać się z moim poleceniem, analizować twoje, a jednocześnie drukować wyniki kogoś innego. Nadzwyczajna szybkość komputerów w praktyce pozwalała im w toku dłuższej pracy radzić sobie od ręki z małymi operacjami (które mogą zajmować tysięczne części sekundy).

Ta właściwość pozwoliła komputerom wykonywać setki poleceń w tym samym czasie, stwarzając często wrażenie, że każdy ze zleceniodawców jest wyłącznym użytkownikiem komputera. Tego typu interakcje z komputerem wymagały na ogół korzystania z tzw. *terminali komputerowych* – urządzeń, które drukowały informacje na papierze albo wyświetlały operacje komputera na monitorze. Terminale jako takie nie były jednak komputerami, lecz urządzeniami służącymi do wchodzenia w kontakt z głównym komputerem, który naprawdę wykonywał zadanie.

Najprawdopodobniej znasz dobrze etap, który nastąpił po współużytkowaniu komputera. Niektóre sieci mają charakter lokalny i łączą komputery w danej instytucji czy organizacji (biuro, klasa szkolna, uniwersytet, przedsiębiorstwo). Dostęp do tych sieci mają tylko upoważnieni użytkownicy. W minionych dziesięcioleciach sieci – i współużytkowanie – stały się zjawiskiem globalnym za pośrednictwem międzynarodowych sieci znanych pod zbiorczą nazwą Internetu i sieci www (*World Wide Web*). Ty i twoi koledzy macie wspólny dostęp do serwera internetowego na waszej uczelni, wasz serwer porozumiewa się z innymi serwerami na całym świecie, co umożliwia niemal bezpośredni kontakt z kimkolwiek podłączonym do któregokolwiek z tych serwerów. Mimo że wynalazek ten stanowi rewolucję komunikacyjną, opiera się on nadal na wprowadzaniu, przechowywaniu, przetwarzaniu i rozpowszechnianiu informacji w formacie czytelnym dla urządzeń elektronicznych. Niniejsze, poprawione wydanie tej książki wymagało na przykład niezliczonych kontaktów drogą elektroniczną z kolegami na całym świecie. Przykłady badań i ilustracje dydaktyczne zaczerpnięte zostały z elektronicznych list dyskusyjnych. Fragmenty tekstu zostały wysłane drogą elektroniczną do recenzentów i krytyków. Załącznik B zawiera bardziej szczegółowe omówienie rozwijającego się świata cyberprzestrzeni.

Sieci komputerowe znacznie poszerzyły zasoby informacji dostępnych badaczom. Umożliwiają one na przykład badaczowi, siedzącemu w domu w Kalifornii, uzyskanie kopii zbioru danych, przechowywanego na komputerze w Londynie, i zanalizowanie tego zbioru za pomocą programu przechowywanego na komputerze w Teksasie. Wyniki analizy mogłyby potem wrócić do badacza w Kalifornii, jak również znaleźć się u jego kolegi z Rosji – w celu późniejszego zbadania. Procedury tego rodzaju są na porządku dziennym – zanim twoja kariera dobiegnie końca, uznasz je za prymitywne.

Rozwój, który udostępnił wszystkie te możliwości indywidualnym badaczom, stał się możliwy dzięki wynalazkowi komputerów osobistych (*Personal Computers* – PC), zwanych pierwotnie „mikrokomputerami" dla odróżnienia ich od dużych komputerów i ich następców, minikomputerów. Rewolucja PC nastąpiła tak nagle, że nietrudno zapomnieć, jak wiele się zmieniło w dość krótkim czasie. Kiedy kończyłem w latach 60. studia na Uniwersytecie Kalifornijskim w Berkeley, Centrum Badań Sondażowych

dysponowało prostym komputerem IBM 1620, który zajmował mniej więcej tyle miejsca, co sześć czy siedem lodówek. Pojemność pamięci tego komputera wynosiła 24Kb, albo 24000 bajtów (znaków) informacji. IBM 1620 obsługiwał od 30 do 40 czynnych badaczy, którzy rokrocznie publikowali wiele książek i artykułów.

W tej chwili pracuję na laptopie o pamięci 128 megabajtów: 5333 razy więcej pamięci niż stary 1620. Komputer kieszonkowy, którego używam do przechowywania adresów i telefonów, ma 167 razy więcej pamięci niż IBM 1620. Nawet mój zegarek ma jedną dwunastą pamięci ogromnej maszyny, która wystarczała na potrzeby 30–40 badaczy w Centrum Badań Sondażowych jeszcze nie tak dawno temu.

Większość komputerów osobistych ma dziś twarde dyski, które przechowują miliony bajtów (megabajty) lub nawet miliardy bajtów (gigabajty) informacji. Mój laptop na przykład ma twardy dysk o pojemności 9 gigabajtów. Jego zewnętrzny twardy dysk, nie większy niż paczka kart, ma 14 gigabajtów pojemności. Poza tą na stałe wbudowaną pojemnością wynalazki takie jak CD-ROM-y, dyski ZIP i inne nośniki danych znacznie zwiększyły możliwości PC w zakresie przechowywania danych. Płyty CD-ROM przechowują setki milionów bajtów na małych, cienkich dyskach. Cząstka „ROM" w skrócie CD-ROM oznacza: *read-only memory* (pamięć tylko do odczytu), co znaczyło początkowo, że możliwe było odczytanie danych zapisanych tam przez producenta, ale nie zapisanie na dysku czegokolwiek nowego. Obecnie ograniczenie to już wyeliminowano dzięki nowej generacji płyt kompaktowych, przeznaczonych także do zapisu.

Abstrakt Statystyczny Biura Spisowego na płycie kompaktowej to świetny przykład tego, jak wartościowe dla socjologów są tego typu nośniki. Przydatne są także liczne inne bibliografie i encyklopedie, jak na przykład *Encarta* firmy Microsoft. Historycy są zapewne najszczęśliwsi, dysponując tomami łatwo dostępnych materiałów na temat dowolnego okresu i zagadnienia w historii.

Wszystkie te technologiczne cudeńka opierają się jednak na tej samej podstawowej zasadzie, co krosna Jacquarda: informacja jest wprowadzana i przetwarzana, po czym otrzymujemy wynik. Dalsza część tego rozdziału zajmuje się etapami i możliwościami przekształcania danych w formę elektroniczną, podatną na komputerowe przetwarzanie i analizę.

Kodowanie

By czar komputera zaczął działać, musi on być w stanie odczytać dane, które zebrałeś w toku badań. Co więcej, komputerom szczególnie odpowiadają liczby. Jeśli badany mówi ci, że jego zdaniem największym obecnie problemem miasta Woodbury w hrabstwie Vermont jest „powiększająca się dziura ozonowa", komputer nie jest w stanie przetworzyć tego typu odpowiedzi metodami liczbowymi. Musisz ją przetłumaczyć, *kodując* odpowiedź. Mówiliśmy już o kodowaniu w związku z analizą treści (rozdział 11) oraz w związku z analizą danych jakościowych (rozdział 13). Teraz przyjrzymy się kodowaniu w kontekście analizy ilościowej.

Aby przeprowadzić analizę ilościową, badacz musi niekiedy po zebraniu danych przystąpić do procesu kodowania. Na przykład w kwestionariuszu z pytaniami otwartymi otrzymujemy odpowiedzi w formie nieliczbowej, które muszą zostać zakodowane przed przystąpieniem do analizy. Podobnie jak przy analizie treści, nasze zadanie polega na redukcji różnorodnych jednostkowych informacji do bardziej ograniczonego zbioru wartości składających się na daną zmienną. Załóżmy na przykład, że badacz zadaje respondentom pytanie: „Jaki jest Pana(i) zawód?". Odpowiedzi na to pytanie będą bardzo zróżnicowane. Mimo że możliwe byłoby przypisanie każdemu zadeklarowanemu zawodowi innego kodu liczbowego, nie uprościłoby to wcale analizy, która na ogół opiera się na fakcie, że wielu badanych ma te same cechy.

Jest wiele ustalonych schematów kodowania zmiennej *zawód*. Jeden z nich rozróżnia zawody kierownicze i specjalistów, zawody urzędnicze, zawody niewykwalifikowane itp. Inny z kolei rozróżnia rozmaite sektory gospodarki, jak przemysł, ochrona zdrowia, edukacja, handel i inne. Jeszcze inne schematy sta-

nowią kombinację tych dwóch. Aby dowiedzieć się więcej na ten temat, odwiedź stronę Ministerstwa Gospodarki, by uzyskać informacje na temat jego Klasyfikacji Zawodów (http://www.mgpips.gov.pl/pliki_do_pobrania/Klasyfikacja_zawodow_i_specjalnosci.rtf). Stosowanie ustalonego schematu kodowania pozwoli ci porównać wyniki twoich badań z pracami innych badaczy.

Schemat kodowania zawodu, który wybierzesz, powinien być dostosowany do pojęć teoretycznych, stosowanych w twoim badaniu. W niektórych przypadkach wystarczy kodowanie zawodów jako pracowników umysłowych i pracowników fizycznych. W innych badaniach użyteczne jest rozróżnienie „samozatrudniających się" i „zatrudnianych przez inną osobę". Badacz pacyfista mógłby być zainteresowany pytaniem, czy praca respondenta wiąże się z przemysłem zbrojeniowym.

Możliwe jest oczywiście przykrawanie schematu kodowania do potrzeb konkretnej analizy, należy jednak mieć na uwadze jedną ogólną wskazówkę: jeśli dane są kodowane w sposób bardzo szczegółowy, kodowane kategorie można zawsze połączyć na potrzeby analizy, która nie wymaga aż takiej szczegółowości (załącznik I opisuje, jak można to zrobić w SPSS). Jeśli dane są kodowane w stosunkowo niewielu ogólnych kategoriach, nie można w drodze analizy odtworzyć ich poprzedniej szczegółowości. Aby zostawić sobie możliwie szeroki wybór, dobrze jest kodować dane bardziej szczegółowo, niż wymaga tego planowana analiza.

Tworzenie kategorii kodowych

Do procesu kodowania można podejść zasadniczo na dwa sposoby. Po pierwsze, można wyjść od dość rozbudowanego schematu kodowania, zgodnego z twoim celem badawczym. W ten sposób wspomniany badacz pacyfista mógłby kodować zawody wedle ich związku z przemysłem zbrojeniowym. Można również użyć gotowego klucza kodowego, aby umożliwić sobie porównanie wyników własnych badań z badaniami poprzedników.

Inna metoda polega na generowaniu kodów na podstawie danych, podobnie jak w rozdziale 13. Załóżmy, że zadaliśmy studentom w uniwersyteckiej ankiecie do samodzielnego wypełnienia pytanie o największy obecnie problem na ich uczelni. Oto niektóre z możliwych odpowiedzi:

Za wysokie czesne
Za mało miejsc parkingowych
Kadra nie wie, co się dzieje
Kierownicy studiów są wiecznie nieosiągalni
Za mało oferowanych wykładów
Karaluchy w akademikach
Zbyt wysokie wymagania
Zepsute jedzenie w stołówce
Za drogie książki
Za mało stypendiów

Przejrzyj przez chwilę te odpowiedzi i sprawdź, czy reprezentują one jakieś szersze kategorie. Musisz zdać sobie sprawę, że nie ma właściwej odpowiedzi: na podstawie tych odpowiedzi można wygenerować wiele schematów kodowania.

Zacznijmy od pierwszej odpowiedzi: „za wysokie czesne". Jakie szersze pole problemowe wskazuje ta odpowiedź? Jedną z oczywistych możliwości są „problemy finansowe". Czy są inne odpowiedzi, które mieszczą się w tej kategorii? W tabeli 14.1 zaznaczono te odpowiedzi, które mieszczą się w kategorii „problemy finansowe".

W ogólniejszym ujęciu możemy także potraktować pierwszą kategorię jako odzwier-

Tabela 14.1

Odpowiedzi studentów, które mogą zostać zakodowane jako „problemy finansowe"

	Problemy finansowe
Za wysokie czesne	X
Za mało miejsc parkingowych	
Kadra nie wie, co się dzieje	
Kierownicy studiów są wiecznie nieosiągalni	
Za mało oferowanych wykładów	
Karaluchy w akademikach	
Zbyt wysokie wymagania	
Zepsute jedzenie w stołówce	
Za drogie książki	X
Za mało stypendiów	X

ciedlającą problemy pozaakademickie. Kategoryzacja taka byłaby trafna, gdyby twoje badania miały za przedmiot rozróżnienie problemów akademickich i pozaakademickich. Wówczas kody odpowiedzi wyglądałyby jak w tabeli 14.2.

Tabela 14.2

Problemy studentów kodowane jako „akademickie" i „pozaakademickie"

	Akademickie	Pozaakademickie
Za wysokie czesne		X
Za mało miejsc parkingowych		X
Kadra nie wie, co się dzieje	X	
Kierownicy studiów są wiecznie nieosiągalni	X	
Za mało oferowanych wykładów	X	
Karaluchy w akademikach		X
Zbyt wysokie wymagania	X	
Zepsute jedzenie w stołówce		X
Za drogie książki		
Za mało stypendiów		X

Zauważ, że nie zakodowałem w tabeli 14.2 odpowiedzi „za drogie książki", ponieważ odpowiedź tę można zakwalifikować do obu kategorii. Książki należą do programu akademickiego, jednak ich ceny już nie. Wskazuje to na potrzebę uściślenia schematu kodowania, który tworzymy. Zależnie od naszego celu moglibyśmy interesować się zwłaszcza określeniem problemów związanych z nauczaniem akademickim – wówczas ten przypadek zakodowalibyśmy jako „akademickie". Równie dobrze moglibyśmy jednak interesować się raczej określeniem problemów pozaakademickich i odpowiednio zakodować tę odpowiedź. W jeszcze innym przypadku moglibyśmy stworzyć odrębną kategorię dla odpowiedzi, zawierających zarówno element akademicki, jak i pozaakademicki.

Jest jednak jeszcze inna możliwość: możemy podzielić problemy pozaakademickie na te, które dotyczą kwestii administracyjnych, oraz te, które wiążą się ze stanem infrastruktury uczelni. Tabela 14.3 przedstawia sposób zakodowania dziesięciu pierwszych odpowiedzi w takim przypadku.

Jak widać na tych kilku przykładach, istnieje wiele schematów kodowania zbioru danych. Twój wybór powinien korespondować z twoim celem badawczym i stanowić odbicie logiki, wyłaniającej się z samych danych. Często będziesz zmuszony modyfikować kategorie kodowe w miarę postępów procesu kodowania. Jednak zawsze, gdy zmieniasz listę kategorii, musisz przejrzeć już zakodowane dane, by sprawdzić, czy zmiany zostały prawidłowo wprowadzone.

Podobnie jak zbiór wartości składających się na zmienną, czy też kategorie odpowiedzi w zamkniętym pytaniu kwestionariusza, kategorie kodów powinny być zarazem wyczerpujące i wzajemnie się wykluczające. Każda kodowana jednostka informacji powinna podpadać pod jedną i tylko jedną kategorię. Problem pojawia się, gdy dana odpowiedź zdaje się pasować do więcej niż jednej lub nie pasować do żadnej kategorii – oba te fakty sygnalizują rozbieżności między twoimi danymi a schematem kodowym.

Jeśli masz szczęście i dysponujesz pomocą w procesie kodowania, będziesz musiał wprowadzić twoich koderów w definicje kategorii kodowych i nauczyć ich poprawnie używać tych kategorii. W tym celu wyjaśnij znaczenie każdej kategorii i podaj kilka przykładów każdej z nich. Aby się upewnić, że koderzy dokładnie cię rozumieją, zakoduj z góry kilka przypadków. Następnie poproś koderów, by zakodowali te same przypadki nie wiedząc, że sam zrobiłeś to wcześniej. Na koniec porównaj ich

Tabela 14.3

Problemy pozaakademickie kodowane jako „administracja" lub „infrastruktura"

	Akademickie	Administracja	Infrastruktura
Za wysokie czesne		X	
Za mało miejsc parkingowych			X
Kadra nie wie, co się dzieje	X		
Kierownicy studiów są wiecznie nieosiągalni	X		
Za mało oferowanych wykładów	X		
Karaluchy w akademikach			X
Zbyt wysokie wymagania	X		
Zatrute jedzenie w stołówce			X
Za drogie książki	X		
Za mało stypendiów		X	

pracę z twoją własną. Jakiekolwiek niezgodności wskazywać będą na nieefektywność komunikacji między tobą a koderami. Jednak nawet jeśli ty i koderzy zgadzacie się doskonale, najlepiej sprawdzić kodowanie przynajmniej części przypadków w toku procesu kodowania.

Jeśli nie masz szczęścia i nikt nie pomaga ci w kodowaniu, powinieneś w każdym razie zweryfikować jakoś twoją wiarygodność jako kodera. Nikt nie jest doskonały, a już na pewno nie pełen entuzjazmu badacz na tropie wielkiego odkrycia. Załóżmy, że studiujesz nowy kult religijny i masz wrażenie, że osoby pozbawione normalnej rodziny będą najprawdopodobniej postrzegać nowy kult jako jej substytut. Niebezpieczeństwo polega na tym, że zawsze, gdy natkniesz się na badanego, który zadeklaruje brak rodziny, będziesz podświadomie starał się znaleźć w jego wypowiedziach dowód faktu, że kult religijny jest substytutem rodziny. Zatem, jeśli jest to w ogóle możliwe, postaraj się, by ktoś inny zakodował część twoich przypadków, by sprawdzić, czy osoba ta dokonała takich samych przyporządkowań jak ty.

Budowa książki kodowej

Produktem końcowym procesu kodowania jest przekształcenie danych w kody liczbowe. Kody te odpowiadają wartościom składającym się na zmienne, którym z kolei przypisane są pozycje w zbiorze danych. **Książka kodowa** to dokument, który opisuje pozycje zmiennych i wymienia kody przyporządkowane wartościom, składającym się na zmienne.

Książka kodowa spełnia dwie zasadnicze funkcje. Po pierwsze, jest podstawową pomocą w procesie kodowania. Po drugie, jest też pomocą w odszukaniu zmiennych i interpretacji kodów w zbiorze danych w toku analizy. Jeśli zdecydujesz się w trakcie analizy danych skorelować dwie zmienne, książka kodowa podpowie ci, gdzie znajdziesz te zmienne i co oznaczają kody.

Rysunek 14.1 przedstawia część książki kodowej, utworzoną z dwóch zmiennych z Generalnego Sondażu Społecznego. Przykład ten przedstawia kilka typowych elementów książki kodowej, choć nie istnieje oczywiście jeden prawidłowy format.

Zauważ przede wszystkim, że każda zmienna jest określona przez skrótową nazwę zmiennej, jak *POLVIEWS* czy *ATTEND*. Uczestnictwo respondentów w obrządkach religijnych możemy określić na przykład posługując się zmienną *ATTEND*. Przykład ten stanowi zastosowanie formatu, ustalonego przez Generalny Sondaż Społeczny, który został przeniesiony do SPSS. Inne zbiory danych i/lub programy do analizy mogą nadawać zmiennym inny format. Niektóre z nich stosują na przykład kody licz-

Rysunek 14.1

Część książki kodowej

POLVIEWS	*ATTEND*
Dużo się ostatnio słyszy o liberałach i konserwatystach. Przedstawię ci siedmiostopniową skalę, porządkującą poglądy ludzi od skrajnie liberalnych (punkt 1) do skrajnie konserwatywnych (punkt 7). Gdzie umieściłbyś siebie na tej skali?	Jak często uczestniczysz w mszach i nabożeństwach?
1. Zdecydowanie liberalny	0. Nigdy
2. Liberalny	1. Rzadziej niż raz do roku
3. Umiarkowanie liberalny	2. Mniej więcej raz lub dwa razy w roku
4. Zwolennik środka, centrum polityczne	3. Kilka razy w roku
5. Umiarkowanie konserwatywny	4. Mniej więcej raz w miesiącu
6. Konserwatywny	5. Dwa, trzy razy w miesiącu
7. Zdecydowanie konserwatywny	6. Prawie co tydzień
8. Nie wiem	7. Co tydzień
9. Brak odpowiedzi	8. Kilka razy w tygodniu lub częściej
	9. Nie wiem, brak odpowiedzi

bowe zamiast skrótowych nazw zmiennych. Musisz jednak dysponować jakimś narzędziem identyfikacji, które pozwoli ci odnaleźć i wykorzystać konkretne zmienne.

Ponadto każda książka kodowa powinna zawierać pełną definicję *każdej* zmiennej. W przypadku kwestionariusza definicja składa się z literalnie przytoczonego pytania, ponieważ, jak się już przekonaliśmy, sformułowanie pytania wywiera duży wpływ na odpowiedź. W przypadku *POLVIEWS* wiesz, że respondent otrzymał stronę z kilkoma kategoriami politycznymi z prośbą, by wybrał tę, do której najlepiej pasuje.

Książka kodowa wskazuje też wartości składające się na każdą ze zmiennych. W *POLVIEWS* na przykład respondent mógł opisać swoje poglądy polityczne jako „zdecydowanie liberalne", „liberalne", „umiarkowanie liberalne" itd.

Zwróć wreszcie uwagę, że każda ze zmiennych ma też oznaczenie liczbowe. W *POLVIEWS* „zdecydowanie liberalny" to kategoria kodowa 1. Te kody liczbowe wykorzystywane są w różnego rodzaju procesach przetwarzania danych. Mógłbyś na przykład postanowić połączyć w jedną kategorie od 1 do 3 (wszyscy respondenci „liberalni"). Łatwiej to przeprowadzić na liczbach niż na długich nazwach.

Jak dotąd, zajmowaliśmy się wykorzystaniem i logiką kodowania na potrzeby analizy ilościowej. Większość powyższych uwag znajduje zastosowanie niezależnie od tego, czy dalsza analiza będzie prowadzona z użyciem komputera, czy też nie. Jednak, podobnie jak w przypadku krosien Jacquarda, jeśli chcemy, żeby maszyna przetwarzała nasze dane, musimy wprowadzić dane w formie, z której może ona korzystać. W następnym podrozdziale wyjaśniam, jak się to robi.

Sposoby kodowania i wprowadzania danych

Wiele lat temu każdy proces wprowadzania danych polegał na ręcznym dziurkowaniu, po czym karty były analizowane przez maszyny sortujące (poprzedniczki komputerów) lub też – w przypadku bardziej złożonych analiz – przez komputery. Obecnie dane są na ogół wstukiwane bezpośrednio do pliku przechowywanego na dysku komputera. Proces ten jest jednak, podobnie jak dawniej, blisko związany z kodowaniem, i istnieje kilka metod efektywnego wykorzystania tego powiązania. Przyjrzyjmy się niektórym spośród nich.

Arkusze kodowe

Tradycyjna metoda przetwarzania danych wymaga zakodowania danych i przeniesienia przyporządkowanych im kodów na arkusz wyników czy inaczej **arkusz kodowy**. Tego typu arkusze były tradycyjnie podzielone na 80 kolumn, odpowiadających kolumnom danych na stronie z danymi, można je jednak adaptować do innych konfiguracji danych, zależnie od metody wprowadzania danych (ilustracja na rysunku 14.2). Przy tej metodzie wprowadzania danych koderzy wpisują liczby, odpowiadające

Rysunek 14.2

Część arkusza kodowego

01	02	03	04	05	06	07	08	09	10	...	71	72	73	74	75	76	77	78	79	80
0	0	1	3	7	8	9	3	1	1		4	5	2	1	1	7	8	7	1	2
0	0	2	4	2	4	2	4	1	2		2	2	5	1	1	2	8	2	2	2
0	0	3	5	6	1	2	3	1	1		1	3	6	2	1	3	9	2	2	3
0	0	4	6	3	4	4	2	2	1		1	4	1	0	3	6	0	4	2	1

pożądanym kategoriom kodowym, w odpowiednie kolumny na stronie. Każda kolumna w macierzy odpowiada zmiennej (na przykład *płeć*), każdy rząd odpowiada przypadkowi (a więc respondentowi w sondażu), liczba zaś wpisana w określonej komórce macierzy odpowiada wartości (na przykład 1 = mężczyzna, 2 = kobieta) danej zmiennej w poszczególnym przypadku. Arkusze kodowe są często stosowane przy wprowadzaniu danych do plików komputerowych. Technika ta jest wciąż użyteczna, zwłaszcza jeśli przedmiotem obróbki są szczególnie złożone kwestionariusze lub inne źródła danych.

Kodowanie brzegowe

Kodowanie brzegowe bywa wykorzystywane w celu wyeliminowania arkuszy kodowych. Przy **kodowaniu brzegowym** zewnętrzny margines każdej strony kwestionariusza lub innego źródła danych pozostawia się wolny lub oznacza się go polami odpowiadającymi nazwom zmiennych lub liczbom. Zamiast przenosić przypisane kody na osobny arkusz, kody wpisuje się w odpowiednich miejscach na marginesie. Dokumenty z kodowaniem brzegowym są potem wykorzystywane przy wprowadzaniu danych. Mógłbyś więc napisać na przykład „3" na marginesie obok pytania o „największy współcześnie problem naszego kraju", gdyby wpisana odpowiedź mieściła się w kategorii zakodowanej jako trzecia.

Bezpośrednie wprowadzanie danych

Dysponując odpowiednio zaprojektowanym kwestionariuszem lub inną formą zbioru danych, możesz niekiedy wprowadzać dane bezpośrednio do komputera, nie używając osobnych arkuszy odpowiedzi czy choćby kodowania brzegowego. Kwestionariusz prekodowany zawiera wskazanie numerów kolumn lub nazw zmiennych w macierzy danych, jak również kodów, które należy przypisać pytaniom i odpowiedziom, tak by dane mogły być wprowadzone bezpośrednio. (Przykład kwestionariusza wstępnie prekodowanego znajdziesz w załączniku D).

Wprowadzanie danych przez ankieterów

Najbardziej bezpośrednia metoda wprowadzania danych w badaniach sondażowych została już omówiona w rozdziale 9: sondaż telefoniczny z wykorzystaniem komputera, czyli CATI (*Computer Assisted Telephone Interviewing*). Jak zapewne pamiętasz, ankieterzy w słuchawkach siedzą przed terminalami komputerowymi. Terminal wyświetla pytania, które należy zadać, ankieter zaś wpisuje odpowiedzi respondentów. W ten sposób dane są wprowadzane bezpośrednio do komputerowego pliku danych już w momencie ich pozyskania. Dane z pytań zamkniętych są natychmiast gotowe do analizy. Dane z pytań otwartych również mogą być w ten sposób wprowadzane, wymagają one jednak dodania jeszcze jednego etapu.

Wyobraź sobie, że w kwestionariuszu pada pytanie: „Jaki jest twoim zdaniem największy współcześnie problem Stanów Zjednoczonych?". Terminal komputerowy podpowiedziałby to pytanie ankieterowi. Następnie, zamiast oczekiwać prostego kodu liczbowego jako odpowiedzi, pozwoliłby ankieterowi wpisać wszystko, cokolwiek powie respondent, na przykład: „przestępczość na ulicach, zwłaszcza przestępstwa popełniane przez handlarzy narkotyków". Następnie koderzy musieliby, siedząc przed terminalami komputerowymi, wyszukać po kolei *pytania otwarte i przypisać* odpowiedziom kody liczbowe wedle metody omówionej powyżej w tym rozdziale.

Kodowanie na karty do skanera optycznego

Dane mogą być niekiedy wprowadzane przy użyciu skanera optycznego. Urządzenie to odczytuje napisane ołówkiem czarne znaki na specjalnych arkuszach kodowych i tworzy pliki danych odpowiadające tym znakom. Koderzy przenoszą dane na specjalne arkusze i zaczerniają odpowiednie pola. Karty te wsuwa się następnie do skanera optycznego, po czym pliki danych tworzone są automatycznie.

Skaner optyczny, mimo że pozwala na pracę szybszą i dokładniejszą niż przy ręcznym

wprowadzaniu danych, ma także wady. Niektórzy koderzy mają trudności z przenoszeniem danych na specjalne arkusze. Bywa, że niełatwo jest znaleźć odpowiednią kolumnę, kiedy zaś koder już ją znajdzie, musi odszukać właściwe pole do zaczernienia.

Poza tym skaner optyczny ma dość sztywne rygory tolerancji. Jeśli czarne znaki nie są dostatecznie czarne, skaner może robić błędy. Co więcej, badacz nie jest w stanie sprawdzić, czy coś takiego miało miejsce, póki nie rozpocznie analizy. Jeśli zaś karty kodowe są zgięte lub uszkodzone, skaner może w ogóle ich nie odczytać.

Bezpośrednie użycie kart do skanera optycznego

Karty do skanera optycznego mogą być używane w nieco inny sposób, tak by uniknąć związanych z tym trudności. Być może znasz już takie formy z testów wielokrotnego wyboru – niekiedy prosi się respondentów, by nanieśli swoje odpowiedzi bezpośrednio na takie karty. Można w tym celu wykorzystać standardowe lub specjalnie zaprojektowane karty wraz z instrukcją ich użycia. Można poprosić respondentów, by, odpowiadając na pytania z kilkoma kategoriami odpowiedzi, zaczernili pola obok wybranej odpowiedzi. Przy odpowiednim wypełnieniu tego typu karty skaner optyczny może bezpośrednio odczytać i wprowadzić odpowiedzi. Metoda ta bywa szczególnie użyteczna przy wprowadzaniu obserwacji eksperymentalnych lub wczytywaniu danych w analizie treści.

Powiązanie z programem do analizy danych

Rozmaite programy komputerowe nadają zbiorom danych różną strukturę. W większości przypadków będziesz zapewne używał przy wprowadzaniu danych twojego programu do analizy danych. SPSS na przykład zaproponuje ci pustą macierz złożoną z rzędów i kolumn. Kolumnom przypiszesz nazwy zmiennych, po czym wprowadzisz dane dla każdego przypadku (na przykład odpowiedzi każdego z respondentów w badaniach sondażowych) w osobnym wierszu. Kiedy skończysz, dane będą gotowe do analizy.

Alternatywną metodą jest stworzenie własnego zbioru danych przy użyciu innych środków (takich jak arkusz kalkulacyjny lub edytor tekstu) i przerzucenie danych do programu do analizy danych. W przypadku SPSS plik tekstowy z pojedynczymi danymi oddzielonymi za pomocą tabulatora (na przykład datafile.dat) może zostać przeniesiony i zapisany w formacie SPSS (jako na przykład datafile.sav). Możesz następnie wczytać plik danych tak, jakby był on stworzony od początku przez SPSS. Większość innych programów do analizy danych ma podobną opcję.

Niezależnie od przyjętej metody obróbki danych, będziesz po wykonaniu tych działań dysponował zbiorem danych czytelnych dla komputera, które najprawdopodobniej odpowiadają informacjom uzyskanym w toku twojej pracy. Kolejny ważny etap to usunięcie błędów: „czyszczenie" danych.

Czyszczenie danych

Niezależnie od metody wprowadzania danych – niezależnie też od staranności wprowadzającego – nie da się uniknąć pewnych błędów. Zależnie od metody obróbki danych błędy te mogą być wynikiem niepoprawnego kodowania, nieprawidłowego odczytania zaczernionych kodów, błędnego odczytania zaczernionych pól na kartach do skanera optycznego itp. **Czyszczenie danych** to proces wykrywania i usuwania błędów w kodowaniu. Powinno się przeprowadzić dwa typy czyszczenia: czyszczenie dotyczące dopuszczalnych kodów i czyszczenie logiczne.

Czyszczenie dotyczące dopuszczalnych kodów

Dla każdej zmiennej istnieje określony zbiór możliwych wartości, przekładający się na zbiór dopuszczalnych kodów. W przypadku zmiennej *płeć* możliwe będą najprawdopodobniej trzy kody: 1 dla mężczyzny, 2 dla kobiety i 0 dla braku odpowiedzi. Jeżeli odpowiedź na pytanie o płeć została zakodowana jako 7, oczywiste jest, że popełniono błąd.

Czyszczenie dotyczące dopuszczalnych kodów to proces sprawdzania, czy w zbiorze danych pojawiają się tylko kody przypisane do poszczególnych wartości zmiennych – czyli kody dopuszczalne. Można je przeprowadzić na dwa sposoby. Po pierwsze, wiele dostępnych programów, służących do wprowadzania danych, sprawdza dane w miarę ich wprowadzania. Jeśli próbujesz w takim programie wprowadzić na przykład 7 dla płci, komputer wydaje dźwięk w rodzaju „bip" i odrzuca błędny kod. Istnieją też programy do sprawdzania prawomocności kodów w zbiorach danych, które nie zostały skontrolowane podczas wprowadzania danych.

Jeśli nie masz dostępu do tego typu programów, możesz zlokalizować niektóre błędy, sprawdzając rozkład odpowiedzi dla każdej zmiennej w twoim zbiorze. Jeśli więc stwierdzisz, że twój zbiór danych zawiera 350 osób, zakodowanych jako 1 dla zmiennej płeć (mężczyźni), 400 osób zakodowanych jako 2 (kobiety) i jedną osobę z kodem 7, uznasz zapewne, że 7 oznacza błąd. Kiedy odkryjesz błąd, następny krok polega na odszukaniu odpowiedniego dokumentu źródłowego (na przykład kwestionariusza), stwierdzeniu, jaki kod powinien był się tam pojawić, i wprowadzeniu koniecznych poprawek.

Czyszczenie logiczne

Drugi typ czyszczenia zbioru danych jest bardziej złożony. Struktura logiczna danych może nakładać pewne ograniczenia na odpowiedzi niektórych respondentów. **Czyszczenie logiczne** ma na celu sprawdzenie, czy dane dla poszczególnych zmiennych występują tylko w przypadkach, w których powinny wystąpić. W kwestionariuszu może się na przykład znaleźć pytanie o liczbę dzieci, które urodziła dana kobieta. Wszystkie respondentki płci żeńskiej powinny więc mieć zakodowaną odpowiedź na to pytanie (lub specjalny kod na brak odpowiedzi), jednak żaden respondent mężczyzna nie powinien udzielić na to pytanie odpowiedzi (lub też powinien pojawić się specjalny kod oznaczający, że pytanie go nie dotyczy). Jeśli zakodowano, że dany respondent płci męskiej urodził troje dzieci, musisz uznać, że albo popełniono błąd i należy go skorygować, albo też twoim badaniom pisana jest sława większa, niż mógłbyś się kiedykolwiek spodziewać.

Mimo że czyszczenie zbioru danych jest koniecznym etapem w obróbce danych, możesz go spokojnie uniknąć w niektórych przypadkach. Być może uznasz, że możesz wykluczyć nieliczne błędy, które pojawiają się w danym pytaniu – jeśli wykluczenie tych przypadków nie wpłynie znacząco na twoje wyniki. Niektóre niewłaściwe z punktu widzenia zgodności odpowiedzi można spokojnie pominąć. Jeśli nawet pewnym mężczyznom przypisano pewną wartość dzietności w procesie kodowania, możesz ograniczyć swoją analizę tej zmiennej do kobiet. Nie powinieneś jednak traktować tych uwag jako usprawiedliwiania niechlujnych badań. „Zanieczyszczenia" zbioru danych niemal zawsze prowadzą do mylnych wniosków badawczych.

Główne zagadnienia

- Kwantyfikacja danych jest konieczna dla celów analizy statystycznej.
- Komputery pozwalają przechowywać, odczytywać i przetwarzać znaczne ilości danych. Wiele programów zaprojektowano specjalnie na użytek badań społecznych.
- Aby zastosować komputerową analizę danych ilościowych, badacz musi przekształcić obserwacje, opisujące każdą jednostkę analizy, w standaryzowane, liczbowe kody, które mogą być odczytywane i przetwarzane przez komputer.
- W procesie kodowania danej zmiennej przypisujemy konkretny identyfikator w zbiorze danych: może to być liczba lub skrót nazwy. Wartościom danej zmiennej odpowiadają kody liczbowe.
- Badacze mogą korzystać z istniejących schematów kodowych, w rodzaju klasyfikacji zawodów Biura Spisowego, lub też tworzyć własne kategorie kodowe. W każdym przypadku klucz kodowy musi odpowiadać naturze i celowi badania.

- Książka kodowa to dokument opisujący identyfikatory przypisane poszczególnym zmiennym wraz z kodami odpowiadającymi wartościom tych zmiennych.
- Wprowadzanie danych można przeprowadzić na wiele sposobów. Badacz może zapisać kody liczbowe na arkuszach kodowych, używanych następnie przy wprowadzaniu danych, lub też przeprowadzić tę samą operację z wykorzystaniem kodowania brzegowego. Inne możliwości obejmują bezpośrednie wprowadzanie danych do plików komputerowych, wprowadzanie danych przez ankieterów używających komputerów czy też wykorzystanie kart do skanera optycznego.
- Dane mogą być wprowadzane bezpośrednio w programie używanym do analizy lub też przeniesione do tego programu z pliku tekstowego, arkusza kalkulacyjnego lub bazy danych.
- Jako że w każdym procesie kodowania zdarzają się błędy, na ogół konieczne jest „czyszczenie" danych. Dwa typy czyszczenia zbioru danych – dotyczące dopuszczalnych kodów i logiczne – zabezpieczają przed jednym z możliwych błędów w obróbce danych.

Kluczowe pojęcia

analiza ilościowa
książka kodowa
arkusz kodowy
kodowanie brzegowe
czyszczenie danych
czyszczenie dotyczące kodów dopuszczalnych
czyszczenie logiczne

Pytania i ćwiczenia

1. Dowiedz się, jakim wyposażeniem w zakresie przetwarzania danych dysponuje twoja uczelnia, i opisz, jak można za jego pomocą opracowywać kwestionariusze z badań sondażowych.
2. Posługując się informatorem uczelnianym lub stroną internetową uczelni określ, jaką strukturę kategorii wykorzystuje organizacja jednostek twojej uczelni.
3. Wybierz losowo dziesięć książek ze swojej biblioteczki. Stwórz trzy różne zbiory kategorii, za pomocą których można sklasyfikować te książki.
4. Stwórz książkę kodową – ze zmiennymi i przyporządkowaniami kodów – dla następujących pytań:
   a. Czy brałeś udział w wyborach w 1996 roku?
      ☐ Tak – Jeśli tak, na kogo głosowałeś?
      ☐ Smith ☐ Brown
      ☐ Nie
   b. Czy uważasz się zasadniczo za demokratę, republikanina, za niezależnego czy za jeszcze kogoś innego?
      ☐ Republikanin
      ☐ Demokrata
      ☐ Niezależny
      ☐ Inny: _____
   c. Jaki jest twoim zdaniem najpoważniejszy obecnie problem twojej gminy?
   d. Wskaż w wyznaczonych miejscach poniżej trzy problemy twojej gminy, które szczególnie cię dotyczą, wpisując 1 dla dotyczącego cię najbardziej, 2 dla drugiego w kolejności i 3 dla ostatniego.
      _____ Przestępczość
      _____ Ruch uliczny
      _____ Nadużywanie narkotyków
      _____ Zanieczyszczenie środowiska
      _____ Uprzedzenia i dyskryminacja
      _____ Recesja
      _____ Bezrobocie
      _____ Brak mieszkań
5. Przez wiele lat Generalny Sondaż Społeczny badał opinię publiczną w sprawie „Poprawki o równouprawnieniu", donosząc o odsetkach jej zwolenników i przeciwników. Twoim zadaniem jest podać niektóre uzasadnienia postaw wobec poprawki, podawane przez respondentów. Zacznij od odwiedzenia strony http://www.icpsr.umich.edu:8080/GSS/homepage.htm. Przejrzyj aneksy, znajdź fragment dotyczący kodowania opinii na temat „Poprawki" i wypisz trzy uzasadnienia podawane przez jej zwolenników oraz trzy podawane przez przeciwników.

Dodatkowe lektury

Bernstein Ira H., Havig Paul, *Computer Literacy. Getting the most from your PC*, Sage, Thousands Oaks, Cal. 1999. Krótki przegląd rozmaitych metod użycia komputerów w naukach społecznych wraz z kilkoma popularnymi programami użytkowymi.

Ferrante Joan, Vaughn Angela, *Let's Go Sociology: Travels on the Internet*, Wadsworth, Belmont, Cal. 1999. Ta przystępna książeczka stanowi świetne wprowadzenie do Internetu i zawiera informacje o wielu stronach internetowych, interesujących z punktu widzenia badacza społecznego.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

15

Podstawy analizy ilościowej

Przegląd holograficzny

Analiza to proces stawiania pytań twoim danym – i otrzymywania od nich odpowiedzi. Analiza ilościowa może być opisowa lub wyjaśniająca, może obejmować jedną, dwie lub więcej zmiennych. Rozpoczniemy nasze rozważania nad prowadzeniem analizy ilościowej od omówienia kilku prostych, lecz niezwykle przydatnych metod przetwarzania danych w celu wyciągnięcia wniosków badawczych.

Wprowadzenie

Analiza jednozmiennowa
Rozkłady
Tendencja centralna
Rozproszenie
Zmienne ciągłe i dyskretne
Szczegółowość a czytelność

Porównywanie podgrup
Łączenie kategorii odpowiedzi
Sposoby postępowania z kategorią „nie wiem"
Opisy liczbowe w badaniach jakościowych

Analiza dwuzmiennowa
Procentowanie w tabeli
Tworzenie i odczytywanie tabel dwuzmiennowych

Wprowadzenie do analizy wielozmiennowej

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

W poprzednim rozdziale zajmowaliśmy się przygotowaniem danych do analizy ilościowej. Jesteśmy w tej chwili gotowi, by przystąpić do tego etapu badań, do którego prowadził cały proces zbierania i opracowywania danych – do analizy. To w trakcie analizy danych odkrywamy i wydobywamy prawidłowości i zależności, sprawdzamy nasze przewidywania, wyciągamy wnioski, które sprawiają, że nasze badania stają się owocne.

Znaczna większość analiz socjologicznych mieści się w rubryce z napisem: **analiza wielozmiennowa** lub: jednoczesne badanie wielu zmiennych. Analiza jednoczesnych powiązań między *wiekiem*, *wykształceniem* i *uprzedzeniami* to przykład analizy wielozmiennowej.

Analiza wielozmiennowa to termin ogólny na określenie analizy kilku zmiennych, nie zaś jakiejś szczególnej formy analizy. Do konkretnych technik używanych w toku analizy wielozmiennowej należą m.in. analiza czynnikowa, analiza metodą najmniejszych przestrzeni, korelacja wielokrotna, regresja wielokrotna i analiza ścieżkowa. Kolejne dwa rozdziały omawiają dokładniej te techniki. Poniższy rozdział zapewnia podstawę dla zrozumienia dalej rozważanych tematów w postaci omówienia bardziej podstawowego modelu analitycznego, to jest analizy jedno- i dwuzmiennowej. Przy tej okazji poznamy też podstawy logiki analizy wielozmiennowej poprzez wykorzystanie prostych tabel zwanych *tabelami krzyżowymi* lub *tabelami kontyngencji*.

Analiza jednozmiennowa

Najprostsza forma analizy ilościowej, **analiza jednozmiennowa** polega na opisywaniu przypadku w kategoriach pojedynczej zmiennej, ze szczególnym uwzględnieniem rozkładu jej wartości. Mierząc na przykład zmienną *płeć*, sprawdzalibyśmy, ilu badanych to kobiety, a ilu – mężczyźni.

Rozkłady

Podstawowy format prezentacji danych jednozmiennowych wymaga uwzględnienia wszystkich przypadków indywidualnych, to znaczy wyszczególnienia wartości badanej zmiennej dla każdego przypadku. Weźmy na przykład dane Generalnego Sondażu Społecznego (GSS) dotyczące uczestnictwa w praktykach religijnych, *ATTEND*. Tabela 15.1 prezentuje wyniki analizy tej zmiennej przez SPSS.

Zbadajmy tę tabelę krok po kroku. Po pierwsze, jeśli spojrzysz na sam dół tabeli, zobaczysz, że analizowana próba składała się ogółem z 1500 przypadków. W ostatnim rzędzie, ponad danymi ogółem, widać, że 17 z 1500 respondentów odpowiedziało, że nie wiedzą (NW) lub nie dało żadnej odpowiedzi (BO) na to pytanie. Tak więc nasze wnioski co do uczestnictwa Amerykanów w praktykach religijnych w latach 1973–1993 będą oparte na odpowiedziach udzielonych przez 1483 respondentów.

Wróćmy teraz do górnej części tabeli. Widać, że 224 osoby podały, że nigdy nie uczestniczą w praktykach religijnych. Liczba ta sama w sobie nic nam nie mówi o praktykach religijnych. Gdyby nasze dane pochodziły od 3000 respondentów zamiast od 1500, moglibyśmy przypuszczać, że 448 osób odpowiedziałoby, że nigdy nie uczestniczą w obrzędach religijnych. Ani liczba 224, ani 448 same w sobie nie mówią nam nic o tym, czy „przeciętny" Amerykanin chodzi do kościoła rzadko czy często.

Załóżmy dla analogii, że twoja najlepsza przyjaciółka mówi ci, że wypiła sześciopak piwa. To dużo czy mało? Odpowiedź zależy naturalnie od tego, czy wypiła to piwo w ciągu miesiąca, tygodnia, dnia czy godziny. W przypadku uczestnictwa w obrzędach religijnych potrzebujemy również jakiejś podstawy dla zrozumienia liczby ludzi, którzy nigdy nie chodzą do kościoła.

Jeden ze sposobów oceny tej liczby wymaga obliczenia, jaki odsetek wszystkich respondentów stanowią ci, którzy powiedzieli, że nigdy nie chodzą do kościoła. Gdybyś podzielił 224 przez 1483 tych, którzy dali jakąkolwiek odpowiedź, otrzymałbyś 15,1%, co pojawia się

Tabela 15.1

Uczestniczenie w mszach i nabożeństwach według GSS

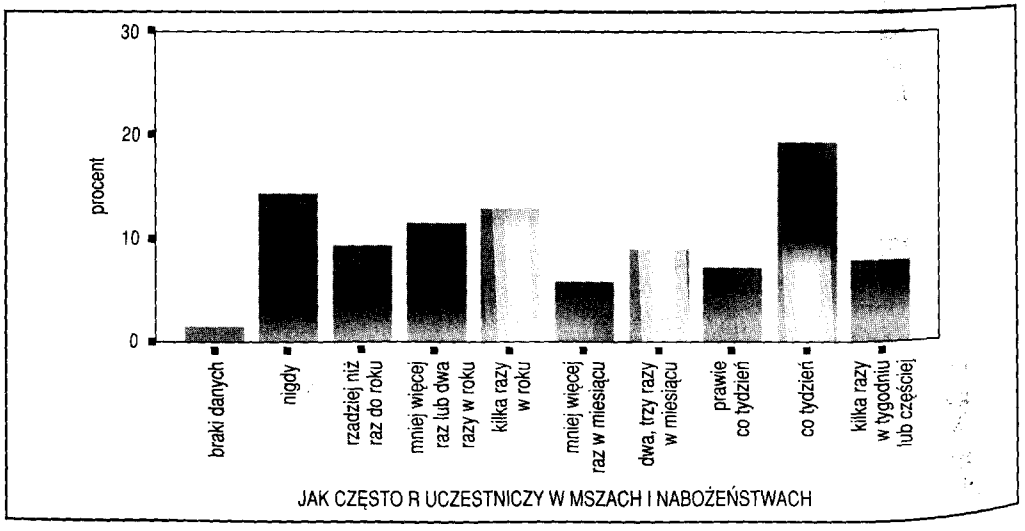
UCZESTNICTWO		Jak często R uczestniczy w mszach i nabożeństwach			
Nazwa wartości	wartość	częstość	procent	procent ważnych	procent skumulowany
Nigdy	0	224	14,9	15,1	15,1
Rzadziej niż raz do roku	1	139	9,3	9,4	24,5
Mniej więcej raz lub dwa razy w roku	2	180	12,0	12,1	36,6
Kilka razy w roku	3	194	12,9	13,1	49,7
Mniej więcej raz w miesiącu	4	84	5,6	5,7	55,4
Dwa, trzy razy w miesiącu	5	136	9,1	9,2	64,5
Prawie co tydzień	6	114	7,6	7,7	72,2
Co tydzień	7	294	19,6	19,8	92,0
Kilka razy w tygodniu lub częściej	8	118	7,9	8,0	100,0
NW, BO	9	17	1,1	brak	
	Ogółem	1500	100,0	100,0	
Ważnych przypadków 1483	Braki odpowiedzi 17				

w tabeli jako „procent ważnych". Możemy teraz stwierdzić, że 15%, czyli z grubsza biorąc co siódmy dorosły Amerykanin podał, że nigdy nie uczestniczy w obrzędach religijnych.

Taki wynik jest bardziej zrozumiały, czy znaczy on jednak, że ludzie w Stanach Zjednoczonych są na ogół mało religijni? Dalszy przegląd tabeli 15.1 pozwala stwierdzić, że najczęściej wybieraną odpowiedzią było „co tydzień". Taką odpowiedź wybrało 19,8% respondentów. Po dodaniu do tego 8% tych, którzy stwierdzili, że chodzą do kościoła częściej niż raz na tydzień, dojdziemy do wniosku, że ponad jedna czwarta (27,8%) dorosłych Amerykanów twierdzi, że uczestniczy w obrzędach religijnych co najmniej raz na tydzień. Jak widzisz, każde nowe porównanie stwarza bardziej kompletny obraz danych.

Rysunek 15.1

Wykres słupkowy zmiennej ATTEND w GSS, 1973–1993



Opis tego, ile razy dana odpowiedź pojawia się w próbie, nazywamy **rozkładem częstości**. Niekiedy najłatwiej przedstawić rozkład częstości na wykresie. Rysunek 15.1 został stworzony przez SPSS na podstawie danych GSS na temat *ATTEND*. Pionowa skala po lewej stronie wykresu wskazuje procent respondentów, wybierających każdą z odpowiedzi wymienionych wzdłuż poziomej osi wykresu. Przyjrzyj się przez chwilę, w jaki sposób procenty w tabeli 15.1 odpowiadają wysokości słupków na rysunku 15.1.

Tendencja centralna

Poza prostym opisem ogólnego rozkładu wartości, zwanym niekiedy *częstościami brzegowymi* lub po prostu *marginesami*, zechcesz być może przedstawić swoje dane w formie **wartości przeciętnej** lub *tendencji centralnej*. Pojęcie tendencji centralnej jest ci już znane z rozmaitych przeciętnych, których używasz w życiu codziennym, by wyrazić „typową" wartość zmiennej. W baseballu na przykład przeciętna celność pałkarza na poziomie 0,300 oznacza, że pałkarz trafia w piłkę przeciętnie trzy razy na dziesięć. W trakcie sezonu pałkarz mógłby mieć dłuższe okresy bez żadnych trafień oraz takie okresy, kiedy trafiałby za każdym razem. Z upływem czasu jednak tendencja centralna celności pałkarza może być opisana jako trzy trafienia w każdych dziesięciu próbach. Podobnie, przeciętna twoich ocen wyraża „typową" wartość wszystkich twoich ocen razem wziętych, mimo że niektóre z nich to piątki, inne – czwórki, a może masz nawet jedną lub dwie trójki (wiem dobrze, że nigdy nie dostajesz mniej niż trójkę).

Przeciętne tego rodzaju nazywa się **średnimi** arytmetycznymi (wynik dzielenia sumy wartości przez ogólną liczbę przypadków). Średnia to tylko jeden ze sposobów pomiaru tendencji centralnej lub wartości „typowych". Dwie inne możliwości to **modalna** (najczęściej występująca wartość) oraz **mediana** (środkowa wartość w uporządkowanym rozkładzie wszystkich zaobserwowanych wartości). Oto w jaki sposób wylicza się te trzy wartości przeciętne na podstawie zbioru danych.

Załóżmy, że prowadzisz eksperyment, w którym badanymi są nastolatki. Ich wiek wynosi od 13 do 19 lat, zgodnie z poniższą tabelą:

Wiek	Liczba
13	3
14	4
15	6
16	8
17	4
18	3
19	3

Teraz, kiedy znasz już faktyczny wiek każdego z 31 badanych, jak określiłbyś ogólnie ich „przeciętny" wiek? Spójrzmy na trzy możliwe odpowiedzi na to pytanie.

Przeciętna, którą najłatwiej obliczyć, to modalna, czyli najczęstsza wartość. Jak widzisz, szesnastolatków było więcej niż osób w jakimkolwiek innym wieku, więc modalna wieku wynosi 16, co ilustruje rysunek 15.2. Technicznie rzecz biorąc, modalna wieku obejmuje kategorię „16", do której mogą należeć także osoby, które są bliżej 17 niż 16 roku życia, ale nie miały jeszcze 17 urodzin.

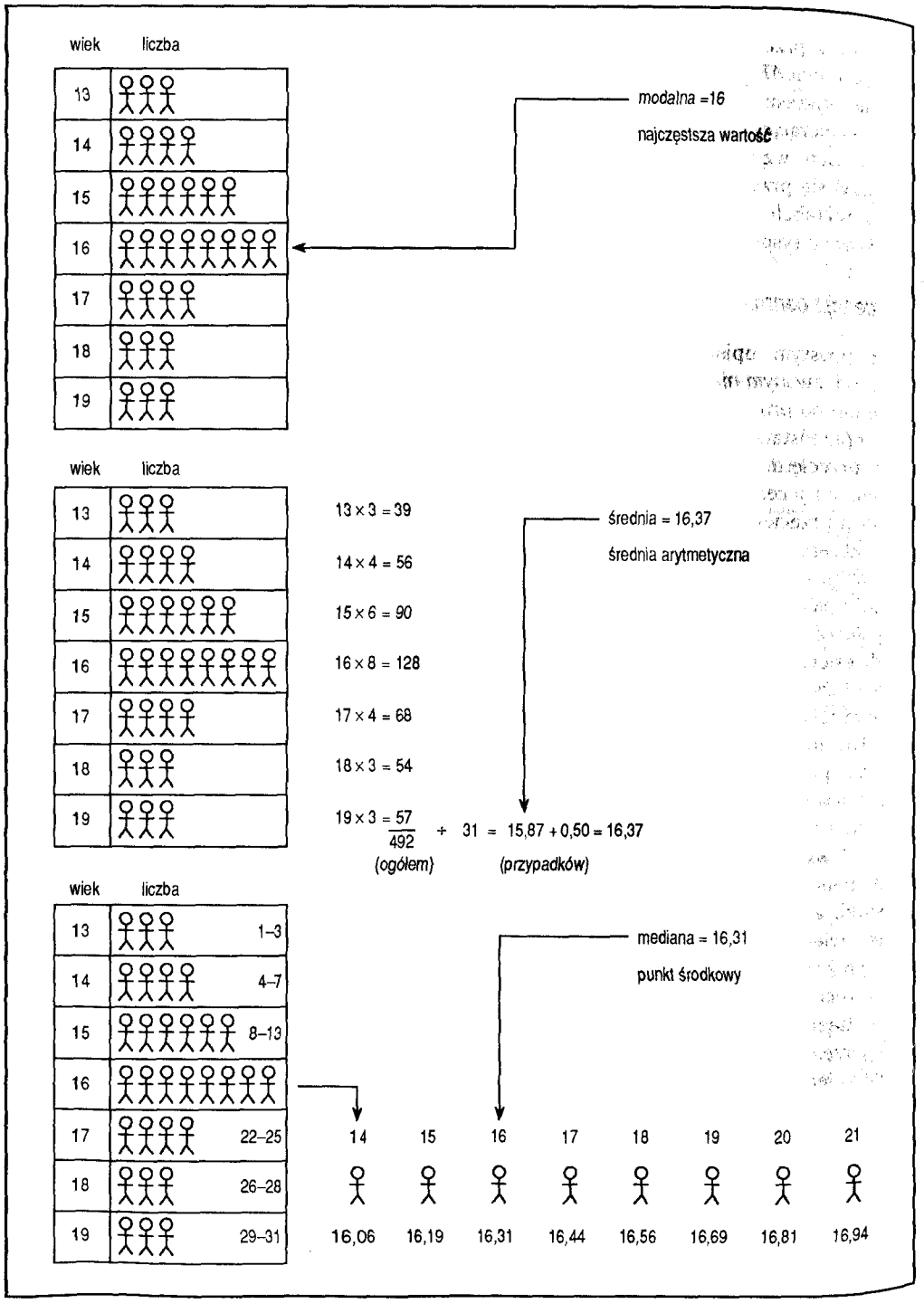
Rysunek 15.2 przedstawia także wyliczenie średniej. Obejmuje ono trzy kroki: (1) pomnóż każdy wiek przez liczbę badanych w tym wieku, (2) dodaj wyniki wszystkich tych mnożeń i (3) podziel sumę przez liczbę badanych.

W przypadku wieku potrzebne jest dodatkowe założenie. Zwróciłem już uwagę przy omówieniu modalnej, że osoby nazywające siebie 13-latkami to zarówno ci, którzy mają dokładnie 13 lat, jak i ci, którzy są tuż przed 14 urodzinami. Co więcej, można zakładać, że w całym kraju grupa 13-latków rozkłada się równomiernie na przestrzeni jednego roku, co daje przeciętny wiek 13,5 roku. Prawidłowość ta zachodzi dla każdej z grup wiekowych. Z tego też powodu wydaje się, że należy dodać do ostatecznego wyniku 0,5 roku, co daje średni wiek 16,37, jak to zilustrowano na rysunku 15.2.

Trzeci miernik tendencji centralnej, mediana, reprezentuje wartość „środkową": połowa jest powyżej, połowa poniżej. Gdybyśmy znali dokładny wiek każdego badanego (na przykład 17 lat i 124 dni), moglibyśmy uporządkować wszystkich 31 badanych według wieku, a me-

Rysunek 15.2

Trzy „przeciętne"



dianą dla całej grupy byłby wiek środkowego badanego.

Jak jednak widzisz, nie znamy dokładnego wieku – nasze dane są pod tym względem „danymi grupowymi". Na przykład trzy osoby, które nie są dokładnie w tym samym wieku, zostały umieszczone w kategorii „13-latki".

Rysunek 15.2 przedstawia zasadę wyliczenia mediany dla danych grupowych. Jako że jest 31 badanych, środkowy byłby badany numer 16, gdyby uszeregowano ich według wieku – 15 nastolatków byłoby starszych, a 15 młodszych. Spójrz na dolny fragment rysunku 15.2, a zobaczysz, że środkowa osoba to jedno z ośmiorga 16-latków. W powiększeniu widać, że numer 16 to trzecia osoba od lewej.

Jako że nie znamy dokładnego wieku badanych w tej grupie, statystyczna konwencja każe nam zakładać, że wiek rozkłada się równomiernie w całej tej grupie. W tym przypadku możliwy wiek badanego wynosi od 16 lat i 0 dni do 16 lat i 364 dni. Ściśle rzecz biorąc skala możliwości wynosi 364/365 dni. Dla celów praktycznych spokojnie możemy to nazwać jednym rokiem.

Gdyby ośmiu badanych w tej grupie rozłożyło się równomiernie od jednej do drugiej granicy, byliby od siebie oddaleni o jedną ósmą roku – interwał wynosiłby 0,125 roku. Spójrz na rysunek, a przekonasz się, że jeśli umieścimy pierwszego badanego o pół interwału od dolnej granicy, po czym dodamy pełną wartość interwału, by obliczyć wiek każdego kolejnego badanego, ostatni wypada pół interwału od górnej granicy.

W ten sposób obliczyliśmy hipotetycznie dokładny wiek każdego z ośmiu badanych – zakładając równy rozkład ich wieku. Następnie możemy zauważyć, że wiek środkowego badanego wynosi 16,31 i że jest to mediana wieku całej grupy.

Kiedy całkowita liczba badanych jest parzysta, nie ma oczywiście przypadku środkowego. By otrzymać medianę, liczy się średnią z dwóch wartości po obu stronach punktu środkowego w uszeregowanym zbiorze danych. Załóżmy na przykład, że w naszej grupie znalazłby się jeszcze jeden 19-latek, co dałoby liczbę ogólną 32 przypadków. Punkt środkowy wypadłby wtedy między badanymi numer 16 i 17. Medianę liczylibyśmy wtedy następująco: (16,31 + 16,44)/2 = 16,38.

Jak widać na rysunku 15.2, trzy mierniki tendencji centralnej dają dla naszego zbioru danych trzy różne wartości, co zdarza się często (acz nie zawsze). Który zatem z tych mierników najlepiej oddaje wartość „przeciętną"? Bardziej ogólnie: który miernik tendencji centralnej powinniśmy preferować? Odpowiedź zależy od typu danych i od celu analizy. Na przykład prezentując średnią, powinieneś być świadom, że jest ona podatna na oddziaływanie wartości skrajnych – niewielu bardzo małych lub bardzo dużych liczb. By podać tylko jeden przykład: średnia wartość majątku mieszkańca Redmond w stanie Waszyngton wynosi ponad 2 miliony dolarów. Gdybyś jednak odwiedził Redmond, stwierdziłbyś zapewne, że „przeciętny" mieszkaniec tej miejscowości odstaje nieco od twoich wyobrażeń o milionerach. Wysokość tej średniej stanowi odbicie wpływu jednego przypadku ekstremalnego wśród mieszkańców Redmond, a mianowicie Billa Gatesa z Microsoftu, którego majątek (w chwili gdy to piszę) jest wart około 90 miliardów dolarów. Z całą pewnością mediana wartości majątku zapewniłaby ci trafniejszy pogląd na temat ogółu mieszkańców Redmond.

Powyższy przykład powinien stanowić dobrą ilustrację konieczności starannego wyboru pomiędzy różnymi miarami tendencji centralnej. W podręczniku do statystyki znajdziesz pełniejsze informacje na temat rozmaitych sytuacji, w których można stosować każdą z nich.

Rozproszenie

Wartości przeciętne pozwalają czytelnikowi zredukować surowe dane do bardziej poręcznej formy: pojedyncza liczba (lub cecha) może reprezentować wszystkie szczegółowe dane zebrane na temat konkretnej zmiennej. Ułatwienie to uzyskujemy jednak oczywiście pewnym kosztem, ponieważ czytelnik nie jest w stanie odtworzyć oryginalnych danych na podstawie wartości przeciętnej. Sumaryczne ujęcie rozproszenia odpowiedzi może jednak nieco zmniejszyć wpływ tej wady.

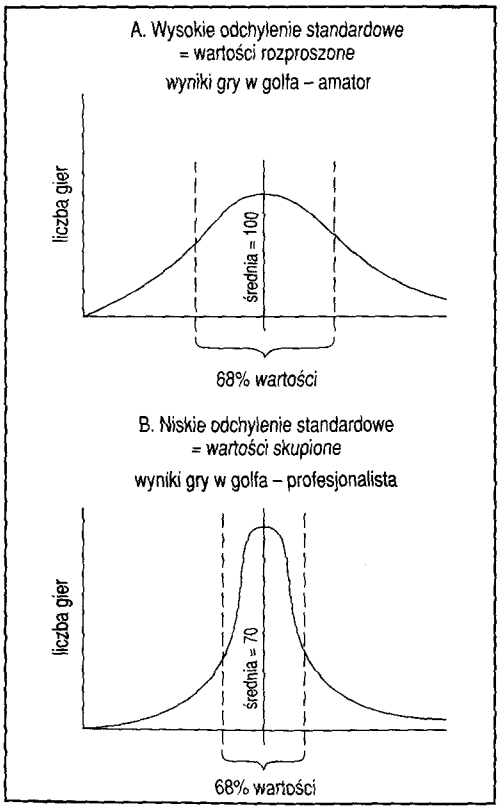
Rozproszenie odnosi się do rozkładu wartości wokół pewnej wartości centralnej, takiej

jak przeciętna. Najprostszym miernikiem rozproszenia jest zakres: odległość między najwyższą a najniższą wartością. W ten sposób, poza stwierdzeniem, że średnia wieku naszych badanych wynosi 15,87, moglibyśmy też wskazać, że ich wiek mieści się w zakresie od 13 do 19 lat.

Bardziej wyrafinowanym miernikiem rozproszenia jest **odchylenie standardowe**. Miernik ten został krótko wspomniany w rozdziale 7 jako błąd standardowy rozkładu z prób. Odchylenie standardowe jest, najprościej mówiąc, wskaźnikiem poziomu zmienności w zbiorze danych. Wyższe odchylenie standardowe oznacza, że dane są bardziej rozproszone; niższe odchylenie standardowe oznacza, że są bardziej skupione. Rysunek 15.3 przedstawia tę zasadniczą cechę. Zauważ, że gracz profesjonalny ma nie tylko niższą średnią wyników – wyniki te są też bardziej spójne, co odzwierciedla niższe odchylenie standardowe. Amator z kolei ma wyższą przeciętną, ale jego wyniki są mniej spójne – czasem idzie mu dużo lepiej, czasem dużo gorzej.

Rysunek 15.3

Wysoka i niska wartość odchylenia standardowego



Jest jeszcze wiele innych mierników rozproszenia. Opisując wyniki testów na inteligencję, badacze mogliby określić *zakres międzykwartylowy* (*interquartile range*), czyli zakres wyników dla środkowych 50% badanych. Gdyby w górnej ćwiartce (górnym kwartylu) wystąpiły wyniki od 120 do 150, a w dolnej ćwiartce (dolnym kwartylu) wyniki od 60 do 90 punktów, raport mógłby stwierdzić, że zakres międzykwartylowy wynosił od 90 do 120 punktów (czyli 30 punktów), ze średnim wynikiem 102.

Zmienne ciągłe i dyskretne

Poprzednio omówione wyliczenia nie stosują się do wszystkich zmiennych. By to zrozumieć, musimy rozróżnić dwa typy zmiennych: ciągłe i dyskretne. **Zmienna ciągła** (lub inaczej zmienna ilorazowa) wzrasta stale w minimalnych przedziałach. Przykładem jest wiek, który wzrasta stale w każdej jednostce czasu. **Zmienna dyskretna** skacze od kategorii do kategorii bez kroków pośrednich. Przykładem może być *płeć*, *stopień wojskowy*, *rok studiów* (przechodzisz z pierwszego na drugi w jednej chwili).

Analizując zmienną dyskretną – na przykład nominalną albo porządkową – nie możesz skorzystać z niektórych omówionych powyżej technik. Ściśle rzecz biorąc, modalną powinno się obliczać dla danych nominalnych, medianę – dla interwałowych, a średnią dla danych ilorazowych, nie dla nominalnych (patrz rozdział 5). Jeśli analizowaną zmienną jest na przykład *płeć*, same liczby (23 osoby w naszej próbie motocyklistów transwestytów to kobiety) lub procenty (7% to kobiety) mogą być użytecznym i stosownym narzędziem analizy, ale ani mediana, ani średnia nie miałyby sensu. Wyliczenie modalnej mogłoby być uzasadnione, acz nie wyjaśniłoby zbyt wiele, dowiedzielibyśmy się bowiem tylko, że „większość to mężczyźni". Modalna z danych na temat przynależności religijnej mogłaby jednak okazać się bardziej interesująca, jak na przykład w przypadku zdania „większość obywateli Stanów Zjednoczonych to protestanci".

Szczegółowość a czytelność

Przedstawiając zarówno dane jednozmiennowe, jak i inne, będziesz się starał osiągnąć dwa cele. Z jednej strony, powinieneś dążyć do tego, aby zapewnić czytelnikowi możliwie najbardziej szczegółowe informacje dotyczące tych danych. Z drugiej jednak strony, dane powinny zostać przedstawione w formie zrozumiałej. Jako że te dwa cele są ze sobą często sprzeczne, będziesz ciągle poszukiwał najlepszego kompromisu między nimi. Praktycznym rozwiązaniem jest przedstawienie konkretnego zbioru danych w więcej niż jednej formie. W przypadku wieku mógłbyś na przykład przedstawić rozkład wieku bez podziału na grupy *oraz* średnią wieku i odchylenie standardowe.

Jak widać na podstawie tego wstępnego omówienia analizy jednozmiennowej, ta pozornie prosta materia bywa dosyć złożona. Wiedza zawarta w niniejszym podrozdziale stanowi jednak pomost do porównań podgrup i analizy dwuzmiennowej.

Porównywanie podgrup

Analiza jednozmiennowa opisuje jednostki analizy, zastosowane w danych badaniach i pozwala, jeśli jednostki te stanowią próbę większej populacji, wyciągnąć wnioski opisowe dotyczące całej populacji. Analizy dwuzmiennowa i wielozmiennowa służą przede wszystkim wyjaśnianiu. Zanim jednak przejdziemy do wyjaśniania, powinniśmy zająć się przypadkiem opisu podgrup.

Często zasadne jest opisywanie podzbiorów przypadków, osób badanych czy respondentów. Tabela 15.2 przedstawia na przykład dane na temat dochodu oddzielnie dla kobiet i mężczyzn. Tabela ta prezentuje dodatkowo stosunek median dochodów kobiet i mężczyzn, wykazując, że we wskazanym okresie zatrudnione kobiety zarabiały nieco ponad połowę tego co zatrudnieni mężczyźni.

W pewnych sytuacjach badacz prezentuje porównania podgrup w celach czysto opisowych. Znacznie częściej celem opisów podgrup jest porównanie: kobiety zarabiają mniej niż mężczyźni. W danym przypadku zakładamy, że jest coś, co wiąże się z byciem kobietą, a co powoduje niższe zarobki. Podobne założenie czynilibyśmy, porównując przychody Afroamerykanów i białych. W takich przypadkach analiza opiera się na założeniu związku przyczynowego: jedna zmienna jest przyczyną drugiej, jak na przykład *płeć* jest przyczyną *dochodu*.

Tabela 15.2

Mediana zarobków całorocznych pełnoetatowych pracowników cywilnych z uwzględnieniem płci: 1967–1977

Rok	Kobiety	Mężczyźni	Stosunek zarobków kobiet do zarobków mężczyzn
1977	8 618	14 626	0,59
1976	8 622	14 323	0,60
1975	8 449	14 175	0,60
1974	8 565	14 578	0,59
1973	8 639	15 254	0,57
1972	8 551	14 778	0,58
1971	8 369	14 064	0,61
1970	8 307	13 993	0,59
1969	8 227	13 976	0,59
1968	7 763	13 349	0,58
1967	7 503	13 021	0,58

Źródło: Opracowanie na podstawie *Statistical Portrait of Women in the United States: 1978* Biura Spisowego, Seria P-23, nr 100, s. 73.

Kiedy Biuro Spisowe opublikowało w 1978 roku dane ujęte w tabeli 15.2, stanowiły one uzasadnienie dla wzrastającego sprzeciwu wobec dyskryminacji kobiet w gospodarce Stanów Zjednoczonych. Zarówno politycy, jak i badacze koncentrują się od owego czasu na tym zagadnieniu, aczkolwiek nie wpłynęło to

Tabela 15.3

Stosunek zarobków kobiet do zarobków mężczyzn wśród pełnoetatowych pracowników i robotników: 1985–1998

Rok	Stosunek mediany tygodniowych zarobków kobiet do mediany tygodniowych zarobków mężczyzn
1985	0,68
1990	0,72
1995	0,75
1998	0,76

Źródło: *Abstrakt statystyczny dla Stanów Zjednoczonych* Biura Spisowego, U.S. Government Printing Office, Waszyngton 1999, s. 445.

zbyt znacząco na rozbieżności w wysokości zarobków kobiet i mężczyzn.

Ostatnie dane statystyczne odzwierciedlają pewną poprawę w tym względzie, nadal jednak wykazują głębokie różnice. W chwili gdy to piszę, przeciętna pełnoetatowa, całoroczna pracownica zarabia 76 centów na każdego dolara zarobionego przez jej kolegę płci męskiej. Aby zapoznać się z aktualnymi proporcjami zarobków, spójrz na tabelę 15.3.

Zanim przejdziemy do omówienia logiki dwuzmiennowej analizy przyczynowej, rozważmy inny przykład porównania podgrup, który pozwoli nam zarazem poruszyć kilka zagadnień z zakresu formatu tabel.

Łączenie kategorii odpowiedzi

„Podręcznikowe przykłady" tabel są na ogół prostsze niż te, które możesz znaleźć w publikowanych raportach z badań albo we własnej analizie danych, dlatego też zarówno ten podrozdział, jak i następny zajmują się dwoma często spotykanymi problemami i sugerują pewne rozwiązania.

Zacznijmy od danych przedstawionych w tabeli 15.4, dotyczących postaw względem ONZ, a zebranych w międzynarodowym sondażu przeprowadzonym przez „New York Times", CBS News i „Herald Tribune" w 1985 roku. Pytanie dotyczyło ogólnego stosunku do tego, jak ONZ spełnia swoje zadania.

Oto pytanie: jak wypada porównanie między ludźmi pięciu narodowości w zakresie ich poparcia dla sposobu działania ONZ? Przeglądając tabelę, uznasz być może, że jest tam zbyt wiele liczb, by zauważyć jakąś znaczącą prawidłowość.

Część problemu z tabelą 15.4 polega na tym, że stosunkowo niewielu respondentów wybierało dwie skrajne kategorie odpowiedzi: bardzo dobrze lub bardzo źle. Co więcej, mimo że kusząca wydaje się możliwość czytania tylko drugiego wiersza tabeli – czyli odpowiedzi „dobrze" – byłoby to zdecydowanie niewskazane. Patrząc tylko na drugi rząd, stwierdzilibyśmy, że najbardziej pozytywny stosunek do ONZ mają Niemcy Zachodnie i USA (46%), tuż za nimi plasowałaby się Francja (45%), Wielka Brytania byłaby mniej przychylna ONZ niż trzy poprzednio wymienione państwa (39%), Japonia zaś (11%) zajmowałaby stanowisko najmniej przychylne ze wszystkich.

Niewłaściwość takiego postępowania polega na tym, że pomijamy wówczas wszystkich respondentów, którzy udzielili najbardziej pozytywnej możliwej odpowiedzi: „bardzo dobrze". W tego typu sytuacji powinno się połączyć lub „scalić" dwa skrajne punkty zakresu zmienności, w tym przypadku połączyć „bardzo dobrze" i „dobrze" oraz „źle" i „bardzo źle". Gdybyś miał tego dokonać w twojej własnej analizie, rozsądne byłoby dodanie częstości i przeliczenie procentów dla połączonych kategorii, jednak analizując opublikowaną tabelę możesz po prostu dodać odpowiednie procenty, jak to zilustrowano w tabeli 15.5.

Dysponując połączonymi kategoriami zilustrowanymi w tabeli 15.5, możemy z łatwością zorientować się w odsetkach obywateli każdego kraju, którzy stwierdzili, że ONZ radzi sobie co najmniej dobrze. Tym razem postawy obywateli Stanów Zjednoczonych wydają się najbardziej pozytywne, Niemcy, Wielka Brytania i Francja są tylko odrobinę mniej pozytywnie usposobione i prawie nie różnią się od siebie nawzajem, Ja-

Tabela 15.4

Postawy względem ONZ: „Jak ONZ radzi sobie z rozwiązywaniem problemów, wobec których staje?"

	Niemcy Zach.	Wielka Brytania	Francja	Japonia	USA
Bardzo dobrze	2%	7%	2%	1%	5%
Dobrze	46%	39%	45%	11%	46%
Źle	21%	28%	22%	43%	27%
Bardzo źle	6%	9%	3%	5%	13%
Nie wiem	26%	17%	28%	41%	10%

Źródło: *5-Nation survey finds hope for U.N.*, „New York Times", 26 czerwca 1985, s. 6.

Tabela 15.5

Łączenie skrajnych kategorii

	Niemcy Zach.	Wielka Brytania	Francja	Japonia	USA
Bardzo dobrze lub dobrze	48%	46%	47%	12%	51%
Źle lub bardzo źle	27%	37%	25%	48%	40%
Nie wiem	26%	17%	28%	41%	10%

ponia zaś jest odosobniona w swoim niskim mniemaniu o działalności ONZ. Mimo że wniosek ten nie różni się radykalnie od tego, który moglibyśmy wyciągnąć na podstawie znajomości tylko drugiego wiersza tabeli 15.4, warto zauważyć, że Wielka Brytania wydaje się teraz względnie bardziej przychylna ONZ.

Chciałbym ci oszczędzić powyżej przedstawionego ryzyka. Załóżmy, że pobieżnie przejrzałeś drugi wiersz odpowiedzi w tabeli 15.4 i uznałeś, że Brytyjczycy mają nieco niższe mniemanie o działalności ONZ niż obywatele Stanów Zjednoczonych, Niemiec i Francji. Mógłbyś więc poszukiwać wyjaśnienia dla tego faktu – być może tworząc genialną teorię psycho-historyczną na temat bolesnego upadku niegdyś potężnego Imperium Brytyjskiego. Gdybyś następnie ujawnił się ze swoją teorią, ktoś inny mógłby ci wykazać, że należyte odczytanie danych dowodzi, że Brytyjczycy nie są wcale bardziej negatywnie usposobieni do ONZ niż którykolwiek z tych trzech narodów. Nie jest to bynajmniej ryzyko czysto hipotetyczne. Błędy tego typu zdarzają się często, można ich jednak uniknąć poprzez scalenie właściwych kategorii odpowiedzi.

Sposoby postępowania z kategorią „nie wiem"

Tabele 15.4 i 15.5 ilustrują zarazem inny powszechny problem, dotyczący analizy danych sondażowych. Pytając ludzi o opinie na dany temat, na ogół dobrze jest pozostawić im możliwość udzielenia odpowiedzi typu: „nie wiem" albo „nie mam zdania". Co jednak zrobić z tymi odpowiedziami w trakcie analizy danych?

Zauważ, że odsetek odpowiedzi „nie wiem" w różnych krajach był w tym przypadku dość zróżnicowany, od 10% w USA do 41% w Japonii. Znaczny udział respondentów, którzy odpowiedzieli „nie wiem", może zafałszować wyniki tego typu tabeli. Czy Japończycy naprawdę znacznie rzadziej twierdzili, że ONZ funkcjonuje dobrze, skoro tak wielu spośród nich nie wyraziło na ten temat żadnej opinii?

Oto prosty sposób przeliczenia procentów z wyłączeniem „nie wiem". Spójrz na pierwszą kolumnę procentową w tabeli 15.5: odpowiedzi Niemców na pytanie o działanie ONZ. Zauważ, że 26% respondentów odpowiedziało „nie wiem". Oznacza to, że ci, którzy odpowiedzieli, że „dobrze" lub „źle", stanowili w sumie 74% (100 minus 26) ogółu. Jeśli podzielimy 48% twierdzących, że ONZ radzi sobie „dobrze" lub „bardzo dobrze", przez 0,74 (odsetek tych, którzy wyrazili jakikolwiek pogląd), możemy stwierdzić, że 65% „mających jakiekolwiek zdanie" uznało, że ONZ radzi sobie dobrze lub bardzo dobrze (48%/0,74 = 65%).

Tabela 15.6 przedstawia całą tabelę z wyłączeniem odpowiedzi „nie wiem". Zauważ, że te nowe dane skłaniają do nieco innej interpretacji niż poprzednie. Wydaje się w szczególności, że najbardziej pozytywną ocenę wysiłków ONZ wyrazili Francuzi i Niemcy, natomiast opinie w USA i Wielkiej Brytanii były nieco gorsze. Japonia, mimo że wciąż pozostaje na ostatniej pozycji, zmieniła swój wynik z 12% na 20% ocen pozytywnych.

W tym miejscu, po przyjrzeniu się trzem wersjom powyższych danych, mógłbyś zadać sobie pytanie: która jest prawidłowa? Odpowiedź zależy od celu twojej analizy i interpretacji danych. Jeśli na przykład nie jest ci potrzebne rozróżnienie odpowiedzi „bardzo dobrze" i „dobrze", połączenie ich ma sens, ponieważ ułatwia odczytanie tabeli.

Znacznie trudniejsze jest podjęcie poza kontekstem konkretnych badań decyzji, czy

Tabela 15.6

Pominięcie odpowiedzi „nie wiem"

	Niemcy Zach.	Wielka Brytania	Francja	Japonia	USA
Bardzo dobrze lub dobrze	65%	55%	65%	20%	57%
Źle lub bardzo źle	35%	45%	35%	81%	44%

uwzględnić, czy też wyłączyć odpowiedzi „nie wiem". Odkrycie, że tak duży odsetek Japończyków nie ma zdania, mogłoby się okazać bardzo istotne – gdybyś chciał się na przykład dowiedzieć, czy ludzie są świadomi działalności ONZ. Gdybyś jednak chciał się na przykład dowiedzieć, jak ludzie będą głosować w danej sprawie, rozsądniej byłoby wyłączyć braki odpowiedzi na mocy przypuszczenia, że osoby te nie będą głosować, lub też że ich głosy podzielą się równo między dwoma stanowiskami w danej kwestii.

Prawda o twoich danych jest w każdym razie taka, że pewien odsetek badanych powiedział, że „nie wie", głosy reszty zaś podzieliły się w pewien określony sposób. Często najlepszym sposobem jest przedstawienie danych w obu formach – z odpowiedziami „nie wiem" i po ich wyłączeniu, co pozwoli czytelnikom wyciągnąć własne wnioski.

Opisy liczbowe w badaniach jakościowych

Mimo że w tym rozdziale zajmujemy się przede wszystkim badaniami ilościowymi, rozważania te znajdują również zastosowanie w analizach jakościowych. Wyniki pogłębionych studiów jakościowych często dają się sprawdzić w drodze testów liczbowych. David Silverman na przykład, kiedy postanowił porównać leczenie pacjentów chorych na raka w klinikach prywatnych i w placówkach brytyjskiej Narodowej Służby Zdrowia, zdecydował się przede wszystkim na pogłębioną analizę interakcji między pacjentami a lekarzami:

> „Moja metoda analizy była w znacznym stopniu jakościowa. (...) Używałem fragmentów wypowiedzi lekarzy i pacjentów, podając zarazem zwięzły opis etnograficzny otoczenia oraz pewnych danych behawioralnych. Dodatkowo skonstruowałem jednak formę kodowania, która pozwoliła mi ustalić kilka ogólnych mierników interakcji między pacjentami a lekarzami". (1993, s. 163)

Dane liczbowe doprecyzowały intuicje Silvermana oparte na jego obserwacjach jakościowych, zarazem jednak pogłębione zrozumienie sytuacji pozwoliło mu na wypracowanie tym trafniejszych metod analizy ilościowej. Przyjrzyj się relacjom między podejściem ilościowym a jakościowym w poniższym obszernym wywodzie:

> „Moje ogólne wrażenie było takie, że konsultacje prywatne trwają znacznie dłużej niż w klinikach publicznych. Po dokładnym zbadaniu danych okazało się, że te pierwsze trwały istotnie prawie dwa razy dłużej (20 minut w stosunku do 11) i że różnica ta była statystycznie bardzo istotna. Przypomniałem sobie jednak, że jedna z klinik publicznych miała, ze szczególnych względów, niezwykle krótki czas konsultacji. Uznałem, że uczciwe porównanie czasu konsultacji w tych dwóch sektorach wymaga wykluczenia tej kliniki oraz porównywania tylko czasu konsultacji z udziałem konkretnego lekarza w obu sektorach. Badanie tej ograniczonej próbki wykazało, że różnica czasu trwania konsultacji prywatnych i publicznych spadła do poniżej 3 minut. Wynik ten był nadal istotny statystycznie, acz w mniejszym stopniu. W końcu jednak gdyby porównać czas poświęcony tylko nowym pacjentom tego samego lekarza, to pacjentom publicznej służby zdrowia poświęcano przeciętnie o 4 minuty więcej – 34 minuty w porównaniu do 30 w klinikach prywatnych". (1993, s. 163–164)

Powyższy przykład ujawnia zalety łączenia różnych podejść w badaniach społecznych. Szczególnie wartościowe może być połączenie metod ilościowych i jakościowych.

Analiza dwuzmiennowa

W przeciwieństwie do analizy jednozmiennowej, porównania podgrup wymagają dwóch zmiennych. Pod tym względem analiza podgrup stanowi rodzaj **analizy dwuzmiennowej** – czyli jednoczesnej analizy dwóch zmiennych. Jednakże, podobnie jak w przypadku analizy jednozmiennowej, celem porównań podgrup jest przede wszystkim opis. Analiza dwuzmiennowa w badaniach społecznych zawiera na ogół jeszcze jeden element: określenie zależności między zmiennymi jako takimi. Tak więc analiza jednozmiennowa i porównania podgrup mają na celu opis ludzi (czy innych jednostek analizy), analiza dwuzmiennowa zaś koncentruje się na zmiennych i empirycznych zależnościach między nimi.

Tabela 15.7 może stanowić przykład porównania podgrup: przedstawia ona niezależnie uczestniczenie w mszach i nabożeństwach kobiet i mężczyzn, zgodnie z wynikami GSS z 1996 roku. Tabela pokazuje – porównawczo i opisowo – dane dowodzące, że badane kobiety uczestniczyły w mszach i nabożeństwach częściej niż mężczyźni. Ta sama tabela jednak, gdy potraktujemy ją jako materiał dla służącej wyjaśnianiu analizy dwuzmiennowej, opowiada o czymś zupełnie innym. Dowiadujemy się z niej, że zmienna *płeć* ma wpływ na zmienną *uczestniczenie w mszach i nabożeństwach*. Oznacza to, że możemy widzieć ten typ zachowania jako zmienną zależną, która jest częściowo zdeterminowana przez zmienną niezależną, czyli *płeć*.

Wyjaśniająca analiza dwuzmiennowa wymaga więc użycia „języka zmiennych", wprowadzonego w rozdziale 1. Przesuwając lekko akcenty, nie mówimy już o mężczyznach i kobietach jako różnych podgrupach, lecz o *płci* jako zmiennej, która ma wpływ na inne zmienne. Teoretyczna interpretacja tabeli 15.7 mogłaby się opierać na hipotezie pocieszenia Charlesa Glocka, omówionej w rozdziale 2:

1. Kobiety są w Stanach Zjednoczonych nadal traktowane jako obywatele drugiej kategorii.
2. Ludzie pozbawieni nagród w dziedzinie statusu społecznego mogą się zwracać ku religii jako alternatywnemu źródłu statusu.
3. Wynika stąd, że kobiety powinny być bardziej religijne niż mężczyźni.

Dane, przedstawione w tabeli 15.7, potwierdzają to rozumowanie; 34% kobiet i tylko 25% mężczyzn uczestniczy w mszach i nabożeństwach co tydzień.

Dodatek w postaci logiki zależności przyczynowych ma istotne implikacje dla tworzenia i odczytywania tabel procentowych. Jednym z zasadniczych problemów początkujących analityków jest ustalenie „kierunku procentowania" danej tabeli. W tabeli 15.7 na przykład podzieliłem grupę badanych na dwie podgrupy – mężczyzn i kobiety – po czym opisałem zachowanie każdej z podgrup. Jest to poprawna metoda tworzenia tego typu tabeli. Zwróć jednak uwagę, że moglibyśmy – w sposób niepoprawny – skonstruować tę tabelę inaczej. Moglibyśmy najpierw podzielić badanych wedle częstości ich uczestnictwa w mszach i nabożeństwach, następnie zaś ustalić odsetek kobiet i mężczyzn w każdej z podgrup. Metoda taka byłaby jednak bezwartościowa z punktu widzenia wyjaśniania. Tabela 15.7 sugeruje, że płeć ma wpływ na częstość uczestnictwa w mszach i nabożeństwach. Gdybyśmy użyli innej metody konstrukcji, tabela sugerowałaby, że uczestnictwo danej osoby w mszach i nabożeństwach ma wpływ na to, czy jest ona mężczyzną, czy kobietą – co nie ma sensu. Zachowanie nie może określać płci.

Pewien podobny problem utrudnia życie początkującym analitykom. Jak czytać tabelę procentową? Istnieje pokusa odczytania tabeli 15.7 w sposób następujący: „spośród kobiet tylko 34% uczestniczyło w mszach i nabożeństwach co tydzień, 66% zaś uczestniczyło w nich rzadziej: bycie kobietą wywołuje zatem spadek prawdopodobieństwa, że dana osoba

Tabela 15.7

Uczestniczenie w mszach i nabożeństwach deklarowane przez mężczyzn i kobiety w 1996 roku

	Mężczyźni	Kobiety
Co tydzień	25%	34%
Rzadziej	75%	66%
100% =	(901)	(1134)

Źródło: Generalny Sondaż Społeczny, Krajowe Centrum Badania Opinii.

będzie często uczestniczyła w mszach i nabożeństwach". Takie odczytanie tabeli jest oczywiście niepoprawne. Jakikolwiek wniosek co do wpływu zmiennej *płeć* na uczestnictwo w mszach i nabożeństwach musi opierać się na porównaniu kobiet i mężczyzn. W szczególności możemy porównać 34% z 25% i stwierdzić, że jest bardziej prawdopodobne, że kobieta będzie uczestniczyć w mszach i nabożeństwach co tydzień. Porównanie podgrup jest więc konieczne dla odczytania tabeli dwuzmiennowej służącej wyjaśnianiu.

Tworząc i przedstawiając tabelę 15.7, zastosowałem konwencję zwaną *procentowaniem w dół (w kolumnach)*. Termin ten oznacza, że w każdej kolumnie odsetki sumują się do pełnych 100%. Tego typu tabelę czyta się rzędami. W rzędzie oznaczonym „co tydzień" sprawdzasz, jaki odsetek mężczyzn uczestniczył w mszach i nabożeństwach co tydzień oraz jaki odsetek kobiet uczestniczył w mszach i nabożeństwach co tydzień.

Kierunek procentowania w tabeli jest arbitralny, niektórzy badacze wolą procentowanie poziome. Ułożyliby oni tabelę 15.7 tak, by „mężczyźni" i „kobiety" znaleźli się po lewej stronie tabeli, oznaczając dwa rzędy, natomiast „co tydzień" i „rzadziej" pojawiłyby się na górze jako oznaczenie kolumn. Liczby w polach tabeli przesunęłyby się odpowiednio, odsetki w każdym rzędzie zaś sumowałyby się do 100%. W takim przypadku tabelę czytałoby się wzdłuż kolumn, szukając odpowiedzi na pytanie, jaki odsetek kobiet i mężczyzn uczestniczył w mszach i nabożeństwach z jaką częstotliwością. Logika i wnioski byłyby w obu przypadkach takie same, zmieniłaby się jedynie forma.

Odczytując tabelę skonstruowaną przez kogoś innego, musisz ustalić, w jakim kierunku przebiegało procentowanie. Na ogół będzie to oznaczone lub będzie wynikać w oczywisty sposób z logiki analizowanych zmiennych. W ostateczności powinieneś dodać odsetki w każdej kolumnie i każdym rzędzie. Jeśli każda z kolumn daje 100%, procentowano w *dół (w kolumnach)*. Jeśli w każdym rzędzie otrzymujesz 100%, procentowanie jest poziome (w wierszach). Zasada jest następująca:

1. Jeśli procentowano w kolumnach, czytaj wierszami.
2. Jeśli procentowano w wierszach, czytaj kolumnami.

Procentowanie w tabeli

Rysunek 15.4 przedstawia logikę tworzenia tabeli procentowej dla dwóch zmiennych. Wykorzystałem zmienne *płeć* i *postawy wobec równouprawnienia mężczyzn i kobiet*.

Oto inny przykład. Załóżmy, że interesujemy się polityką redakcyjną gazet, dotyczącą kwestii legalizacji marihuany. Prowadzimy analizę treści artykułów redakcyjnych na ten temat, które pojawiły się w ciągu danego roku w próbie dzienników w całym kraju. Każdy artykuł redakcyjny klasyfikujemy jako sprzyjający, neutralny lub wrogi względem legalizacji marihuany. Moglibyśmy na przykład zbadać zależność między poglądami wyrażonymi w artykule a typem miejscowości, w której dany dziennik jest wydawany, zakładając, że społeczności wiejskie powinny być pod tym względem bardziej konserwatywne niż miejskie. Tak więc każdy dziennik – a więc i każdy artykuł redakcyjny – został zaklasyfikowany również w kategoriach liczby ludności w miejscowości, w której jest wydawany.

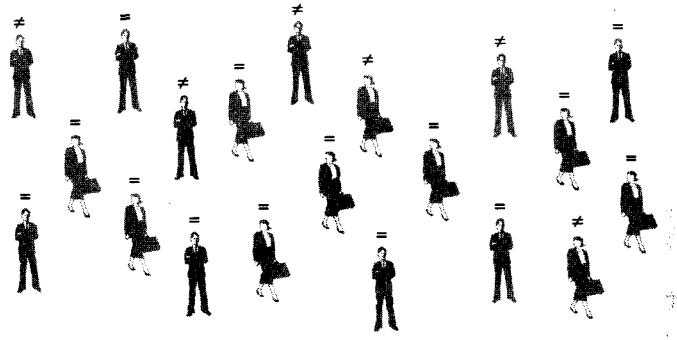
Tabela 15.8 przedstawia hipotetyczne dane opisujące poglądy wyrażane w artykułach redakcyjnych w dziennikach miejskich i wiejskich. Zauważ, że jednostką analizy jest w tym przypadku pojedynczy artykuł. Tabela 15.8 zawiera informację, że 127 artykułów redakcyjnych na temat marihuany w naszej próbie dzienników zostało opublikowanych w spo-

Tabela 15.8

Hipotetyczne dane dotyczące treści artykułów redakcyjnych na temat legalizacji marihuany

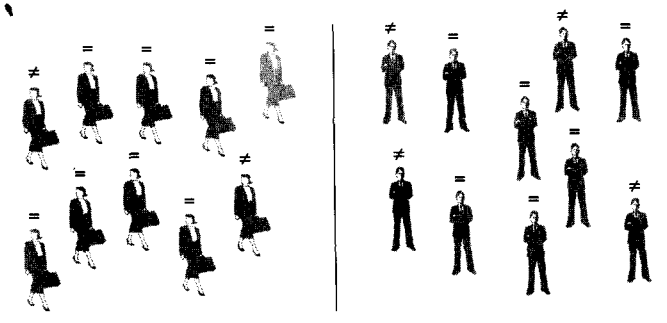
Polityka redakcyjna wobec legalizacji marihuany	Liczba mieszkańców	
	poniżej 100 000	powyżej 100 000
Sprzyjająca	11%	32%
Neutralna	29%	40%
Wroga	60%	28%
100% =	(127)	(438)

A. Mężczyźni i kobiety popierający (=) lub nie popierający (≠) równouprawnienia płci



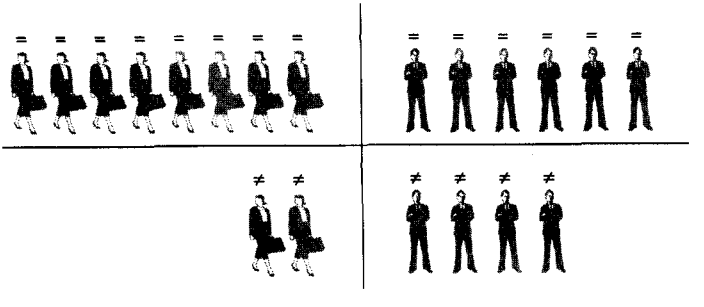
B. Oddzielamy mężczyzn od kobiet (zmienna niezależna)

kobiety mężczyźni



C. W obrębie każdej z płci oddzielamy tych, którzy popierają równouprawnienie, od tych, którzy go nie popierają (zmienna zależna)

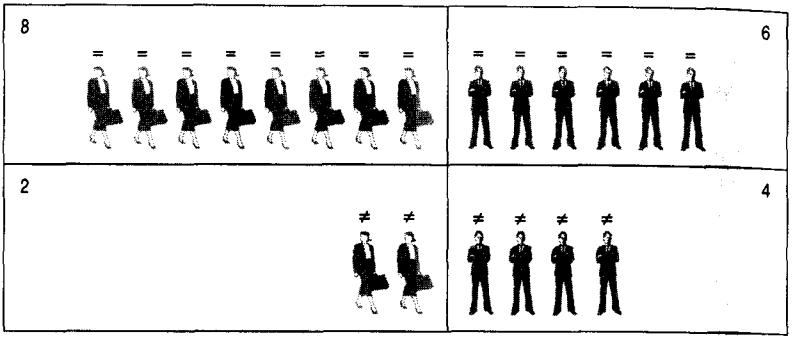
kobiety mężczyźni



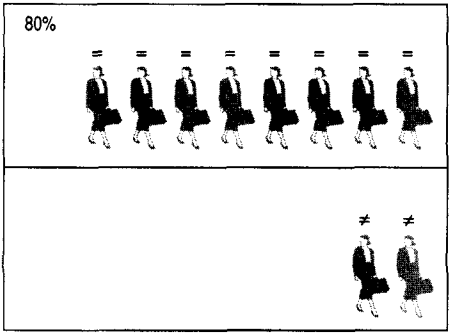
Rysunek 15.4 (ciąg dalszy)

Procentowanie w tabeli

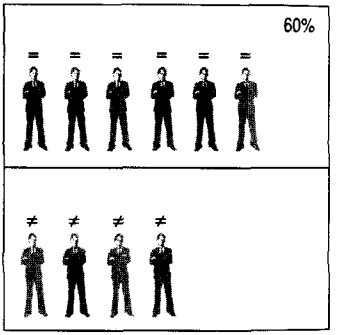
D. Liczymy, ile przypadków jest w każdej komórce tabeli



E. Jaki procent kobiet popiera równouprawnienie?



F. Jaki procent mężczyzn popiera równouprawnienie?



G. Wnioski

Prawdopodobieństwo, że kobieta popiera równouprawnienie, jest większe niż w przypadku mężczyzn, aczkolwiek większość przedstawicieli obu płci popiera równouprawnienie.

Wydaje się więc, że płeć jest jedną z przyczyn postaw wobec równouprawnienia płci.

	kobiety	mężczyźni
popierają równouprawnienie	80%	60%
nie popierają równouprawnienia	20%	40%
ogółem	100%	100%

łecznościach o populacji poniżej 100 tys. mieszkańców. (*Uwaga*: ten punkt graniczny został wybrany dla uproszczenia przykładu i nie oznacza, że słowo „wiejski" w jakimkolwiek obiektywnym sensie odnosi się do społeczności poniżej 100 tys. mieszkańców). Spośród nich 11% (14 artykułów) sprzyjało legalizacji marihuany, 29% było neutralnych, 60% zaś – wrogich. Spośród 438 artykułów, które pojawiły się w próbie dzienników wydawanych w miejscowościach powyżej 100 tys. mieszkańców, 32% (140 artykułów) zaliczono do sprzyjających legalizacji marihuany, 40% do obojętnych, 28% zaś – do wrogich.

Porównując politykę redakcyjną gazet miejskich i wiejskich w naszych hipotetycznych badaniach, stwierdzamy – zgodnie z oczekiwaniami – że dzienniki wiejskie mniej niż miejskie sprzyjają legalizacji marihuany. Określamy to, odnotowując fakt, że więcej (32%) artykułów redakcji miejskich niż wiejskich (11%) sprzyjało legalizacji marihuany. Moglibyśmy również zauważyć, że więcej artykułów redakcji wiejskich niż miejskich było wrogich temu pomysłowi (60% do 28%). Zauważ, że tabela ta zakłada, że *wielkość miejscowości może wpływać* na poglądy wyrażane przez redakcję dziennika na ten temat, nie zaś, że poglądy redakcji wpływają na wielkość miejscowości.

Tworzenie i odczytywanie tabel dwuzmiennowych

Przyjrzyjmy się teraz kolejnym etapom tworzenia tabeli dwuzmiennowej służącej wyjaśnianiu:

1. Dzielimy przypadki na grupy zgodnie z wartościami zmiennej niezależnej.
2. Opisujemy następnie każdą z podgrup w kategoriach wartości zmiennej zależnej.
3. Na koniec odczytujemy tabelę, porównując między sobą podgrupy zmiennej niezależnej w kategoriach danej wartości zmiennej zależnej.

Powtórzmy analizę płci i postaw wobec równouprawnienia, postępując zgodnie z powyższym opisem. Z powyżej przedstawionych względów *płeć* jest zmienną niezależną, *postawy wobec równouprawnienia płci* stanowią zmienną zależną. Postępujemy zatem w sposób następujący:

1. Dzielimy przypadki na mężczyzn i kobiety.
2. Każda podgrupa płci zostaje opisana w kategoriach poparcia lub braku poparcia dla równouprawnienia płci.
3. Porównujemy mężczyzn i kobiety pod kątem odsetka osób popierających równouprawnienie.

W przykładzie dotyczącym polityki redakcyjnej wobec legalizacji marihuany *wielkość miejscowości* była zmienną niezależną, *polityka redakcyjna gazety* zaś – zmienną zależną. Tabelę tworzyłbyś wówczas następująco:

1. Podziel artykuły na podgrupy według wielkości miejscowości, w których wydawane są badane dzienniki.
2. Opisz każdą podgrupę artykułów według odsetka artykułów przyjaznych, obojętnych lub wrogich wobec legalizacji marihuany.
3. Porównaj obie podgrupy w kategoriach odsetków artykułów przyjaznych legalizacji marihuany.

Analizy dwuzmiennowe mają na ogół na celu wyjaśnianie przyczynowe. Powyższe hipotetyczne przykłady rzucają nieco światła na naturę przyczynowości w naukach społecznych. Rozwinięcie tego dość uproszczonego opisu znajdziesz w rozdziale 16, w którym zajmiemy się znacznie dokładniej złożoną naturą przyczynowości.

Tabele podobne do przed chwilą analizowanych nazywamy na ogół **tabelami krzyżowymi (tabelami kontyngencji)**. Wartości zmiennej zależnej są uzależnione od wartości zmiennej niezależnej. Pomimo popularności tego typu tabel w naukach społecznych ich format nie został nigdy zestandaryzowany, wskutek czego znajdziesz wiele rozmaitych formatów w literaturze badawczej. Nie ma zresztą prawdopodobnie sensu dążyć do standaryzacji, póki tabele są łatwe do odczytania i interpretacji. Jest jednak kilka wskazówek, którymi powinieneś się kierować podczas prezentacji większości danych tabelarycznych:

1. Tabela powinna mieć nagłówek lub tytuł opisujący zwięźle jej zawartość.
2. Oryginalna treść zmiennej powinna zostać jasno przedstawiona, jeśli to możliwe – w samej tabeli lub w tekście wraz z parafrazą w tabeli. Informacja ta jest szczególnie istotna, gdy zmienna została wyprowadzona z odpowiedzi na pytania o postawy, wówczas bowiem znaczenie odpowiedzi zależy w znacznym stopniu od sformułowania pytania.
3. Należy wyraźnie wskazać wartości każdej zmiennej. Mimo że konieczne okaże się skrócenie rozbudowanych kategorii, ich znaczenie powinno zostać objaśnione w tabeli, oprócz tego zaś pełny opis powinien oczywiście znaleźć się w tekście.

4. Jeśli tabela przedstawia odsetki, należy podać podstawę ich obliczenia. Nie ma sensu podawać wszystkich surowych liczb dla każdej kategorii, ponieważ można je odtworzyć na podstawie procentów i podstaw procentowych. Co więcej, prezentacja zarówno liczb, jak i odsetków często komplikuje tabelę i utrudnia jej czytanie.
5. Jeśli w tabeli pominięto jakieś przypadki ze względu na brakujące dane (na przykład brak odpowiedzi), ich liczba powinna zostać podana w tabeli.

Tabela 15.9

Zależność wielozmiennowa: uczestniczenie w mszach i nabożeństwach, płeć i wiek

	Jak często uczestniczysz w mszach i nabożeństwach?			
	poniżej 40 lat		40 lat i więcej	
	mężczyźni	kobiety	mężczyźni	kobiety
Mniej więcej co tydzień *	24%	32%	33%	48%
Rzadziej	76%	68%	67%	52%
100 % =	(325)	(383)	(323)	(452)

* „Mniej więcej co tydzień" = „częściej niż raz na tydzień", „co tydzień" i „prawie co tydzień".

Źródło: Próba losowa respondentów GSS z lat 1973, 1983 i 1993.

Wprowadzenie do analizy wielozmiennowej

Logika *analizy wielozmiennowej*, czyli jednoczesnej analizy więcej niż dwóch zmiennych, jest głównym tematem rozdziałów 16 i 17. W tym miejscu wydaje się jednak celowe rozwinięcie dotychczasowych rozważań na temat tabel dwuzmiennowych poprzez rzut oka na budowę tabel wielozmiennowych.

Tabelę wielozmiennową tworzy się na podstawie bardziej złożonego opisu podgrup w zasadniczo ten sam sposób, jak tabelę dwuzmiennową. Jednak zamiast jednej zmiennej zależnej i jednej zmiennej niezależnej występuje więcej niż jedna zmienna niezależna. Zamiast wyjaśniania zmiennej zależnej za pomocą jednej zmiennej niezależnej, poszukujemy wyjaśnienia, wykorzystując więcej niż jedną zmienną niezależną.

Wróćmy do przykładu uczestniczenia w mszach i nabożeństwach. Załóżmy, że jesteśmy przekonani, że również wiek wpływa na tego rodzaju zachowania (według hipotezy pocieszenia Glocka ludzie starsi są bardziej religijni niż młodsi). W pierwszym kroku podzielibyśmy więc próbę na podgrupy na podstawie wartości obu zmiennych niezależnych jednocześnie: młodzi mężczyźni, starsi mężczyźni, młode kobiety i starsze kobiety. Następnie kilka podgrup zostałoby opisanych w kategoriach zmiennej zależnej, *uczestniczenia w mszach i nabożeństwach*, po czym nastąpiłby etap porównawania. Efekt przedstawiono w tabeli 15.9, sporządzonej na podstawie analizy wyników danych Generalnego Sondażu Społecznego z lat 1973–1993.

W tabeli 15.9 procentowania dokonano w kolumnach, dlatego należy ją czytać poziomo. Interpretacja tej tabeli prowadzi do kilku wniosków.

1. Zarówno wśród mężczyzn, jak i kobiet ludzie starsi uczestniczą w praktykach religijnych częściej niż młodzi. 32% kobiet poniżej 40 roku życia i 48% kobiet w wieku 40 lat lub więcej uczestniczyło w mszach i nabożeństwach co tydzień. Liczby te dla mężczyzn wynoszą odpowiednio 24% i 33%.
2. W obrębie każdej z grup wiekowych kobiety uczestniczyły w mszach i nabożeństwach częściej niż mężczyźni. Wśród respondentów poniżej 40 roku życia co tydzień uczestniczyło w nich 32% kobiet i tylko 24% mężczyzn. Wśród osób w wieku 40 lat i więcej co tydzień uczestniczyło w mszach i nabożeństwach 48% kobiet i 33% mężczyzn.
3. Zgodnie z danymi w tabeli płeć i wiek mają porównywalny wpływ na uczestnictwo w mszach i nabożeństwach.
4. Płeć i wiek wpływają na uczestnictwo w mszach i nabożeństwach niezależnie. Przy danej wartości jednej ze zmiennych niezależnych zmiana wartości drugiej nadal wywiera wpływ na zachowanie.

Dwie zmienne niezależne wywierają również łączny wpływ na zachowanie. Starsze kobiety uczestniczą w mszach i nabożeństwach najczęściej (48%), młodzi mężczyźni najrzadziej (24%).

W rozdziale 16, dotyczącym modelu analizy rozbudowanej, zbadamy dokładniej logikę analizy wielozmiennowej. Zanim jednak zamkniemy tę część, warto zauważyć możliwość prezentacji tego typu danych w alternatywnym formacie.

Niektóre tabele w tym rozdziale nie są dostatecznie użyteczne. Kiedy zmienna zależna, *uczestniczenie w mszach i nabożeństwach*, jest dychotomiczna (ma tylko dwie wartości), znajomość jednej z nich pozwala czytelnikowi z łatwością odtworzyć drugą. W ten sposób, wiedząc, że 32% kobiet poniżej 40 roku życia uczestniczy w mszach i nabożeństwach co tydzień, wiemy również automatycznie, że 68% uczestniczy w mszach i nabożeństwach rzadziej. Tak więc nie ma potrzeby podawania odsetka tych, które uczestniczą w mszach i nabożeństwach rzadziej.

Zgodnie z powyższym rozumowaniem, tabela 15.9 może być przedstawiona w alternatywnym formacie tabeli 15.10. W tabeli 15.10 w polach odpowiadających kombinacjom dwóch zmiennych niezależnych przedstawione zostały odsetki osób twierdzących, że uczestniczą w mszach i nabożeństwach co tydzień. Liczby przedstawione w nawiasach poniżej każdego odsetka odpowiadają liczbie przypadków, na podstawie której obliczono dany odsetek. W ten sposób czytelnik wie na przykład, że próba obejmuje 383 kobiety poniżej 40 roku życia, spośród których 32%

Tabela 15.10

Uproszczenie tabeli 15.9

	Odsetek uczestniczących w mszach i nabożeństwach mniej więcej co tydzień	
	mężczyźni	kobiety
Poniżej 40 lat	24%	32%
	(325)	(383)
40 lat i więcej	33%	48%
	(323)	(452)

Tabela 15.11

Płeć, czas trwania stosunku pracy a zarobki (pracownicy pełnoetatowi w wieku 21–64 lata)

Staż pracy u aktualnego pracodawcy	Przeciętna stawka za godzinę pracy		Kobiety/mężczyźni
	mężczyźni	kobiety	proporcja
Poniżej 2 lat	8,46 USD	6,03 USD	0,71
Od 2 do 4 lat	9,38 USD	6,78 USD	0,72
Od 5 do 9 lat	10,42 USD	7,56 USD	0,73
10 lat i więcej	12,38 USD	7,91 USD	0,64

Źródło: Biuro Spisowe Stanów Zjednoczonych, Current Population Reports, Seria P-70, nr 10, *Male-Female Differences in Work Experience, Occupation and Earning, 1984*, U.S. Governement Printing Office, Washington, DC 1987, s. 4.

uczestniczy w praktykach religijnych co tydzień. Możemy na tej podstawie obliczyć, że 123 spośród tych 383 kobiet uczestniczy w mszach i nabożeństwach co tydzień, a pozostałe 260 młodszych kobiet (lub 68%) uczestniczy w nich rzadziej. Nowa tabela jest łatwiejsza do odczytania niż poprzednia, przy tym zaś nie tracimy żadnego szczegółu.

By posłużyć się innym prostym przykładem analizy wielozmiennowej, wróćmy do kwestii płci i dochodu. Podejmowano już wiele prób wyjaśnienia trwałości zjawiska niższych zarobków pracujących kobiet niż pracujących mężczyzn. Jedno z wyjaśnień odwołuje się do faktu, że ze względu na tradycyjny wzorzec rodziny kobiety mają mniejszy udział w ogólnej liczbie pracowników, a wiele spośród nich zaczyna pracę dopiero po wykonaniu pewnych obowiązków związanych z wychowaniem dzieci. Tak więc staż pracy w grupie kobiet będzie prawdopodobnie niższy niż w przypadku mężczyzn, a zarobki rosną wraz z czasem pracy. Badania Biura Spisowego z 1984 roku wykazały, że rozumowanie to jest *częściowo* prawdziwe, jak to uwidoczniono w tabeli 15.11.

Tabela 15.11 pozwala przede wszystkim stwierdzić, że czas trwania stosunku pracy istotnie wpływa na zarobki. Zarówno mężczyźni, jak i kobiety z dłuższym stażem zarabiają więcej. Można to zauważyć, czytając pionowo dwie pierwsze kolumny tabeli.

Tabela wskazuje również, że kobiety zarabiają mniej niż mężczyźni, niezależnie od stażu

pracy. Można to zauważyć, porównując przeciętne stawki w rzędach tabeli oraz proporcje stawek kobiet i mężczyzn, przedstawione w trzeciej kolumnie. Tak więc czas trwania stosunku pracy jest istotną determinantą zarobków, jednak staż pracy nie stanowi adekwatnego wyjaśnienia prawidłowości, zgodnie z którą kobiety zarabiają mniej niż mężczyźni. W rzeczywistości kobiety z ponaddziesięcioletnim stażem zarabiają wyraźnie mniej (7,91 dolara za godzinę) niż mężczyźni ze stażem nie przekraczającym dwóch lat (8,46 dolara za godzinę).

Skoro czas trwania stosunku pracy nie wyjaśnia różnicy między płacami mężczyzn i kobiet, możliwe są również inne wyjaśnienia: poziom wykształcenia, obowiązki związane z opieką nad dziećmi itp. Badacze, którzy stworzyli tabelę 15.11, zbadali także inne zmienne, które mogłyby wyjaśnić różnice zarobków, nie będąc przy tym wyrazem dyskryminacji ze względu na płeć. Oprócz stażu pracy u obecnego pracodawcy, wzięli oni pod uwagę następujące zmienne:

- Liczba lat pracy w obecnym zawodzie
- Ogólny staż pracy (niezależnie od zawodu)
- Fakt, czy respondent(ka) pracował(a) zazwyczaj na pełny etat
- Stan cywilny
- Wielkość miejsca zamieszkania
- Fakt bycia objętym zbiorowym układem pracy
- Rodzaj pracy
- Liczba pracowników w zakładzie pracy
- Pracodawca prywatny lub publiczny
- Przyczyny zwolnienia z poprzedniej pracy
- Upływ czasu między obecną a poprzednią pracą
- Rasa
- Niepełnosprawność
- Stan zdrowia
- Wiek dzieci
- Uczestnictwo w zajęciach akademickich w szkole średniej
- Liczba zajęć z matematyki, nauk ścisłych i przyrodniczych oraz języków obcych w szkole średniej
- Fakt uczęszczania do prywatnej lub publicznej szkoły średniej
- Osiągnięty poziom wykształcenia
- Odsetek kobiet zatrudnionych w danym zawodzie
- Główny kierunek studiów

Każda z tych zmiennych mogłaby wpływać na zarobki i służyć jako wyjaśnienie różnic zarobków kobiet i mężczyzn, gdyby okazało się, że różnią się oni pod którymś z powyższych względów. Wziąwszy wszystkie te zmienne pod uwagę, badacze zdołali wyjaśnić 60% nierówności zarobków kobiet i mężczyzn. Pozostałe 40% jest więc funkcją innych „rozsądnych" zmiennych oraz/lub uprzedzeń. Do tego typu wniosków można dojść tylko badając wpływ wielu zmiennych jednocześnie – to znaczy w drodze analizy wielozmiennowej.

Omówiliśmy już cały proces analizy danych. W rozdziałach 16 i 17 będziemy dalej zgłębiać materiał przedstawiony w ostatnich kilku rozdziałach.

Główne zagadnienia

- Analiza jednozmiennowa to analiza pojedynczej zmiennej. Ponieważ analiza jednozmiennowa nie obejmuje związków między dwiema i więcej zmiennymi, służy ona raczej celom opisowym niż wyjaśnianiu.
- Wiele technik pozwala badaczom podsumować ich pierwotne dane, by uczynić je bardziej zrozumiałymi przy zachowaniu tak wielu szczegółów, jak to tylko możliwe. Rozkład częstości, wartości przeciętne, grupowanie danych, miary rozproszenia – to sposoby podsumowania danych dotyczących jednej zmiennej.
- Porównania podgrup mogą być wykorzystywane do opisu podobieństw i różnic między podgrupami w zakresie tej samej zmiennej.
- Analiza dwuzmiennowa koncentruje się raczej na zależnościach między zmiennymi niż na porównywaniu podgrup. Analiza dwuzmiennowa odsłania powiązania statystyczne między zmienną zależną i niezależną. Służy ona na ogół raczej wyjaśnianiu niż tylko opisowi.

Wyniki analizy dwuzmiennowej przedstawiane są często w formie tabel krzyżowych (tabel kontyngencji), które tworzy się w celu wyjaśnienia wpływu zmiennej niezależnej na zmienną zależną.

- Analiza wielozmiennowa jest metodą analizy jednoczesnych zależności między wieloma zmiennymi. Może być także wykorzystywana dla pełniejszego zrozumienia zależności między dwiema zmiennymi.
- Mimo że omówione w tym rozdziale tematy kojarzone są zwykle przede wszystkim z badaniami ilościowymi, przedstawione powyżej zasady i techniki badawcze mogą być również cenne dla badań jakościowych.

Kluczowe pojęcia

analiza wielozmiennowa
analiza jednozmiennowa
rozkład częstości
wartość przeciętna
średnia
analiza dwuzmiennowa
modalna
mediana
rozproszenie
odchylenie standardowe
zmienna ciągła
zmienna dyskretna
tabele krzyżowe (tabele kontyngencji)

Pytania i ćwiczenia

1. Odwiedź stronę Biura Spisowego (www.census.gov) i znajdź przykład analizy jednozmiennowej.
2. Znajdź przykład analizy jednozmiennowej w gazecie lub czasopiśmie.
3. Stwórz i zinterpretuj tabelę kontyngencji na podstawie następujących informacji: 150 demokratów popiera podniesienie płacy minimalnej, a 50 sprzeciwia się temu; 100 republikanów popiera podniesienie płacy minimalnej, a 300 jest temu przeciwnych.
4. Używając hipotetycznych danych z poniższej tabeli, stwórz i zinterpretuj tabele ilustrujące następujące zależności:
   a. Dwuzmiennowa zależność między wiekiem a postawą wobec aborcji.
   b. Dwuzmiennowa zależność między orientacją polityczną a postawą wobec aborcji.
   c. Wielozmiennowa zależność między wiekiem, orientacją polityczną i postawą wobec aborcji.

Wiek	Orientacja polityczna	Postawa wobec aborcji	Częstość
Młody	liberalny	zwolennik	90
Młody	liberalny	przeciwnik	10
Młody	konserwatywny	zwolennik	60
Młody	konserwatywny	przeciwnik	40
Stary	liberalny	zwolennik	60
Stary	liberalny	przeciwnik	40
Stary	konserwatywny	zwolennik	20
Stary	konserwatywny	przeciwnik	80

5. Znajdź w sieci lub w bibliotece książkę lub artykuł z badań zawierający procentową tabelę dwuzmiennową. Skopiuj tabelę i wyjaśnij, czego się z niej można dowiedzieć.

Dodatkowe lektury

Babbie Earl, Halley Fred, Zaino Jeanne, *Adventures in Social Research*, Pine Forge Press, Newbury Park, Cal. 2000. Książka wprowadza w analizę danych z badań społecznych za pomocą SPSS dla Windows. Zawiera omówienie i przykłady kilku podstawowych technik statystycznych, używanych przez badaczy społecznych.

Davis James, *Elementary Survey Analysis*, Prentice Hall, Englewood Cliffs, N.J. 1971. Nadzwyczajnie napisane i świetnie przemyślane wprowadzenie do analizy danych. Jako dodatek do omówienia partii materiału zaprezentowanej w rozdziale 15, książka Davisa warta jest uwagi w zakresie pomiaru, statystyki oraz modelu analizy rozbudowanej.

Lewis-Beck Michael, *Data Analysis: An Introduction*, t. 103, seria Quantitative Application in the Social Sciences, Sage, Thousand Oaks, Cal. 1995. Bardzo popularna krótka książka, która przybliża nowicjuszowi język statystyki. Jasność wyjaśnień i szerokie wykorzystanie przykładów powinny ci przypaść do gustu.

Newton Rae R., Rudestam Kjell Erik, *Your Statistical Consultant: Answers to Your Data Analysis Questions*, Sage, Thousands Oaks, Cal.

1999. Znakomity, łatwy w użyciu podręcznik, który odpowie na rozmaite pytania, które przyszły ci albo przyjdą do głowy w trakcie analizy danych ilościowych.

Ziesel Hans, *Say It with Figures*, Harper & Row, New York 1957. Świetne omówienie tworzenia tabel i innych podstaw analizy. Mimo swego wieku książka ta stanowi być może najlepsze spośród dostępnych omówień tego właśnie tematu. Wyróżnia się zrozumiałością i przystępnym językiem, zawiera wiele konkretnych przykładów.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 16

Model analizy rozbudowanej

Przegląd holograficzny

Model analizy rozbudowanej jest przykładem zastosowania podstawowych zasad analizy wielozmiennowej i przyczynowej. Zapoznanie się z zastosowaniem tych zasad w formie prostych tabel procentowych to pierwszy krok na drodze do zrozumienia bardziej złożonych metod analizy.

Wprowadzenie

Początki modelu analizy rozbudowanej

Paradygmat analizy rozbudowanej
Replikacja
Eksplanacja
Interpretacja
Specyfikacja
Doskonalenie paradygmatu

Analiza rozbudowana i stawianie hipotez *post factum*

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Poniższy rozdział poświęcony jest pewnej perspektywie analitycznej w naukach społecznych, zwanej niejednolicie **modelem analizy rozbudowanej** (*elaboration model*), metodą interpretatywną, metodą szkoły Columbia lub metodą Lazarsfelda. Wielość owych określeń stanowi odbicie faktu, że metoda ta ma na celu *rozbudowanie* naszej wiedzy o empirycznych zależnościach pomiędzy zmiennymi w celu dokonania *interpretacji* tych zależności metodą opracowaną przez Paula Lazarsfelda w czasie, gdy przebywał on na Uniwersytecie Columbia. Model analizy rozbudowanej jest więc jedną z metod prowadzenia analizy wielozmiennowej.

Badacze stosują model analizy rozbudowanej, aby zrozumieć zależności między dwiema zmiennymi, dzięki jednoczesnemu wprowadzeniu dodatkowych zmiennych. Model ten został co prawda stworzony początkowo tylko na użytek analizy danych z tabel krzyżowych, można go jednak z powodzeniem wykorzystywać także w połączeniu z innymi technikami statystycznymi, jak się okaże w dalszych rozdziałach tej książki.

Wierzę, że model analizy rozbudowanej umożliwia stworzenie najbardziej klarownego z możliwych obrazów logiki analizy przyczynowej w badaniach społecznych. Metoda ta stanowi obraz logicznego procesu analizy naukowej, zwłaszcza dzięki wykorzystaniu tabel krzyżowych. Co więcej, pełne zrozumienie zastosowania modelu analizy rozbudowanej do danych z tabel krzyżowych pozwala łatwiej stosować i lepiej rozumieć bardziej wyrafinowane techniki statystyczne. W ramce „Po co nam analiza rozbudowana?", autorstwa jednej z twórczyń tego modelu, Patricii Kendall, przedstawione zostały kolejne przekonujące argumenty.

Początki modelu analizy rozbudowanej

Historyczne początki modelu analizy rozbudowanej stanowią dobry przykład tego, jak wygląda praktyka badań naukowych. Jak już wspomniałem w rozdziale 1, podczas II wojny światowej Samuel Stouffer zorganizował w ramach armii amerykańskiej specjalną grupę badawczą, nad którą objął osobiście kierownictwo. W czasie wojny grupa owa prowadziła wiele rozmaitych sondaży wśród żołnierzy amerykańskich. Mimo że cele owych badań nieco się różniły, wszystkie one koncentrowały się wokół problemu czynników wpływających na sprawność bojową żołnierzy.

Kilka z tych badań dotyczyło nastrojów w armii. Badacze uważali, że skoro nastroje wpływają na sprawność bojową żołnierzy, poprawa nastrojów powinna wpłynąć na poprawę skuteczności armii na polu walki. Stouffer i jego badacze usiłowali odkryć zmienne, wpływające na morale. Między innymi starano się uzyskać empiryczne potwierdzenie kilku powszechnie akceptowanych poglądów, jak na przykład:

1. Awanse z pewnością wpływają na nastroje wśród żołnierzy, więc nastroje wśród żołnierzy jednostek o niskim wskaźniku awansów powinny być relatywnie gorsze.
2. Biorąc pod uwagę segregację i dyskryminację rasową na południu Stanów Zjednoczonych, nastroje wśród żołnierzy afroamerykańskich szkolonych w obozach na północy kraju powinny być lepsze niż wśród tych szkolonych na południu.
3. Jest bardziej prawdopodobne, że wykształceni żołnierze będą odczuwali większą niechęć do służby w armii, do której wcieleni zostali z poboru, niż ich mniej wykształceni koledzy.

Każdy z tych poglądów jest logiczny. Zdrowy rozsądek każe uważać je za prawdziwe. Stouffer postanowił sprawdzić każdy z nich empirycznie. Ku jego zaskoczeniu żaden nie znalazł potwierdzenia.

Pierwszy pogląd omawiałem w rozdziale 1. Jak być może pamiętasz, Stouffer odkrył, że żołnierze służący w Żandarmerii Wojskowej (gdzie wskaźnik awansów był najniższy w całej armii) skarżyli się na system awansowania znacznie rzadziej niż żołnierze Wojskowego Korpusu Powietrznego (gdzie czas oczekiwania na awans był najkrótszy). Pozostałe dwa sądy wypadły równie kiepsko: Afroamerykanie szkole-

Po co nam analiza rozbudowana?

Patricia Kendall

Wydział Socjologii,
Queens College, Uniwersytet Miasta Nowy Jork

Prawdziwy eksperyment kontrolowany musi spełniać wiele wymagań. Najważniejsze z nich to: (a) stworzenie grup eksperymentalnych i kontrolnych, *identycznych* w granicach prawdopodobieństwa (osiągamy to, przypisując losowo jednostki do poszczególnych grup: za pomocą tablic losowych, rzucając monetą itp.), (b) pewność, że to *eksperymentator*, nie zaś czynniki zewnętrzne, uruchamia bodziec eksperymentalny i (c) oczekiwanie, czy bodziec wywołał zakładany skutek.

Możemy na przykład postawić hipotezę, że studenci uczelni należących do Ivy League osiągają większe sukcesy zawodowe niż studenci innych uczelni. W jaki sposób moglibyśmy przetestować tę hipotezę w drodze eksperymentu? Załóżmy, że odpowiedziałeś: „Bierzemy grupę ludzi około czterdziestki, sprawdzamy, którzy z nich studiowali w uczelniach z Ivy League, sprawdzamy, czy odnieśli większy sukces niż ci, którzy kończyli inne uczelnie". Jeśli tak brzmiała twoja odpowiedź, jesteś w błędzie.

Prawdziwy eksperyment wymagałby w danym przypadku, aby badacz wybrał kilkanaście starszych klas w szkołach średnich, podzielił każdą z tych klas losowo na grupy eksperymentalne i kontrolne, i wysłał grupy eksperymentalne do szkół z Ivy League (nie biorąc w ogóle pod uwagę ani sytuacji finansowej i możliwości edukacyjnych uczestników, ani też gotowości do ich przyjęcia przez daną uczelnię). Grupy kontrolne studiowałyby na innych uczelniach. Eksperymentator musiałby poczekać jakieś 20 lat, póki członkowie obu grup nie osiągną pełni zawodowych możliwości, po czym mógłby zmierzyć względny sukces każdej z grup. Procedura ogólnie dość dziwaczna.

Socjologowie badają też często hipotezę, że pochodzenie z rozbitej rodziny sprzyja skłonnościom przestępczym młodocianych. Jak zabrałbyś się do eksperymentalnego testowania tej hipotezy? Gdybyś postępował zgodnie z powyższym przykładem, musiałbyś stwierdzić, że sprawdzenie tej hipotezy w drodze prawdziwego eksperymentu jest całkowicie niemożliwe. Pomyśl tylko, co musiałby zrobić eksperymentator!

Wymogi prawdziwego eksperymentu w badaniach socjologicznych są tak nierealistyczne, że zmuszeni jesteśmy używać innych, mniej idealnych metod we wszystkich właściwie sytuacjach – może z wyjątkiem najbardziej banalnych. Możemy badać, czy studenci uczą się więcej podczas wykładów tego czy innego typu, albo też, czy obejrzenie filmu zmienia poglądy widza. Nie zawsze jednak jesteśmy naprawdę zainteresowani tego typu pytaniami.

Uciekamy się zatem do przybliżeń – przede wszystkim do sondaży – które także mają wady. Model analizy rozbudowanej pozwala nam jednak badać dane sondażowe, zdać sobie sprawę z ich możliwych wad i wyciągnąć dość skomplikowane wnioski w wielu istotnych kwestiach.

ni w obozach na północy i na południu nie różnili się właściwie pod względem nastrojów, mniej zaś wykształceni żołnierze z poboru mieli bardziej niechętny stosunek do służby w armii niż ich wykształceni koledzy.

Stouffer, zamiast schować wyniki do szuflady albo też po prostu opublikować je wraz ze wskaźnikami ich istotności statystycznej, zadał sobie pytanie „dlaczego?". Odpowiedzi dostarczyła mu teoria grup odniesienia i deprywacji względnej. Najprościej mówiąc, Stouffer uznał, że żołnierze nie oceniali swojej sytuacji życiowej wedle obiektywnych i absolutnych standardów. Posługiwali się raczej kryterium swego względnego położenia w odniesieniu do ludzi wokół siebie. Osoby, z którymi się porównywali, wchodziły w skład ich grupy odniesienia. Deprywacja względna jest efektem niekorzystnego wyniku tak przeprowadzonego porównania.

Zgodnie z logiką powyższego rozumowania Stouffer wyjaśnił wszystkie anomalie pojawiające się w danych empirycznych. Jeśli chodzi o awans, uznał, że żołnierze oceniali sprawiedliwość systemu awansowania, porównując własne

doświadczenia z doświadczeniami innych ludzi w swoim otoczeniu. W żandarmerii, gdzie awanse były rzadkie, niewielu żołnierzy miało szansę znać kogoś z gorszymi kwalifikacjami, który zostałby szybciej awansowany. W siłach powietrznych natomiast wysoki wskaźnik awansów oznaczał, że wielu żołnierzy mogło znać ludzi gorzej wykwalifikowanych, którzy zostali awansowani szybciej, niż na to zasługiwali. Tak więc, jak na ironię, żandarmi twierdzili, że system awansowania jest w zasadzie sprawiedliwy, lotnicy zaś byli przeciwnego zdania.

Podobna analiza posłużyła wyjaśnieniu przypadku Afroamerykanów. Żołnierze afroamerykańscy nie porównywali warunków na Północy i na Południu kraju, lecz swoją własną pozycję życiową z sytuacją innych obywateli pochodzenia afroamerykańskiego w swoim otoczeniu. Na Południu, gdzie dyskryminacja była najbardziej dotkliwa, żołnierze ci odczuwali, iż fakt bycia żołnierzem izolował ich w pewnym stopniu od wrogości otaczających ich norm kulturowych. Afroamerykańscy żołnierze cieszyli się nieco wyższym statusem niż cywile, których dyskryminowano na każdym kroku, odmawiając im godności własnej, lepiej płatnych stanowisk pracy itp. Na Północy natomiast wielu napotkanych cywili – Afroamerykanów miało dobrze płatną pracę w przemyśle zbrojeniowym. Przy niższym poziomie dyskryminacji fakt bycia żołnierzem nie podnosił statusu jednostki w zbiorowości.

Koncepcje grupy odniesienia i deprywacji względnej zapewniały, jak się zdawało, także wyjaśnienie paradoksalnego faktu, iż żołnierze z wyższym wykształceniem akceptują fakt powołania do służby łatwiej niż żołnierze mniej wykształceni. Stouffer rozumował tu w sposób następujący:

„1. Przyjaciel danej osoby ma na ogół podobne do niej wykształcenie.
2. Gorzej wykształceni mężczyźni w wieku poborowym częściej niż lepiej wykształceni podejmują niewymagającą kwalifikacji pracę w przemyśle i rolnictwie.
3. Podczas wojny przemysł i rolnictwo to branże kluczowe dla interesu narodowego. Robotnicy zatrudnieni w tych sektorach są zwolnieni od poboru.
4. Jest bardziej prawdopodobne, że człowiek gorzej wykształcony ma przyjaciół w zawodach zwalniających od poboru niż człowiek z wyższym wykształceniem.
5. Porównując się ze swoimi przyjaciółmi, gorzej wykształcony poborowy ma większe szanse poczuć się dyskryminowanym, niż poborowy o wyższym poziomie wykształcenia".

(1949–1950, s. 122–127)

Wyjaśnienie Stouffera rozwiązywało zagadkę trzech dziwacznych wyników. Brakowało mu jednak danych empirycznych dla ich sprawdzenia, ponieważ nie wchodziły one w skład ustalonego z góry projektu badawczego. Wywody logiczne Stouffera zapewniły jednak podstawy dla późniejszego rozwoju modelu analizy rozbudowanej: wyjaśnienie zależności między dwiema zmiennymi dzięki kontrolowanemu wprowadzeniu innych zmiennych.

Paul Lazarsfeld i jego współpracownicy na Uniwersytecie Columbia opracowali formalnie model analizy rozbudowanej w 1946 roku. Komentując metodologię badań armii Stouffera, Lazarsfeld i Patricia Kendall wykorzystali logikę analizy rozbudowanej, by stworzyć hipotetyczne tabele, które dowodziłyby twierdzeń Stouffera co do wykształcenia i stosunku poborowego względem armii, gdyby dane empiryczne były dostępne (Kendall, Lazarsfeld 1950).

Centralnym elementem logiki modelu analizy rozbudowanej jest przyjęcie za punkt wyjścia zaobserwowanej zależności między zmiennymi oraz możliwości, że jedna ze zmiennych może być przyczyną drugiej. W przykładzie Stouffera dwie początkowe zmienne to *poziom wykształcenia* i *akceptacja faktu powołania do służby wojskowej*. Jako że wykształcenie żołnierzy było ustalone już przed poborem (czyli przed wyrobieniem zdania na temat poboru), należałoby uznać, że to *wykształcenie* było przyczyną, czyli zmienną niezależną, *akceptacja poboru* zaś skutkiem, zatem zmienną zależną. Jak już jednak widzieliśmy, obserwowane zależności przeczyły oczekiwaniom badacza.

Analiza rozbudowana bada wpływ innych zmiennych na pierwotnie zaobserwowane zależności. Analiza ta odsłania niekiedy mecha-

nizmy, poprzez które zachodzi związek przyczynowy. Analiza rozbudowana może też niekiedy dowieść braku jakiegokolwiek związku przyczynowego.

W naszym przykładzie zmienną dodatkową był fakt, że koledzy żołnierza zostali również wcieleni do armii – lub też nie. W spekulatywnym wyjaśnieniu Stouffera zmienna ta pozwalała uznać za logiczne, że lepiej wykształceni żołnierze w wyższym stopniu akceptowali fakt powołania do armii: prawdopodobne było, że ich koledzy zostaną również powołani. Żołnierze z niższym wykształceniem częściej pracowali w zawodach, których wykonywanie powodowało w wielu przypadkach odroczenie poboru, co wytwarzało u poborowych rekrutujących się z tej grupy poczucie krzywdy.

Kendall i Lazarsfeld zaczęli od danych Stouffera na temat pozytywnego związku wykształcenia z akceptacją poboru (patrz tabela 16.1). W tej i kolejnych tabelach stwierdzenia „powinienem był zostać zwolniony" i „nie powinienem był zostać zwolniony" odpowiadają poglądom poborowych na ich własną sytuację, przy czym ta druga grupa uznaje powołanie ich do służby za słuszne.

Kendall i Lazarsfeld stworzyli następnie kilka hipotetycznych tabel, które ilustrują, jak wyglądałaby analiza, gdyby żołnierzom zadano pytanie, czy większość ich przyjaciół została objęta poborem, czy też z niego zwolniona. W tabeli 16.2 19% lepiej wykształconych żołnierzy stwierdza, że ich przyjaciele zostali zwolnieni, w porównaniu z aż 79% twierdzących tak żołnierzy z niższym wykształceniem.

Zauważ, że liczba żołnierzy z wyższym i niższym poziomem wykształcenia jest taka sama, jak w prawdziwych danych Stouffera. W kolejnych tabelach podane są również zgodne z prawdą odsetki tych, którzy uważali objęcie ich poborem za słuszne lub niesłuszne. Wymyślone są tylko dane co do liczby żołnierzy zeznających, że ich przyjaciele zostali powołani lub zwolnieni.

Wyjaśnienie Stouffera zakładało następnie, że żołnierze, których przyjaciele zostali zwolnieni, będą bardziej niechętnie usposobieni wobec faktu powołania ich samych, niż ci, którzy nie mieli zwolnionych od poboru przyjaciół.

Tabela 16.1

Podsumowanie danych Stouffera dotyczących wykształcenia i akceptacji faktu powołania

	Poziom wykształcenia	
	wysoki	niski
Nie powinienem był zostać zwolniony	88%	70%
Powinienem był zostać zwolniony	12%	30%
	100%	100%
	(1761)	(1876)

Źródło: Tabele 16.1, 16.2, 16.3 i 16.4 zostały zmienione za zezwoleniem Macmillan Publishing Co., Inc., na podstawie: *Continuities in Social Research: Studies in the Scope and Method of „The American Soldier"*, Robert K. Merton, Paul F. Lazarsfeld (red.), Copyright 1950, The Free Press, Corporation, wznowienie: Robert Merton 1978.

Tabela 16.2

Hipotetyczne zależności między wykształceniem a zwolnieniem przyjaciół od poboru

	Poziom wykształcenia	
Czy twoi przyjaciele zostali zwolnieni?	wysoki	niski
Tak	19%	79%
Nie	81%	21%
	100%	100%
	(1761)	(1876)

Tabela 16.3

Hipotetyczne zależności między faktem posiadania zwolnionych od poboru przyjaciół a akceptacją własnego powołania do armii

	Czy twoi przyjaciele zostali zwolnieni?	
	tak	nie
Nie powinienem był zostać zwolniony	63%	94%
Powinienem był zostać zwolniony	37%	6%
	100%	100%
	(1819)	(1818)

Tabela 16.3 przedstawia hipotetyczne dane wspierające to założenie.

Hipotetyczne dane w tabelach 16.2 i 16.3 potwierdzałyby związki, które Stouffer uwzględnił w swoim wyjaśnieniu. Po pierwsze, gorzej wykształceni żołnierze częściej niż żołnierze lepiej wykształceni miewaliby przyjaciół zwolnionych od poboru. Po drugie, posiadanie

Tabela 16.4

Hipotetyczne dane wiążące wykształcenie z akceptacją faktu poboru poprzez czynnik posiadania zwolnionych od poboru

	Zwolnieni przyjaciele		Brak zwolnionych przyjaciół	
	wysoki poziom wykształcenia	niski poziom wykształcenia	wysoki poziom wykształcenia	niski poziom wykształcenia
Nie powinienem był zostać zwolniony	63%	63%	94%	95%
Powinienem był zostać zwolniony	37%	37%	6%	5%
	100%	100%	100%	100%
100% =	(335)	(1484)	(1426)	(392)

zwolnionych od poboru przyjaciół sprawiałoby, że prawdopodobieństwo, iż żołnierz sądzi, że powinien zostać zwolniony, byłoby wyższe. Stouffer sugerował, że te dwie zależności wyjaśniłyby pierwotną zależność między wykształceniem a stosunkiem do poboru. Kendall i Lazarsfeld stworzyli hipotetyczną tabelę, która stanowiłaby potwierdzenie wyjaśnienia Stouffera (tabela 16.4).

Przypomnijmy, że wyniki początkowych badań dowodziły, że lepsze wykształcenie częściej niż gorsze szło w parze z pogodzeniem się z faktem powołania do armii. W tabeli 16.4 zauważamy jednak, że poziom wykształcenia nie ma wpływu na akceptację powołania wśród tych, którzy twierdzą, że ich przyjaciele zostali zwolnieni: 63% badanych w obu kategoriach wykształcenia twierdzi, że akceptuje fakt powołania do służby (czyli uznają oni, że nie powinni byli zostać zwolnieni). Podobnie, poziom wykształcenia nie ma widocznego wpływu na akceptację powołania wśród osób, które podały, że nie mają zwolnionych od poboru przyjaciół: 94% i 95% twierdzi, że nie powinni byli zostać zwolnieni.

Z kolei wśród żołnierzy o wyższym poziomie wykształcenia akceptacja poboru jest silnie powiązana z faktem posiadania lub nieposiadania zwolnionych od służby przyjaciół: odpowiednio 63% i 94%. Prawidłowość ta zachodzi również wśród gorzej wykształconych żołnierzy. Hipotetyczne dane w tabeli 16.4 wspierają zatem tezę Stouffera, iż wykształcenie wpływa na akceptację powołania do służby jedynie za pośrednictwem faktu posiadania zwolnionych od poboru przyjaciół. Wykształceni poborowi rzadziej miewali zwolnionych od poboru przyjaciół i z tego właśnie powodu byli bardziej skłonni zaakceptować fakt, że również oni sami zostali objęci poborem. Żołnierze o niższym poziomie wykształcenia częściej miewali zwolnionych od służby przyjaciół i z tego względu częściej uważali objęcie ich samych poborem za niesłuszne.

Zauważ, że ani wyjaśnienie Stouffera, ani też hipotetyczne dane nie przeczyły prawdziwości pierwotnej zależności. W miarę wzrostu poziomu wykształcenia rosła także akceptacja powołania do służby w armii. Naturę tej empirycznej zależności zinterpretowano jednak dzięki wprowadzeniu trzeciej zmiennej. Zmienna ta, *posiadanie zwolnionych od służby wojskowej przyjaciół*, nie niweczy pierwotnej zależności, wyjaśnia jedynie mechanizm, który prowadzi do jej zaistnienia.

Tak zatem wygląda sedno modelu analizy rozbudowanej i analizy wielozmiennowej. Po zaobserwowaniu empirycznej zależności między dwiema zmiennymi (na przykład *poziomem wykształcenia* a *akceptacją faktu powołania do służby wojskowej*), poszukujemy wyjaśnienia natury tej zależności za pomocą wpływu wywieranego przez wprowadzenie innych zmiennych (takich jak *posiadanie zwolnionych od służby wojskowej przyjaciół*). Technicznie rzecz biorąc, dokonujemy tego dzieląc naszą próbę na podzbiory wedle wartości **zmiennej testowej**, zwanej także **zmienną kontrolną**. W naszym przykładzie posiadanie zwolnionych od służby przyjaciół to zmienna kontrolna, próbę zaś podzielono na tych, którzy mają zwolnionych przyjaciół, i tych, których przyjaciele nie byli zwolnieni od służby w wojsku. Zależność między dwiema zmiennymi wyjściowymi (*poziomem wykształcenia* a *akceptacją faktu powołania do służby wojskowej*) przeliczamy następnie dla każdej podpróby osobno. Tabele, które powstają w ten sposób, nazywamy

tabelami cząstkowymi, a zależności w nich zilustrowane to **zależności cząstkowe**. Zależności cząstkowe są następnie porównywane z **zależnością początkową** wykrytą w całej próbie, którą nazywa się często zależnością zerową, dla zaznaczenia, że nie była ona kontrolowana ze względu na żadne zmienne testowe.

Mimo że model analizy rozbudowanej został pierwszy raz zastosowany do danych hipotetycznych, stanowił on początek logicznej metody analizy zależności między faktycznie zmierzonymi zmiennymi. Jak się wkrótce przekonamy, nasz pierwszy, hipotetyczny przykład opisuje tylko jeden z możliwych wyników zastosowania modelu analizy rozbudowanej – są bowiem i inne.

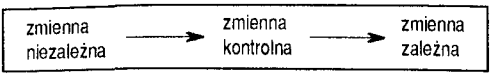
Paradygmat analizy rozbudowanej

W poniższym podrozdziale przedstawiam wskazówki, ułatwiające zrozumienie analizy rozbudowanej. Na początku musimy się dowiedzieć, czy zmienna kontrolna jest *uprzednia* (wcześniejsza w czasie) w stosunku do pozostałych dwóch zmiennych, czy też *pośrednicząca* między nimi, ponieważ te dwie możliwości wskazujące na różne zależności logiczne w modelu wielozmiennowym. Jeśli zmienna kontrolna jest pośrednicząca, jak w przypadku wykształcenia, zwolnienia przyjaciół od służby wojskowej i akceptacji powołania do armii, analiza oparta jest na modelu przedstawionym na rysunku 16.1. Logika tej zależności wielozmiennowej jest następująca: zmienna niezależna (*poziom wykształcenia*) wpływa na zmienną kontrolną (*zwolnienie przyjaciół*), ta zaś z kolei wpływa na zmienną zależną (*akceptacja faktu powołania do służby wojskowej*).

Jeśli zmienna kontrolna jest uprzednia w stosunku do obu zmiennych, zależnej i niezależnej, trzeba zastosować inny model (patrz rysunek 16.2). Zmienna kontrolna wpływa w tym przypadku zarówno na zmienną „zależną", jak i „niezależną". Z pewnością zauważysz, że na rysunku terminy *zmienna zależna* i *zmienna niezależna* zostały, ściśle rzecz biorąc, zastosowane niepoprawnie. W istocie mamy bowiem tylko

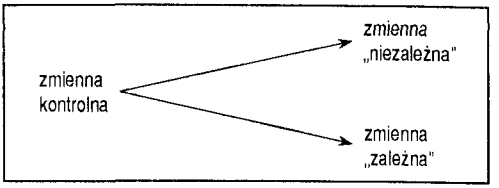
Rysunek 16.1

Zmienna kontrolna pośrednicząca



Rysunek 16.2

Zmienna kontrolna uprzednia



jedną zmienną niezależną (zmienną kontrolną) i dwie zmienne zależne. Niepoprawna terminologia służy jednak zachowaniu ciągłości przykładu. Ze względu na indywidualne zależności między zmiennymi „zależną" i „niezależną" a zmienną kontrolną, zmienne „zależna" i „niezależna" pozostają w zależności empirycznej, przy czym nie występuje między nimi związek przyczynowy. Zależność empiryczna między nimi jest więc tylko skutkiem ich przypadkowego powiązania ze zmienną kontrolną (kolejne przykłady pomogą wyjaśnić tę zależność).

Tabela 16.5 stanowi wprowadzenie do analizy rozbudowanej. Dwie kolumny tabeli wskazują, czy zmienna kontrolna jest uprzednia, czy pośrednicząca, w sensie przedstawionym powyżej. Lewa strona tabeli określa naturę zależności cząstkowych w porównaniu z pierwotną zależnością między zmienną zależną a niezależną. Wewnątrz tabeli podane są techniczne

Tabela 16.5

Paradygmat analizy rozbudowanej

Zależności cząstkowe w porównaniu z pierwotną	Zmienna kontrolna	
	uprzednia	pośrednicząca
Ta sama zależność	replikacja	replikacja
Zależność słabsza lub brak zależności	eksplanacja	interpretacja
Wynik mieszany*	specyfikacja	specyfikacja

* Jedna z zależności cząstkowych jest taka sama lub silniejsza, druga zaś jest słabsza lub nie występuje.

określenia – replikacja, eksplanacja, interpretacja i specyfikacja – przyporządkowane każdej poszczególnej sytuacji. Będziemy je omawiać po kolei.

Replikacja

Kiedy zależności cząstkowe są zasadniczo takie same, jak zależność początkowa, wynik ten określany jest terminem **replikacja**, niezależnie od tego, czy zmienna kontrolna jest uprzednia, czy pośrednicząca. Oznacza to, że zależność początkowa uległa replikacji, czyli powtórzeniu w warunkach kontrolnych. Gdyby w naszym poprzednim przykładzie okazało się, że wykształcenie nadal wpływa na akceptację faktu powołania do służby wojskowej zarówno wśród tych, których przyjaciele zostali zwolnieni od poboru, jak i wśród tych, których przyjaciele nie zostali zwolnieni, stwierdzilibyśmy, że otrzymaliśmy replikację zależności wyjściowej. Zauważ jednak, że odkrycie to nie potwierdzałoby wyjaśnienia pierwotnej zależności autorstwa Stouffera. Fakt posiadania lub nieposiadania zwolnionych przyjaciół *nie* stanowiłby wówczas mechanizmu, za pomocą którego wykształcenie wpływa na akceptację powołania do służby wojskowej.

Aby zobaczyć, jak wygląda replikacja, spójrz jeszcze raz na tabele 16.3 i 16.4. Wyobraź sobie, że naszym pierwszym odkryciem było, że posiadanie zwolnionych od służby przyjaciół silnie wpływa na stosunek żołnierzy do własnego powołania do armii, jak to uwidoczniono w tabeli 16.3. Gdybyśmy odkryli tę zależność jako pierwszą, być może postanowilibyśmy sprawdzić, czy była ona równie prawdziwa dla żołnierzy z różnym zapleczem edukacyjnym. By się tego dowiedzieć, wprowadzilibyśmy *wykształcenie* jako zmienną kontrolną.

Tabela 16.4 przedstawia wyniki tego typu badania, mimo że została skonstruowana nieco inaczej, niż gdyby powstała w wyniku potraktowania wykształcenia jako zmiennej kontrolnej. Tak czy inaczej, widzimy w tabeli, że posiadanie bądź nieposiadanie zwolnionych od służby przyjaciół nadal wpływa na postawy wobec poboru wśród żołnierzy z niskim i wysokim poziomem wykształcenia (porównaj kolumny 1 i 3, a następnie 2 i 4). Wynik ten oznacza replikację zależności między posiadaniem zwolnionych od służby wojskowej przyjaciół a postawą wobec własnego powołania.

Badacze stosują często model analizy rozbudowanej w sposób dość rutynowy, mając nadzieję na replikację swoich wyników w podzbiorach próby. Gdybyśmy na przykład odkryli zależność między wykształceniem a uprzedzeniami, moglibyśmy wprowadzić takie zmienne kontrolne jak wiek, region, rasa, wyznanie itp., by zbadać stabilność zależności początkowej. Gdyby zależność replikowała się wśród starych i młodych, wśród ludzi pochodzących z różnych stron itd., wówczas mielibyśmy podstawy, by wnioskować, że zależność początkowa jest prawdziwa i ogólnie ważna.

Eksplanacja

Eksplanacja to termin stosowany w celu opisania *zależności pozornej*: dzięki wprowadzeniu zmiennej kontrolnej dowodzimy, że zależność początkowa była fałszywa. Wymaga to spełnienia dwóch warunków: (1) zmienna kontrolna musi być uprzednia zarówno wobec zmiennej zależnej, jak i niezależnej; (2) zależności cząstkowe muszą być zerowe lub istotnie słabsze niż zależności początkowe. Sytuację tę zilustruję kilkoma przykładami.

Istnieje zależność empiryczna między liczbą bocianów a wskaźnikiem urodzeń w danej miejscowości. Im więcej bocianów, tym wyższy wskaźnik urodzeń. Ta zależność empiryczna mogłaby doprowadzić do przypuszczenia, że liczba bocianów wpływa na wskaźnik urodzeń. Kontrola za pomocą zmiennej uprzedniej pozwala jednak na eksplanację tej zależności. Obszary wiejskie mają zarazem więcej bocianów i wyższe wskaźniki urodzeń niż obszary miejskie. W ramach obszarów wiejskich nie występuje jednak zależność między liczbą bocianów a wskaźnikiem urodzeń, nie pojawia się też ona w ramach obszarów miejskich.

Rysunek 16.3 ilustruje proces wywoływania przez zmienną wieś/miasto pozornej zależności między bocianami a urodzeniami. Część I obrazuje zależność początkową. Zauważ, że z jednym wyjątkiem wszystkie miejscowości w ru-

Rysunek 16.3

Z życia wzięte fakty na temat bocianów i dzieci

I. Wskaźnik urodzeń w miejscowościach z dużą lub małą liczbą bocianów

W – wysoki wskaźnik urodzeń
N – niski wskaźnik urodzeń

LICZBA BOCIANÓW

niewiele	wiele
N N NNN N NNN NW N N N NNNN	W W W N W W W W W W W W

II. Zmienna kontrolna: obszar miejski (miasto) lub wiejski (wieś)

LICZBA BOCIANÓW

	niewiele	wiele
miasto	W	W W W W W W W W W W W
wieś	N N N N NN N N N N N N N N N N	N

bryce „wiele bocianów" mają wysokie wskaźniki urodzeń, jak również wszystkie – poza jednym przypadkiem – miejscowości w rubryce „niewiele bocianów" mają niski wskaźnik urodzeń. Ujmując tę zależność w procentach, można powiedzieć, że 93% miejscowości o dużej liczbie bocianów miało zarazem wysoki wskaźnik urodzeń, przy zaledwie 7% takich przypadków w miejscowościach o niewielkiej liczbie bocianów. Ta dość znacząca różnica jest odbiciem silnego związku między obiema zmiennymi.

W części II rysunku oddzielono wsie od miast (obszary miejskie od wiejskich) i zbadano zależność między bocianami a dziećmi w każdej z tych grup oddzielnie. Widać teraz, że wszystkie obszary wiejskie mają wysoki wskaźnik urodzeń, na wszystkich obszarach miejskich zaś wskaźniki te są niskie. Zauważ też, że tylko jedna wieś ma mało bocianów, tylko jedno miasto zaś ma ich wiele.

Oto inny podobny przykład. Istnieje prosta zależność między liczbą wozów strażackich przyjeżdżających do pożaru a poziomem szkód przez ten pożar wyrządzonych. Im więcej przybywa wozów strażackich, tym większe są wyrządzane szkody. Można by na tej podstawie przypuszczać, że wozy strażackie powodują szkody. Jednak wprowadzenie uprzedniej zmiennej kontrolnej, rozmiaru pożaru, eksplanuje (wyjaśnia) zależność początkową. Wielkie pożary powodują większe szkody niż małe i przyjeżdża do nich więcej wozów strażackich. Gdy bierzemy pod uwagę tylko wielkie pożary, zależność znika (czy wręcz ulega odwróceniu), podobnie w przypadku małych pożarów.

Weźmy na koniec przykład z prawdziwych badań. Wiele lat temu odkryłem zależność empiryczną pomiędzy regionem kraju, w którym kadra naukowa uczelni medycznych zdobywała wykształcenie, a stosunkiem członków tejże do państwowej służby zdrowia (Babbie 1970). Dla uproszczenia, zbadajmy tylko Wschód i Południe kraju. 78% spośród członków kadry uczelni medycznych wykształconych na wschodnich

Uczęszczanie do uczelni należącej do Ivy League a późniejszy sukces zawodowy

Patricia Kendall

Wydział Socjologii,
Queens College, Uniwersytet Miasta Nowy Jork

Być może największym zagrożeniem dla badacza, analizującego dane sondażowe, jest sytuacja, kiedy zależność, co do której miał nadzieję, iż jest przyczynowa, okazuje się pozorna, czyli pierwotna zależność między *X* a *Y* da się wyjaśnić oddziaływaniem uprzedniego czynnika kontrolnego. Dokładniej mówiąc, zależności cząstkowe między *X* a *Y* zostają zredukowane do zera przy stałym poziomie uprzedniego czynnika kontrolnego.

Możliwość ta okazała się decydująca dla głównego wyniku badań przeprowadzonych kilkadziesiąt lat temu. Jedna z moich koleżanek, doktorantek na Uniwersytecie Columbia, Patricia Salter West, oparła się w swojej pracy na kwestionariuszach, otrzymanych przez „Time Magazine" od 10 000 jego prenumeratorów – mężczyzn. Jedna z wielu hipotez West była następująca: mężczyźni, wykształceni na uczelniach wchodzących w skład Ivy League (Brown, Columbia, Cornell, Dartmouth, Harvard, Uniwersytet Pennsylvania, Princeton i Yale), odnoszą w późniejszym życiu zawodowym większe sukcesy mierzone wielkością rocznych zarobków, niż ci, którzy ukończyli inne uczelnie.

Pierwsza tabela czteropolowa przemawiała za rozumowaniem West. Podane cyfry wymyśliłam co prawda sama, ale odpowiadają one dokładnie wynikom badań West. Wydaje się, że fakt uczęszczania do uczelni należącej do Ivy League stanowi drogę do wyraźnie większych sukcesów zawodowych niż ukończenie jakiejkolwiek innej uczelni.

Momencik, momencik. Czy to nie jest przypadkiem typowa zależność, która może być pozorna? Kogo stać na posłanie syna do szkół, wchodzących w skład Ivy League? Oczywiście, że bogatych*. A kto może zapewnić kontakty zawodowe, które pozwolą synowi osiągnąć sukces i zrobić karierę? Również rodziny bogate i zamożne.

Inaczej mówiąc, status społeczno-ekonomiczny rodziny studenta mógłby posłużyć eksplanacji rzekomej zależności przyczynowej. Pewne obserwacje West rzeczywiście pozwalają sądzić, że tak się właśnie sprawy mają.

Tabela 1*

Późniejsze sukcesy zawodowe (*Y*)	Uczelnia, do której uczęszczał respondent: uczelnia z Ivy League	inna uczelnia
Sukces	1300 (65%)	2000 (25%)
Brak sukcesu	700 (35%)	6000 (75%)
Ogółem	2000 (100%)	8000 (100%)

* Musiałam wymyślić odpowiednie liczby, ponieważ jedyna opublikowana wersja badań West nie podawała liczb ogółem. Por. Havemann Ernest, West Particia Salter, *They Went to College*, Harcourt, Brace, New York 1952.

Tabela 2

Uczęszczanie do uczelni wchodzących w skład Ivy League w zależności od statusu społeczno-ekonomicznego rodziny (SSE)

Uczelnia, do której uczęszczał respondent (*X*)	SSE rodziny respondenta: wysoki SSE	niski SSE
Ivy League	1500 (33%)	500 (9%)
Inna uczelnia	3000 (67%)	5000 (91%)
Ogółem	4500 (100%)	5500 (100%)

uczelniach miało pozytywny stosunek do służby zdrowia, podczas gdy spośród ich kolegów wykształconych na Południu zdania tego było 59%. Odkrycie to wydawało się zrozumiałe w świetle faktu, iż na Południu programy publicznej służby zdrowia napotykały zawsze na znacznie większy opór niż na Wschodzie, okres studiów na uczelni medycznej powinien zaś co do zasady wpływać na poglądy lekarza w kwestiach medycznych. Zależność ta uległa jednak eksplanacji po wprowadzeniu zmiennej uprzedniej: regionu pochodzenia przyszłego członka kadry uczelni medycznej. Spośród wychowanych na Wschodzie 89% kształciło się również na Wschodzie, a tylko 11% na Południu. Spośród badanych wychowanych na Południu 53% ukończyło wschod-

Tabela 3

Zależności cząstkowe między *X* a *Y* przy stałej wartości *T*

	Wysoki SSE (*T*)		Niski SSE (*T*)	
Późniejsze sukcesy zawodowe (*Y*)	Uczelnia z Ivy League (*X*)	Inna uczelnia (*X*)	Uczelnia z Ivy League (*X*)	Inna uczelnia (*X*)
Sukces	1000 (67%)	1000 (33%)	300 (60%)	1000 (20%)
Brak sukcesu	500 (33%)	2000 (67%)	200 (40%)	4000 (80%)
Ogółem	1500 (100%)	3000 (100%)	500 (100%)	5000 (100%)

Zgodnie z tabelą 2, do uczelni z Ivy League uczęszczała jedna trzecia studentów z rodzin opisanych jako zamożne, w porównaniu z jedną jedenastą studentów pochodzących ze sfer gorzej sytuowanych. Tak więc istnieje bardzo wysoka korelacja między dwiema zmiennymi, *X* i *T*. Podobnie wysoka korelacja istnieje między statusem społeczno-ekonomicznym rodziny (*T*) a późniejszym sukcesem zawodowym studenta (*Y*).

Siła tych tak zwanych zależności brzegowych skłania ku przypuszczeniu, że hipoteza West co do przyczynowej natury uczęszczania do szkoły z Ivy League mogła być błędna; skłania ona za to do przypuszczenia, że wpływ na obserwowaną przez West zależność brzegową miał status społeczno-ekonomiczny rodziny studenta.

To jednak jeszcze nie koniec. Pytanie brzmi: co się stanie z zależnościami cząstkowymi po wprowadzeniu kontroli zmiennej testowej. Sytuację tę obrazuje tabela 3.

Zależności cząstkowe wskazują, że nawet przy stałej wartości statusu społeczno-ekonomicznego rodziny nadal daje się zauważyć zależność między uczęszczaniem do uczelni wchodzącej w skład Ivy League a sukcesem w późniejszym życiu zawodowym. Ostatecznie zatem początkowa hipoteza West znalazła poparcie w wynikach jej analizy.

Mimo to West nie zdołała w żadnym razie udowodnić swojej hipotezy. Niemal zawsze istnieją jakieś dodatkowe czynniki uprzednie, które mogłyby wyjaśniać zależność początkową. Zwróć na przykład uwagę na inteligencję studentów (mierzoną testami IQ lub na skali SAT). Szkoły z Ivy League szczycą się doskonałym poziomem swoich studentów. Z tego też względu są być może skłonne przyznawać stypendia naukowe nadzwyczaj zdolnym uczniom pozbawionym środków finansowych na utrzymanie i czesne. Błyskotliwi studenci, skoro już znajdą się na tych uczelniach, mogą rozwinąć swoje zdolności i nawiązać kontakty, które przyczynią się do ich późniejszego sukcesu zawodowego. West nie mogła zbadać, czy zależności cząstkowe znikną po wprowadzeniu kontroli zmiennej testowej w postaci inteligencji, ponieważ nie dysponowała danymi na temat inteligencji badanych.

W sumie trzeba stwierdzić, że paradygmat analizy rozbudowanej pozwala badaczowi wyeliminować niektóre możliwości i znaleźć dodatkowe argumenty na rzecz innych. Nie pozwala jednak czegokolwiek dowodzić.

* West, nie mając żadnych bezpośrednich danych na temat statusu społeczno-ekonomicznego rodziny, zdefiniowała jako bogate lub posiadające wysoki status społeczno-ekonomiczny te rodziny, które utrzymywały całkowicie swoich synów przez cały okres trwania studiów. Jako mniej zamożne lub mające niższy status społeczno-ekonomiczny zdefiniowała te rodziny, których synowie pracowali w czasie studiów na cały etat lub część etatu.

nie uczelnie, 47% zaś kształciło się na Południu. Co więcej, region pochodzenia miał też wpływ na postawy względem państwowej służby zdrowia. Pozytywny stosunek wobec niej zadeklarowało 84% wychowanych na Wschodzie i 49% wychowanych na Południu.

Tabela 16.6 przedstawia zależność między trzema zmiennymi: (a) regionem pochodzenia, (b) regionem pobierania wykształcenia medycznego i (c) postawą wobec państwowej służby zdrowia. Badani wychowani na Wschodzie częściej miewają pozytywny stosunek do służby zdrowia, niezależnie od regionu kraju, w którym odebrali wykształcenie medyczne. Badani wychowani na Południu rzadziej mają pozytywny stosunek do państwowej służby

Tabela 16.6

Region pochodzenia, region kształcenia i stosunek do służby zdrowia

		Odsetek nastawionych pozytywnie wobec państwowej służby zdrowia	
		region pochodzenia	
		Wschód	Południe
region wykształcenia	Wschód	84%	50%
	Południe	80%	47%

Źródło: Earl R. Babbie, *Science and Morality in Medicine*, University of California Press, Berkeley 1970, s. 181.

zdrowia, przy czym region, w którym się kształcili, nie ma właściwie wpływu na tę zależność. Dane te wskazują zatem, że zależność początkowa między regionem, w którym dany badany odebrał wykształcenie medyczne, a jego stosunkiem do służby zdrowia, była pozorna. Jej występowanie było spowodowane wyłącznie przypadkowym wpływem regionu *pochodzenia* zarówno na region kształcenia, jak i na postawę wobec państwowej służby zdrowia. Przy tym samym regionie kształcenia, jak widzimy w tabeli 16.6, zależność początkowa znika przy badaniu zależności cząstkowych.

W ramce „Uczęszczanie do uczelni należącej do Ivy League a późniejszy sukces zawodowy" Patricia Kendall, należąca do grona twórców modelu analizy rozbudowanej, przypomina badania, w których badacz podejrzewał możliwość eksplanacji, ale otrzymał replikację. Mimo że dane są już nieaktualne, temat w dalszym ciągu zajmuje żywo studentów: do jakiego stopnia sukces zawodowy zależy od wyboru „właściwej" szkoły?

Interpretacja

Interpretacja przypomina eksplanację, z wyjątkiem momentu wystąpienia zmiennej kontrolnej oraz konsekwencji tej różnicy. Interpretacja oznacza taki wynik badań, zgodnie z którym zmienna kontrolna okazuje się czynnikiem pośredniczącym między zmienną niezależną a zależną. Doskonałym przykładem interpretacji jest omówiony wcześniej przykład wykształcenia, posiadania zwolnionych od poboru przyjaciół i akceptacji faktu powołania do armii. W modelu analizy rozbudowanej wpływ wykształcenia na akceptację powołania nie ulega eksplanacji – jest to zatem zależność prawdziwa. Realnie rzecz biorąc, różnice w wykształceniu powodują różnice postaw wobec powołania. Zmienna pośrednicząca, jaką jest posiadanie zwolnionych od poboru przyjaciół, pomaga jedynie zinterpretować mechanizm, za pośrednictwem którego powstaje owa zależność. Tak więc interpretacja nie przeczy ważności początkowej zależności przyczynowej, a jedynie objaśnia proces, dzięki któremu ta zależność funkcjonuje.

Oto inny przykład interpretacji: badacze zaobserwowali, że dzieci z rozbitych rodzin częściej wchodzą na drogę przestępstwa niż dzieci z rodzin pełnych. Zależność tę można jednak zinterpretować dzięki wprowadzeniu *nadzoru* jako zmiennej pośredniczącej. Wśród dzieci należycie nadzorowanych wskaźnik przestępczości nie ma związku z faktem, że ich rodzice są lub nie są rozwiedzeni. Tak samo wygląda sprawa z dziećmi pozbawionymi nadzoru. Zależność początkowa powstaje więc dzięki zależności między rozbiciem rodziny a brakiem należytego nadzoru nad dziećmi.

Specyfikacja

Analiza rozbudowana prowadzi niekiedy do wykrycia zależności cząstkowych, które w istotny sposób różnią się między sobą. Na przykład jedna z zależności cząstkowych może być taka sama jak początkowa zależność dwuzmiennowa lub silniejsza od niej, druga zaś zależność cząstkowa może być słabsza od początkowej – lub wręcz równa zeru. W paradygmacie analizy rozbudowanej nazywamy taką sytuację **specyfikacją**: wyszczególnieniem warunków, w których zachodzi zależność początkowa.

Przypomnij sobie cytowane wcześniej badania źródeł zaangażowania religijnego (Glock, Ringer, Babbie 1967, s. 92). Odkryliśmy, że wśród wiernych Kościoła Episkopalnego zaangażowanie to jest tym mniejsze, im wyższa jest klasa społeczna. Wynik ten przedstawiony

został w tabeli 16.7, która zawiera dane dotyczące średniego poziomu zaangażowania religijnego parafianek należących do różnych klas społecznych.

Glock interpretował ten wynik w kontekście innych elementów analizy i doszedł do wniosku, że zaangażowanie religijne zapewnia alternatywną formę nagrody dla ludzi, którzy w społeczeństwie świeckim są nagród pozbawieni. Wniosek ten wyjaśniał, dlaczego kobiety są bardziej religijne niż mężczyźni,

Tabela 16.7

Klasa społeczna a średni poziom zaangażowania religijnego wśród wyznawczyń Kościoła Episkopalnego

	Klasa społeczna				
	niska				wysoka
	0	1	2	3	4
Średni poziom zaangażowania religijnego	0,63	0,58	0,49	0,48	0,45

Źródło: Tabele 16.7, 16.8 i 16.9 pochodzą z książki Charlesa Y. Glocka, Benjamina B. Ringera i Earla R. Babbiego *To Comfort and to Challenge*, University of California Press, Berkeley 1967. Wykorzystane za pozwoleniem Regentów Uniwersytetu California.

Uwaga: Tabela zawiera średnie wyniki punktowe, nie zaś procenty.

Tabela 16.8

Klasa społeczna a pełnienie funkcji w organizacjach świeckich

	Klasa społeczna				
	niska				wysoka
	0	1	2	3	4
Odsetek respondentek pełniących kiedykolwiek funkcje w organizacjach świeckich	46%	47%	54%	60%	83%

Tabela 16.9

Zaangażowanie religijne w zależności od klasy społecznej i pełnienia funkcji w organizacjach świeckich

	Średni poziom zaangażowania religijnego dla danej klasy społecznej				
	niska				wysoka
	0	1	2	3	4
Pełnią/pełniły funkcje	0,46	0,53	0,46	0,46	0,46
Nie pełnią/pełniły funkcji	0,62	0,55	0,47	0,46	0,40

ludzie starsi bardziej niż młodzież itd. Glock rozumował następująco: ludzie z niższych klas społecznych (mierzonych za pomocą dochodu i wykształcenia) mają mniejsze szanse na utrzymanie dobrej samooceny w społeczeństwie świeckim niż ludzie z klas wyższych. Na poparcie tej tezy zauważył, że klasa społeczna była mocno związana z prawdopodobieństwem pełnienia kiedykolwiek przez kobietę jakiejś funkcji w organizacji świeckiej (patrz tabela 16.8).

Glock doszedł następnie do wniosku, że zależność pomiędzy klasą społeczną a religijnością powstaje tylko dlatego, że kobiety z klas niższych pozbawione są możliwości uzyskania nagród w życiu świeckim. Tak więc zależność początkowa nie powinna zachodzić w grupie kobiet, które są w życiu świeckim nagradzane. Jako zmienną stanowiącą uproszczony wskaźnik uzyskiwania nagród w życiu świeckim Glock wykorzystał pełnienie jakiejś funkcji w organizacji świeckiej. W jego teście nie powinna istnieć zależność między zaangażowaniem religijnym a klasą społeczną wśród kobiet pełniących tego typu funkcje.

Tabela 16.9 przedstawia przykład specyfikacji. W grupie kobiet, które pełniły jakieś funkcje w organizacjach świeckich, nie dało się zasadniczo zaobserwować żadnej zależności między klasą społeczną a poziomem zaangażowania religijnego. Tabela zawiera więc specyfikację warunków, w których zachodzi zależność początkowa: wśród kobiet pozbawionych nagród w życiu świeckim.

Termin *specyfikacja* stosowany jest w paradygmacie analizy rozbudowanej bez względu na to, czy zmienna kontrolna jest uprzednia, czy pośrednicząca. Znaczenie jest w obu przypadkach to samo: określenie szczególnych warunków, w których zachodzi zależność początkowa.

Doskonalenie paradygmatu

Poprzednie partie tego rozdziału zawierały prezentację modelu analizy rozbudowanej, takiego, jakim go stworzył Lazarsfeld i jego współpracownicy. Spójrzmy teraz na kilka logicznie możliwych odmian tego modelu, część których znajdziesz w książce Morrisa Rosenberga (1968).

Po pierwsze, podstawowy paradygmat zakłada istnienie zależności początkowej między dwiema zmiennymi. Niekiedy, w modelach bardziej złożonych, może się jednak okazać użyteczne rozróżnienie zależności ujemnych i dodatnich. Co więcej, Rosenberg sugeruje, że można stosować model analizy rozbudowanej nawet przy zależności początkowej równej zeru. Podaje on jako przykład badania przynależności do związków zawodowych i postaw wobec Żydów we władzach związku (patrz tabela 16.10). Analiza początkowa wykazała, iż czas przynależności do związku nie był związany z owymi postawami: osoby należące do związku przez okres krótszy niż cztery lata były w takim samym stopniu skłonne akceptować obecność Żydów we władzach związkowych, jak osoby należące doń od ponad czterech lat. Okazało się jednak, iż *wiek* związkowców *osłabiał* zależność między stażem związkowym a postawami wobec Żydów. Młodsi członkowie związku byli ogólnie rzecz biorąc bardziej przychylnie usposobieni wobec Żydów niż starsi. Młodsi związkowcy mieli zarazem na ogół krótszy staż związkowy od starszych kolegów. Jednak w obrębie poszczególnych grup wiekowych badani z najdłuższym stażem związkowym byli najbardziej pozytywnie usposobieni do obecności Żydów we władzach związku. W tym przypadku wiek był **zmienną tłumiącą** (*suppressor variable*), maskującą zależność między stażem związkowym a postawami wobec Żydów.

Po drugie, podstawowy paradygmat koncentruje się na zależnościach cząstkowych o sile takiej samej lub mniejszej od siły zależności początkowej, nie podaje jednak wskazówek dla określenia, co stanowi istotną różnicę między zależnością początkową a zależnościami cząstkowymi. Stosując analizę rozbudowaną, często będziesz zmuszony podejmować arbitralne decyzje co do tego, czy dana zależność cząstkowa jest znacznie słabsza niż początkowa. Fakt ten skłania do uzupełnienia paradygmatu o dodatkowy wymiar.

Po trzecie, ograniczenie paradygmatu podstawowego do zależności cząstkowych takich samych lub słabszych od zależności początkowej pomija dwie inne możliwości. Zależność cząstkowa może być silniejsza niż początkowa. Może też, z drugiej strony, być jej odwróceniem – jeśli na przykład jest ujemna, a początkowa była dodatnia.

Rosenberg podaje hipotetyczny przykład tej ostatniej sytuacji, wychodząc z założenia, że badacz mógłby odkryć, iż respondenci z klasy robotniczej w większym stopniu popierają ruch obrony praw obywatelskich niż członko-

Tabela 16.10

Przykład zmiennej tłumiącej

I: Brak obserwowalnej zależności między postawami wobec Żydów a stażem związkowym

	mniej niż 4 lata	4 lata i więcej
Odsetek tych, którym nie robi różnicy obecność Żydów we władzach związku	**49,2%** (126)	50,5% (256)

II: W każdej z grup staż związkowy wpływa na wzrost akceptacji Żydów we władzach związku

	staż związkowy	
Odsetek tych, którym nie robi różnicy obecność Żydów we władzach związku	mniej niż 4 lata	4 lata i więcej
wiek:		
29 lat i mniej	56,4% (78)	62,7% (51)
30-49 lat	37,1% (35)	48,3% (116)
50 lat i więcej	38,4% (13)	56,1% (89)

Źródło: Zaadaptowano na podstawie Morris Rosenberg, *The Logic of Survey Analysis*, Basic Books, New York 1968, s. 88–89.

wie klasy średniej (patrz tabela 16.11). Następnie sugeruje on, że rasa mogłaby być w tym przypadku **zmienną zniekształcającą** (*distorter variable*), odwracającą prawdziwą zależność między klasą a postawami. Afroamerykanie są prawdopodobnie większymi zwolennikami ruchu obrony praw obywatelskich, ale też Afroamerykanie byliby najprawdopodobniej nadreprezentowani w próbie członków klasy robotniczej, niedostatecznie zaś reprezentowani w próbie członków klasy średniej. Afroamerykanie należący do klasy średniej mogliby jednak popierać ruch bardziej niż Afroamerykanie z klasy robotniczej; z podobną zależnością mielibyśmy do czynienia wśród białych. Przy stałej wartości zmiennej *rasa* badacz musiałby więc stwierdzić, że poparcie dla ruchu obrony praw obywatelskich jest większe wśród członków klasy średniej niż wśród robotników.

Oto inny przykład działania zmiennej zniekształcającej. Kiedy Michele de Seve postanowiła zbadać początkowe zarobki mężczyzn i kobiet w tej samej organizacji, była bardzo zaskoczona, gdy odkryła, że kobiety otrzymują przeciętnie wyższe pierwsze wynagrodzenie niż ich koledzy. Zmienną zniekształcającą okazał się *moment pierwszego zatrudnienia*. Wiele kobiet zostało zatrudnionych stosunkowo niedawno, gdy stawki wynagrodzeń były ogólnie wyższe niż w latach ubiegłych, kiedy to zatrudniono wielu mężczyzn (Cook 1995).

Wszystkie te nowe aspekty dodatkowo komplikują pojęcie specyfikacji. Jak powinien zareagować badacz, jeśli jedna z zależności cząstkowych jest taka sama jak zależność początkowa, a druga jest jeszcze silniejsza? Udało się wprawdzie określić jeden z warunków, w których utrzymana jest zależność początkowa, zarazem jednak wyspecyfikowany został kolejny warunek, który czyni ją jeszcze wyraźniejszą.

Na koniec należy zauważyć, że podstawowy paradygmat skupia się przede wszystkim na dychotomicznych zmiennych kontrolnych. W rzeczywistości model analizy rozbudowanej nie jest tak ograniczony – ani teoretycznie, ani też praktycznie – jednak podstawowy paradygmat staje się bardziej złożony, gdy zmienna kontrolna dzieli próbę na trzy podgrupy i więcej. Paradygmat staje się jeszcze bardziej

Tabela 16.11

Przykład zmiennej zniekształcającej (hipotetyczny)

I: Badani z klasy robotniczej reprezentują bardziej liberalne podejście do ruchu obrony praw obywatelskich niż badani z klasy średniej

Prawa obywatelskie (wynik w punktach)	Klasa średnia	Klasa robotnicza
Wysoki	37%	45%
Niski	63%	55%
	100%	100%
100% =	(120)	(120)

II: Wprowadzenie rasy jako zmiennej kontrolnej dowodzi, że klasa średnia jest bardziej liberalna niż klasa robotnicza

	Klasa społeczna			
	czarni		biali	
Prawa obywatelskie (wynik w punktach)	klasa średnia	klasa robotnicza	klasa średnia	klasa robotnicza
Wysoki	70%	50%	30%	20%
Niski	30%	50%	70%	80%
	100%	100%	100%	100%
100% =	(20)	(100)	(100)	(20)

Źródło: Morris Rosenberg, *The Logic of Survey Analysis*, Basic Books, New York 1968, s. 94–95. Wykorzystano za pozwoleniem.

skomplikowany, gdy używamy jednocześnie więcej niż jednej zmiennej kontrolnej.

Nie chcę bynajmniej negować użyteczności podstawowego paradygmatu analizy rozbudowanej. Wręcz przeciwnie, chciałbym podkreślić, że model analizy rozbudowanej nie jest prostym algorytmem – zbiorem procedur, w toku których należy analizować wyniki badań. Jest to raczej przede wszystkim narzędzie logiczne, pomocne badaczowi w zrozumieniu uzyskanych danych. Solidne opanowanie modelu analizy rozbudowanej ułatwia bardziej wyrafinowaną analizę. Model ten nie podpowiada jednak, jakie należy wprowadzić zmienne kontrolne, ani też, jakie są ostateczne wnioski co do znaczenia wyników analizy. Wszystko to musi ci podpowiedzieć własna pomysłowość. Pomysłowości owej nie nabędziesz jednak inaczej, jak tylko dzięki długiemu doświadczeniu. Wskazując na nadmierne uproszczenia w podstawowym paradygmacie analizy rozbudowanej, starałem się dowieść, że model ten stanowi jedynie pewien ogólny schemat logiczny. Zetkniesz się z przy-

kładami wyrafinowanych analiz znacznie bardziej skomplikowanych niż wykorzystane w tym rozdziale.

Zarazem jednak, jeśli w pełni zrozumiesz model podstawowy, znacznie łatwiej ci będzie zrozumieć też inne techniki, takie jak korelacje, regresję i analizę czynnikową. W rozdziale 17 techniki te są prezentowane w kontekście modelu analizy rozbudowanej jako korelacje i regresje cząstkowe.

Analiza rozbudowana i stawianie hipotez *post factum*

Zanim zakończymy rozważania na temat modelu analizy rozbudowanej, powinniśmy przyjrzeć się jego roli w kontekście jednej z form błędnego rozumowania, zwanej stawianiem **hipotez *post factum***. Czytelnik literatury metodologicznej znajdzie bez trudu niezliczone przykłady wypowiedzi przestrzegających przed tym błędem. Jednak mimo że intencje tego typu ostrzeżeń są słuszne, niedoświadczeni badacze mogą niekiedy poczuć się zagubieni, próbując zrozumieć ich znaczenie.

„*Post factum*" oznacza „po fakcie". Kiedy obserwujesz zależność empiryczną między dwiema zmiennymi i następnie proponujesz po prostu wyjaśnienie dla tej zależności, wówczas stawiasz *hipotezę post factum*. Stworzyłeś hipotezę łączącą dwie zmienne po tym, jak już poznałeś zależność między nimi. Zapewne przypominasz sobie rozważania we wcześniejszych partiach tej książki, gdy pisałem, że każda znacząca hipoteza musi być obalalna. Jeśli nie jesteś w stanie określić, jakie odkrycia empiryczne stanowiłyby *zaprzeczenie* twojej hipotezy, nie jest to prawdziwa *hipoteza* w naukowym znaczeniu tego słowa. Możesz bowiem przyjąć, że skoro już zaobserwowałeś daną zależność, żadna hipoteza na jej temat nie może zostać obalona.

Jest to uczciwe stwierdzenie, jeśli nie robisz nic innego, niż ubieranie – po fakcie – twoich obserwacji empirycznych w obalalne hipotezy. Dokonawszy obserwacji, iż kobiety są bardziej religijne niż mężczyźni, nie powinieneś tylko po prostu stwierdzić, że kobiety są bardziej religijne ze względu na jakąś ogólną dynamikę zachowań społecznych i pozostać przy tej wstępnej obserwacji.

Argumentacja przeciwko hipotetyzowaniu *post factum* zwraca się przeciwko samej sobie, ponieważ deprecjonuje dobre, uczciwe hipotezy postawione po fakcie. Niedoświadczeni badacze często skłaniają się więc ku przekonaniu, że należy postawić wszelkie hipotezy przed zbadaniem danych – nawet jeśli oznacza to, że hipotezy te będą w znacznej części nieprzemyślane. Co więcej, badacze owi często ignorują empirycznie obserwowalne zależności tylko dlatego, że nie potwierdzają one wcześniejszych hipotez.

Z całą pewnością niewielu badaczy życzyłoby sobie, by Sam Stouffer zachował dla siebie dziwaczne wyniki swoich badań na temat nastrojów w armii. Stouffer dokonał szczególnych obserwacji empirycznych i zabrał się do stawiania hipotez na temat możliwych przyczyn obserwowanych wyników. Jego rozumowanie okazało się nieocenione dla wszystkich późniejszych badaczy. Kluczowe znaczenie ma w tym przypadku fakt, że jego hipotezy *post factum* same były również sprawdzalne.

Należy jednak zwrócić w tym miejscu uwagę na jeszcze inny problem. Każdy może oczywiście stworzyć hipotezę wyjaśniającą obserwowane zależności empiryczne, jednak model analizy rozbudowanej dostarcza narzędzi logicznych dla *testowania* tych hipotez z użyciem tych samych danych. Dobrym przykładem testowania może być omówiona wcześniej problematyka klasy społecznej i zaangażowania religijnego. Glock wyjaśnił zależność początkową w kategoriach teorii deprywacji społecznej. Gdyby zatrzymał się w tym punkcie, jego uwagi byłyby zapewne interesujące, lecz niezbyt przekonujące. Glock poszedł jednak dalej. Zauważył on mianowicie, że jeśli jego hipoteza jest prawdziwa, zależność między klasą społeczną a zaangażowaniem religijnym powinna zniknąć w grupie tych kobiet, które są nagradzane w społeczeństwie świeckim – czyli na przykład tych, które pełnią jakieś funkcje w świeckich organizacjach. Hipotezę tę przetestował następnie empirycznie. Gdyby ta

nowa hipoteza nie znalazła potwierdzenia w danych, Glock byłby zmuszony rozważyć ją od nowa.

Te dodatkowe uwagi stanowią dalsze uzasadnienie poglądu, że analiza danych jest procesem ciągłym, wymagającym zaangażowania całej dostępnej pomysłowości i wytrwałości. Obraz badacza, starannie formułującego hipotezy i sprawdzającego je następnie z rytualną pieczołowitością, przynosi w efekcie jedynie rytualne badania.

Jeśli obawiasz się, że wartość dowodów *post factum* jest mniejsza niż dowodów tradycyjnych, pozwól mi raz jeszcze powtórzyć to, co mówiłem już wcześniej, a mianowicie, że „dowód naukowy" to sprzeczność sama w sobie. W nauce nigdy niczego się nie *dowodzi*. Hipotezy, wyjaśnienia, teorie, przypuszczenia – wszystko to może się ostać w ogniu prób obalenia, nigdy jednak nie może zostać dowiedzione w jakimkolwiek sensie absolutnym. Przyjęcie hipotezy jest zatem tak naprawdę funkcją zakresu, w jakim została ona sprawdzona i nie została obalona. Nie ma więc hipotezy, którą można by uznać za prawdziwą na podstawie jednego testu – niezależnie od tego, czy hipoteza ta powstała przed obserwacją zjawisk empirycznych, czy po takiej obserwacji. Mając to na uwadze, nie powinieneś odmawiać sobie niekiedy szczególnie obfitujących w odkrycia ścieżek analizy danych. Zawsze powinieneś starać się dojść do rzetelnego zrozumienia twoich danych, wypracować znaczące teorie w celu uogólnienia wyjaśnienia i nie martwić się o sposób dojścia do zrozumienia.

Główne zagadnienia

- Model analizy rozbudowanej jest metodą analizy wielozmiennowej, właściwą dla nauk społecznych. Jest to przede wszystkim model logiczny, który pozwala zrozumieć podstawy logiczne innych metod wielozmiennowych.
- Podstawowe etapy analizy rozbudowanej są następujące: (1) zaobserwowana zostaje zależność między dwiema zmiennymi, (2) wartość trzeciej zmiennej utrzymuje się na stałym poziomie w tym znaczeniu, że badane przypadki dzieli się wedle wartości tej trzeciej zmiennej, (3) początkowa zależność dwuzmiennowa jest liczona od nowa w każdej z podgrup i (4) porównanie zależności początkowej z zależnościami w każdej z podgrup umożliwia pełniejsze wyjaśnienie zależności wyjściowej.
- Zależność logiczna między zmiennymi zależy od tego, czy zmienna kontrolna jest uprzednia w stosunku do pozostałych dwóch zmiennych, czy też pośrednicząca między nimi.
- Zależność zerowa to zależność obserwowana między dwiema zmiennymi bez kontrolowania wartości trzeciej zmiennej.
- Zależności cząstkowe to zależności obserwowane między dwiema zmiennymi w obrębie podgrupy przypadków, wyodrębnionej według wartości zmiennej kontrolnej.
- Wynikiem analizy rozbudowanej może być replikacja (gdy zbiór zależności cząstkowych jest taki sam, jak odpowiednia zależność początkowa), eksplanacja (gdy zbiór zależności cząstkowych zostaje sprowadzony do zera przy stałym poziomie zmiennej kontrolnej uprzedniej), interpretacja (gdy zbiór zależności cząstkowych zostaje sprowadzony do zera przy stałym poziomie zmiennej kontrolnej pośredniczącej), lub też specyfikacja (gdy jedna z zależności cząstkowych jest zredukowana, najlepiej do zera, druga zaś pozostaje taka sama jak zależność początkowa, lub jest silniejsza).
- Zmienna tłumiąca osłabia zależność między dwiema zmiennymi, zmienna zniekształcająca powoduje pozorne odwrócenie zależności między dwiema zmiennymi (z ujemnej na dodatnią lub na odwrót).
- Stawianie hipotez *post factum*, czyli budowanie hipotez „przewidujących" zależności już zaobserwowane, jest w nauce błędem, ponieważ hipotez takich nie można obalić. Nic nie stoi na przeszkodzie proponowaniu powodów, dla których zachodzą obserwowane zależności, nie powinno się jednak ujmować tych propozycji w formę „hipotez". Co więcej, jedna zaobserwowa-

na zależność i jej możliwe wyjaśnienia mogą podpowiadać hipotezy co do innych, jeszcze nie zbadanych zależności. Model analizy rozbudowanej jest wspaniałym narzędziem logicznym dla tego rodzaju postępującej analizy danych.

Kluczowe pojęcia

model analizy rozbudowanej
zmienna testowa (kontrolna)
zależności cząstkowe
zależność początkowa
replikacja
eksplanacja
interpretacja
specyfikacja
zmienna tłumiąca
zmienna zniekształcająca
hipoteza *post factum*

Pytania i ćwiczenia

1. Przypomnij sobie przykład z badań Stouffera, Kendall i Lazarsfelda na temat wykształcenia, zwolnionych od poboru przyjaciół i postaw badanych wobec faktu powołania do służby wojskowej. Załóżmy, że wyszli oni od związku między posiadaniem zwolnionych przyjaciół a postawami wobec faktu powołania do służby wojskowej, jako zmiennej kontrolnej użyli zaś wykształcenia. Do jakich doszliby wówczas wniosków?
2. Opisz własnymi słowami logikę analizy rozbudowanej w przypadku: (a) replikacji, (b) interpretacji, (c) eksplanacji i (d) specyfikacji.
3. Przypomnij sobie ramkę o Ivy League i sukcesach w późniejszym życiu zawodowym. Wyjaśnij własnymi słowami, co ma na myśli Patricia Kendall, gdy pisze: „Pomimo to [pomimo wsparcia uzyskanego dzięki analizie zależności cząstkowych] West nie zdołała w żadnym razie udowodnić swojej hipotezy". Jakie wnioski można w sposób uprawniony wyciągnąć z analizy West?
4. Podaj hipotetyczne przykłady zmiennych tłumiących i zmiennych zniekształcających.
5. Poszukaj w Internecie raportu z badań, zawierającego odkrycie zależności pozornej. Podaj adres strony i przytocz lub zreferuj, czego dotyczyły te badania.

Dodatkowe lektury

Continuities in the Language of Social Research, Paul F. Lazarsfeld, Ann Pasanella, Morris Rosenberg (red.), Free Press, New York 1972. Bardzo dobry, klasyczny zbiór rozważań teoretycznych i przykładów empirycznych. Część II jest szczególnie ważna, chociaż logika analizy rozbudowanej zajmuje przeważającą część tomu.

Lazarsfeld Paul F., *The Varied Sociology of Paul F. Lazarsfeld*, Patricia L. Kendall (red.), Columbia University Press, New York 1982. Zbiór esejów na temat logiki badań społecznych, autorstwa głównego twórcy modelu analizy rozbudowanej, w redakcji jej współtwórczyni.

Qualitative and Quantitative Social Research, Robert K. Merton, James S. Coleman, Peter H. Rossi (red.), Free Press, New York 1979. Ten zbiór artykułów napisanych ku czci Lazarsfelda przedstawia porządek logiczny, jaki wniósł on do badań społecznych.

Rosenberg Morris, *The Logic of Survey Analysis*, Basic Books, New York 1968. Najpełniejsza prezentacja analizy rozbudowanej. Rosenberg przedstawia podstawowy paradygmat i sugeruje logiczne możliwości jego poszerzenia. Niełatwo powiedzieć, co jest bardziej wartościowe w tej książce – sama prezentacja czy ilustrujące ją liczne przykłady. Jedno i drugie jest znakomite.

Survey Research in the Social Sciences, Glock Charles (red.), Russell Sage, New York 1967, rozdział 1. Pierwszorzędne omówienie logiki analizy rozbudowanej. Rozdział napisany przez Glocka zawiera prezentację modelu analizy rozbudowanej wraz z rzeczowymi przykładami.

Tacq Jaques, *Multivariate Analysis Techniques in Social Science Research: From Problem to Analysis*, Sage, Thousand Oaks, Cal. 1997. Bardzo przystępna książka dla tych, którzy chcieliby poznać nieco dokładniej statystyczne zaplecze analizy wielozmiennowej

w modelu analizy rozbudowanej. Autor używa prawdziwych przykładów badań z różnych dziedzin, by wyjaśnić, jak i do czego można stosować analizę rozbudowaną.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział 17

Statystyka w naukach społecznych

Przegląd holograficzny

Dzięki statystyce możemy syntetycznie opisywać dane, mierzyć zależności między zmiennymi oraz formułować na podstawie wyników uzyskanych w próbie wnioski dotyczące całej populacji. Opracowanie kilku najprostszych i najczęściej wykorzystywanych w badaniach społecznych umiejętności statystycznych jest znacznie mniej bolesne (i mniej groźne dla twojego życia towarzyskiego), niż mógłbyś sądzić.

Wprowadzenie

Niebezpieczeństwo sukcesów matematycznych

Statystyka opisowa
Redukcja danych
Miary siły związku
Analiza regresji

Inne techniki wielozmiennowe
Analiza ścieżkowa
Analiza szeregów czasowych
Analiza czynnikowa

Wnioskowanie statystyczne
Wnioskowanie jednozmiennowe
Testy istotności statystycznej
Logika istotności statystycznej
Test chi-kwadrat

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Od lat obserwuję, że statystyka onieśmiela wielu studentów. Niekiedy statystyka sprawia, że czują się oni:

- Jak clowni, ale bez cyrku
- Głupsi od własnych butów
- Jak pióra, które zgubiły swoją kaczkę
- Jak piana bez piwa
- Jakby przestało im działać zdalne sterowanie
- Jak fasolki, które wypadły ze strączka
- Błyskotliwi jak kula do kręgli
- Zakręceni jak kij bejsbolowy
- Jakby im brakowało 4 centów, żeby mieć piątaka
- Jak samochód na pełnych obrotach, który nie może ruszyć z miejsca*

Wielu ludzi boi się badań ilościowych tylko dlatego, że odpycha ich stosowana w nich matematyka i statystyka. I rzeczywiście – wiele raportów z badań składa się głównie z rachunków podanych bez żadnego komentarza. Statystyka często odgrywa w badaniach społecznych istotną rolę, ale równie ważne jest, by widzieć tę rolę we właściwych proporcjach.

Analiza wyników badań empirycznych ma charakter raczej logiczny niż matematyczny. Matematyka jest jedynie wygodnym i skutecznym językiem, który możemy stosować w operacjach logicznych, właściwych analizie danych ilościowych. *Statystyka* jest gałęzią matematyki stosowanej, która może być szczególnie użyteczna w różnorodnych analizach.

W tym rozdziale zaczniemy od pewnej niezbyt formalnej analizy skojarzeń wywoływanych u wielu ludzi przez kontakt ze statystyką. Mam nadzieję, że po tym ćwiczeniu będzie ci łatwiej zrozumieć relatywnie proste podstawy statystyki, zaprezentowane w dalszej części rozdziału, i że nie będziesz się ich bać. Zapoznamy się z dwiema dziedzinami statystyki: statystyką opisową oraz wnioskowaniem statystycznym. *Statystyka opisowa* jest narzędziem opisywania danych w poręcznej formie, tak, by można było łatwo zobaczyć, co w istocie zawierają. *Wnioskowanie statystyczne* natomiast pomaga badaczowi formułować wnioski na podstawie obserwacji; z reguły dotyczy ono wyciągania wniosków o populacji na podstawie badań przeprowadzonych na próbie.

* Powyższe zestawienie zawdzięczam licznym autorom internetowych stron z humorem.

Niebezpieczeństwo sukcesów matematycznych

Kiedy zaczynałem uczyć metod badawczych, które choć trochę zahaczały o statystykę, byłem zaskoczony ogromną liczbą studentów tłumaczących, że są „niezbyt dobrzy z matematyki". Tak jak o niektórych ludziach wiemy, że kompletnie nie mają słuchu, a o innych, że nie są w stanie nauczyć się żadnego języka obcego, tak 90% moich studentów sprawiało wrażenie dotkniętych Syndromem Wrodzonego Upośledzenia Matematycznego (SWUM), z jego typowymi symptomami – frustracją, znudzeniem i sennością. Z przyjemnością donoszę, że udało mi się wreszcie znaleźć główną przyczynę tej choroby, a nawet opracować skuteczną kurację. Ponieważ i ty możesz cierpieć na tę chorobę, chciałbym podzielić się z tobą moją wiedzą, zanim zaczniemy nasze zmagania ze statystyką w badaniach społecznych.

Być może zetknąłeś się już z historią Mary Mallon, znanej jako Tyfusowa Mary. Była ona nosicielką pałeczki duru brzusznego, zmarłą w Nowym Jorku w 1938 roku. Przed śmiercią pracowała jako kucharka; w kolejnych domach, gdzie pracowała wywołała dziesięć ognisk choroby – zaraziła łącznie 50 osób, z których trzy zmarły.

Syndrom Wrodzonego Upośledzenia Matematycznego ma podobne źródło. Po wyczerpujących badaniach znalazłem winnego, którego przedstawię wam pod nazwiskiem Matematycznego Marvina (choć używa również wielu innych). Jeśli cierpisz na SWUM, najprawdopodobniej miałeś już z nim do czynienia. Wytęż pamięć – czy pamiętasz z czasów szkoły średniej osobnika, którego twoi koledzy i nauczyciele określali jako „geniusza matematycznego", który nie tylko dostawał z matematyki same piątki, ale zdawał się znać ją lepiej od nauczyciela?

Jeśli już go sobie przypomniałeś, to mam parę pytań. Płeć? Zgaduję, że pewnie był to chłopak – tak w każdym razie twierdziła większość moich studentów. Ale dalej:

1. Czy był dobrze zbudowany?
2. Czy nosił okulary?
3. Na ile imprez (w całym okresie chodzenia do szkoły) był zaproszony?
4. Jeśli już był na imprezie, to czy ktokolwiek z nim wtedy rozmawiał?
5. Ile razy miałeś poczucie, że zazdrościsz matematycznemu geniuszowi jego pozycji i chciałbyś zająć jego miejsce?

Zadawałem te pytania moim studentom (także takim, którzy ukończyli szkołę wiele lat wcześniej) przez wiele lat; odpowiedzi, które dostawałem, są zaskakująco zgodne. Marvin z reguły jest kiepsko zbudowany (albo za chudy, alba za gruby). Zwykle nosi okulary i w ogóle sprawia wrażenie słabeusza. Przez cały okres chodzenia do szkoły zapraszany jest na (średnio) 1, 2 imprezy, na których nikt z nim nie rozmawia. Ma fatalną cerę. Jest niedostosowany społecznie, nikt nie chciałby być na jego miejscu i w ogóle budzi raczej litość niż zazdrość.

W miarę jak rozmawiałem o Marvinie z moimi studentami, stawało się dla mnie coraz bardziej oczywiste, że większość z nich podświadomie wiąże sprawność w matematyce z jego godną pożałowania charakterystyką. Zdaje się, że doszli oni do wniosku, że gdyby zaczęli coś z matematyki i statystyki rozumieć, to sami staliby się niedostosowani społecznie – a to byłaby zbyt wysoka cena.

Wszystko to, co powiedziałem o Matematycznym Marvinie, odpowiada powszechnemu stereotypowi – ale tylko stereotypowi: to nie ma nic wspólnego z biologią. Są kobiety świetne w matematyce. Atrakcyjni ludzie potrafią liczyć nie gorzej niż nieatrakcyjni. Nie zmienia to faktu, że stereotypy mają przemożny wpływ na nasze zachowania – co pokazują tragiczne przykłady kobiet demonstrujących matematyczną ignorancję w przekonaniu, że są przez to atrakcyjniejsze.

Tak więc jeśli należysz do tych, którzy uważają się za „niezbyt dobrych z matematyki", to być może chodzi tu o twój podświadomy strach, że jeśli na kursie statystyki pójdzie ci lepiej niż źle, to, powiedzmy, dostaniesz pryszczy. Jeśli tak, to zapewne czytając dalszy ciąg tego rozdziału, będziesz się strasznie męczyć starając się jednocześnie cokolwiek zrozumieć (na tyle, żeby zdać egzamin), a przy tym nic nie zrozumieć (żeby nie stracić przyjaciół).

Nie ma jednak powodu do obaw. Badania wykazują, że dawka statystyki, zawarta w tym rozdziale, nie jest niebezpieczna dla ludzi, a poza tym nie udokumentowano żadnego przypadku, w którym zrozumienie, co to jest lambda, gamma, chi-kwadrat lub którakolwiek z innych rzeczy, którymi będziemy się teraz zajmować, wywołało choć jeden pryszcz. Wręcz przeciwnie – taka porcja statystyki, na którą będziesz tutaj wyeksponowany, jest podobno dla młodych badaczy bardzo zdrowa.

A swoją drogą, odkrycie Marvina pozwoliło mi wyjaśnić wiele tajemnic (nota bene „mój" Marvin nie nosił okularów, ale za to miał potwornego zeza). W pierwszej mojej książce o metodologii badań zawarłem dokładnie trzy obliczenia matematyczne – w tym jedno błędne. W pierwszym wydaniu tej książki popełniłem błąd w innym miejscu. Ale w tej wcześniejszej książce zrobiłem jedną jeszcze gorszą rzecz: uznałem, że zamiast przepisywać od kogoś tabelę liczb losowych, przyjemniej będzie napisać program komputerowy, który stworzy moją własną tabelę. W rezultacie osiągnąłem wątpliwy zaszczyt opublikowania pierwszej na świecie tabeli liczb losowych, w której nie występowała ani jedna dziewiątka! Dopiero kiedy wyśledziłem Marvina, odkryłem źródło moich kłopotów, statystyka zaś zaczęła sprawiać mi więcej przyjemności (i mniej trudności). Czego i tobie życzę.

Statystyka opisowa

Jak już wspomniałem, **statystyka opisowa** służy do przedstawiania danych ilościowych w syntetycznej formie. Czasami chcemy opisać jedną zmienną, czasami zaś chcemy opisać związek, łączący jedną zmienną z inną. Przyjrzyjmy się różnym technikom, które możemy do tego wykorzystywać.

Redukcja danych

Badania naukowe często wiążą się z gromadzeniem olbrzymich ilości danych. Wyobraźmy sobie, że przeprowadziliśmy sondaż (sto pytań w kwestionariuszu) na próbie 2000 respondentów. Nie byłyby to jakieś wyjątkowo duże badania, ale zgromadzilibyśmy w nim 200 000 odpowiedzi! Nikt nie byłby w stanie przeczytać ich wszystkich i do tego jeszcze wyciągnąć z nich jakiekolwiek sensowne wnioski. Analiza naukowa wymaga zatem redukcji danych do formy bardziej syntetycznej.

Na początek przyjrzyjmy się macierzy danych surowych, będącej produktem ilościowego projektu badawczego. Tabela 17.1 zawiera fragment takiej macierzy. Każdy jej wiersz odpowiada jednej osobie (czy ogólniej – jednostce analizy), każda kolumna odpowiada zmiennej, każda komórka zaś zawiera zakodowany atrybut lub wartość przypisaną danej osobie dla danej zmiennej. Pierwsza kolumna tabeli 17.1 reprezentuje płeć badanego. Powiedzmy, że „1" oznacza mężczyznę, a „2" oznacza kobietę. Tak więc osoby 1 i 2 to mężczyźni, osoba 3 jest kobietą itd.

Jeśli chodzi o wiek, „3" przypisane osobie 1 mogłoby oznaczać przedział wieku 30–39 lat, a „4" osoby 2 przedział 40–49 lat. Niezależnie od tego, w jaki sposób wiek byłby zakodowany (patrz rozdział 14), wartości kodów z tabeli 17.1 opisują wszystkie reprezentowane w niej osoby.

Zauważ, że dane zostały zredukowane już na etapie przygotowywania tej tabeli. Jeśli wiek był zakodowany tak, jak opisałem to przed chwilą, konkretna odpowiedź „33 lata" została przypisana do kategorii „30–39 lat". W badaniu mogło paść 60 czy 70 różnych odpowiedzi na pytanie o wiek, ale zredukowaliśmy je wszystkie do 6 lub 7 kategorii.

W rozdziale 15 omawialiśmy niektóre metody dalszej redukcji danych, dotyczących jednej zmiennej, poprzez użycie miar wartości przeciętnej, takich jak modalna, mediana czy średnia oraz miar rozproszenia, takich jak zakres, odchylenie standardowe itd. W podobny sposób można również syntetycznie opisywać związki między zmiennymi.

Miary siły związku

Także związek między dwiema zmiennymi można przedstawić w postaci macierzy danych, na przykład zawierającej łączne rozkłady częstości tych dwóch zmiennych – tak jak w tabeli 17.2, która zawiera wszystkie informacje niezbędne do określenia charakteru i siły związku między wykształceniem a uprzedzeniami.

Możemy się z niej dowiedzieć, że 23 badanych (1) nie miało żadnego wykształcenia oraz (2) charakteryzowało się wysokim poziomem uprzedzeń, 77 badanych zaś (1) miało pełne wyższe wykształcenie oraz (2) charakteryzowało się niskim poziomem uprzedzeń.

Podobnie jak macierz danych surowych z tabeli 17.1, tabela 17.2 zawiera zbyt dużo informacji, by dało się je łatwo przyswoić. Uważnie czytając dane w tabeli, moglibyśmy wprawdzie zauważyć, że gdy wykształcenie rośnie od „żadnego" do „pełnego wyższego", poziom uprzedzeń generalnie się obniża – ale nie byli-

Tabela 17.1

Fragment macierzy danych surowych

	Płeć	Wiek	Wykształcenie	Dochód	Zawód	Przynależność polityczna	Orientacja polityczna	Przynależność religijna	Znaczenie przypisywane religii
Osoba 1	1	3	2	4	1	2	3	0	4
Osoba 2	1	4	2	4	4	1	1	1	2
Osoba 3	2	2	5	5	2	4	4	2	3
Osoba 4	1	5	4	4	3	2	2	2	4
Osoba 5	2	3	7	8	6	1	1	5	1
Osoba 6	2	1	3	3	5	5	5	1	1

Tabela 17.2

Hipotetyczne dane surowe dotyczące wykształcenia i uprzedzeń

	Wykształcenie				
Uprzedzenia	brak	podstawowe	średnie	wyższe (college)	wyższe (pełne)
Silne	23	34	156	67	16
Średnie	11	21	123	102	23
Słabe	6	12	95	164	77

byśmy w stanie sformułować żadnego bardziej konkretnego wniosku. Aby dokładniej podsumować tę macierz danych, możemy się odwołać do którejś z miar oferowanych przez statystykę opisową; wybór właściwego współczynnika zależy przede wszystkim od charakteru obu analizowanych zmiennych.

W bieżącym rozdziale przyjrzymy się kilku metodom opisu związku między dwiema zmiennymi. Współczynniki, z którymi się zapoznamy, skonstruowane są na podstawie tego samego modelu – **proporcjonalnej redukcji błędu**[1] (**PRE** – *proportional reduction of error*).

Co to jest model PRE? Wyobraź sobie, że poprosiłem cię, byś spróbował odgadnąć, jakie wartości przyjmuje jakaś zmienna dla poszczególnych respondentów, na przykład czy dany badany na któreś z pytań kwestionariusza odpowiedział „tak", czy „nie". Aby ułatwić ci zadanie, załóżmy, że znasz ogólny rozkład odpowiedzi w całej próbie – powiedzmy, 60% „tak", 40% „nie". W takiej sytuacji, aby popełnić najmniej pomyłek, powinieneś za każdym razem przewidywać, że badany udzielił odpowiedzi będącej modalną (tj. najczęstszą odpowiedzią), a więc w tym przypadku „tak".

Załóżmy teraz, że poznasz także empiryczną zależność między zmienną, której wartość odgadujesz, a jakąś inną zmienną, na przykład płcią, po czym za każdym razem, gdy masz dla jakiejś osoby odgadnąć, czy na pytanie kwestionariusza odpowiedziała ona „tak", czy „nie", informuje cię, czy jest to kobieta, czy mężczyzna. Jeśli między obiema zmiennymi zachodzi jakiś związek, to informacja ta powinna pozwolić ci ograniczyć liczbę popełnianych błędów. Znając siłę związku między zmiennymi, możemy obliczyć, jaki procent błędu udało się wyeliminować (czyli obliczyć PRE): im związek ten jest silniejszy, tym większa jest redukcja błędu.

Podstawowy model PRE podlega modyfikacjom zależnym od typu skali pomiaru wartości zmiennej – nominalnej, porządkowej lub interwałowej. Zapoznamy się teraz z miarami siły związku właściwymi dla każdej z tych skal, musisz jednak pamiętać, że trzy współczynniki, o których będziemy mówili, to tylko niektóre z bardzo wielu stosowanych w praktyce.

Zmienne nominalne

Jeżeli dwie zmienne mierzone są na skali nominalnej (zawierają informacje dotyczące na przykład płci, przynależności religijnej, rasy), jednym z dopuszczalnych współczynników siły związku jest λ (lambda). (Lambda jest jedną z liter alfabetu greckiego, odgrywającą w nim podobną rolę jak *l* w naszym. Fakt, że w statystyce często używa się greckich liter, może po części wyjaśniać, dlaczego tak wiele osób słysząc o statystyce udaje Greka). Współczynnik λ jest to wartość proporcjonalnej redukcji błędu odgadywania wartości zmiennej, osiągniętej dzięki znajomości wartości innej zmiennej.

Wyobraźmy sobie następującą sytuację: w pokoju jest 100 osób; twoim zadaniem jest odgadnięcie płci każdej z nich. Jeśli połowa z nich to mężczyźni, a druga połowa to kobiety, będziesz się zapewne mylił w połowie przypadków, a w połowie przypadków odgadniesz właściwie. Ale załóżmy, że dodatkowo, za każdym razem gdy masz odgadnąć płeć jakiejś osoby, jesteś przedtem informowany o jej zawodzie.

Jaką płeć podasz dla osoby, o której wiesz, że pracuje jako kierowca ciężarówki? Zapewne jesteś dość inteligentny, by powiedzieć „mężczyzna" – chociaż i w tym zawodzie jest już trochę kobiet, większość kierowców ciężarówek to jednak mężczyźni. Jeśli kolejna osoba będzie pracowała jako opiekun do dzieci, na tej samej zasadzie będziesz zapewne zgadywał, że jest to kobieta. Co prawda wciąż bę-

[1] Ściślej rzecz biorąc: proporcjonalnej redukcji funkcji błędu przewidywania (przyp. tłum.).

Tabela 17.3

Hipotetyczne dane dotyczące związku między płcią a zatrudnieniem

	Mężczyźni	Kobiety	Razem
Pracujący	900	200	1100
Niepracujący	100	800	900
Razem	1000	1000	2000

dziesz się czasami mylić, ale znacznie rzadziej, niż gdybyś nie znał zawodów poszczególnych osób. To, w jakim stopniu udało się ograniczyć pomyłki (proporcjonalna redukcja błędu), potraktujemy jako miarę związku łączącego płeć i zawód.

Przyjrzyjmy się innemu prostemu przykładowi, ilustrującemu logikę i metodę wyznaczania współczynnika λ. Tabela 17.3 zawiera fikcyjne dane dotyczące relacji między statusem zatrudnienia a płcią. Generalnie, w naszej populacji jest 900 osób niepracujących oraz 1100 pracujących. Gdybyś miał przewidzieć dla każdego z badanych, jaki jest jego status zatrudnienia, korzystając jedynie ze znajomości ogólnego rozkładu zmiennej, powinieneś zawsze przewidywać „pracujący", gdyż w ten sposób popełnisz mniej błędów, niż gdybyś o każdej osobie przewidywał, że nie pracuje. Ale i tak przyjmując taką strategię, popełniasz błędy w 900 z 2000 przypadków.

Gdybyś jednak miał dostęp do danych z tabeli 17.3 oraz, przewidując status zatrudnienia, znał płeć danej osoby, przewidywałbyś inaczej. O każdym mężczyźnie zgadywałbyś, że jest on „pracujący", a o każdej kobiecie, że jest „niepracująca". W tej sytuacji pomyliłbyś się 300 razy – 100 razy w odniesieniu do niepracujących mężczyzn i 200 razy w odniesieniu do pracujących kobiet, a zatem zrobiłbyś o 600 błędów mniej w porównaniu z sytuacją, gdy nie znałeś płci poszczególnych badanych.

Współczynnik λ równy jest proporcji między wielkością, o jaką zmniejszyła się liczba popełnionych błędów, a liczbą błędów, które zostałyby popełnione, gdyby przewidywanie było robione wyłącznie na podstawie rozkładu zmiennej przewidywanej. W naszym przykładzie λ = 0,67, a więc wartości 600 (o tyle zmniejszyliśmy liczbę błędów) podzielonej przez 900 (tyle błędów popełnilibyśmy, opierając się jedynie na rozkładzie statusu zatrudnienia). Wartość λ możemy zatem traktować jako miarę siły związku między płcią a zatrudnieniem.

Gdyby płeć i zatrudnienie były statystycznie niezależne, to rozkłady statusu zatrudnienia wśród mężczyzn i kobiet byłyby identyczne. W takiej sytuacji znajomość płci poszczególnych osób nie miałaby wpływu na liczbę błędów popełnianych przy przewidywaniu ich statusu zatrudnienia i w efekcie wartość λ wynosiłaby 0. Z kolei, gdyby wszyscy mężczyźni pracowali, a wszystkie kobiety by nie pracowały, znając płeć, nie popełnilibyśmy w takim przewidywaniu żadnego błędu. Oznaczałoby to, że popełniliśmy o 900 błędów mniej (w odniesieniu do 900 popełnionych pierwotnie), a więc uzyskalibyśmy λ = 1, co odpowiada pełnej zależności statystycznej.

Współczynnik λ jest tylko jedną z wielu miar siły związku, stosowaną przy analizowaniu związku między dwiema zmiennymi nominalnymi. O innych takich miarach możesz przeczytać w podręcznikach do statystyki.

Zmienne porządkowe

Jeśli mamy zbadać związek między zmiennymi mierzonymi na skali porządkowej (takimi jak *klasa społeczna*, *religijność*, *alienacja*), możemy użyć współczynnika γ (gamma). Podobnie jak λ, γ odwołuje się do możliwości odgadywania wartości jednej zmiennej na podstawie znajomości wartości innej zmiennej. Jednak podczas gdy λ odwołuje się do zgadywania dokładnej wartości zmiennej, to w przypadku γ odgadujemy uporządkowanie wartości. Dla każdej możliwej pary przypadków przewidujemy, że ich uporządkowanie ze względu na wartości jednej zmiennej jest takie samo (bądź też odwrotne), jak ich uporządkowanie ze względu na wartości drugiej zmiennej.

Wyobraźmy sobie, że mamy do czynienia z grupą uczniów szkoły podstawowej. Właściwa wydaje się hipoteza, że w grupie tej zachodzi związek między wiekiem a wzrostem. Możemy zbadać jej zasadność porównując wszystkie pary uczniów: Jasia i Anię, Jasia i Stasia, Anię i Stasia itd. Pomińmy wszystkie pary,

w których oboje są tego samego wzrostu i/lub są w tym samym wieku. Wszystkie pozostałe pary (czyli pary uczniów, różniących się zarówno pod względem wieku, jak i wzrostu) możemy teraz zaklasyfikować do jednej z dwóch kategorii – par „zgodnych" (gdy dziecko starsze jest wyższe od młodszego) oraz par „niezgodnych" (gdy dziecko starsze jest niższe od młodszego). Tak więc jeśli Jaś jest starszy i wyższy od Ani, para Jaś–Ania będzie liczona jako „zgodna"; gdyby zaś Jaś był starszy, ale niższy od Ani, byłaby to para „niezgodna". (Gdyby zaś byli w tym samym wieku lub tego samego wzrostu, pominęlibyśmy ich).

Aby ocenić, czy między wiekiem a wzrostem zachodzi związek, porównajmy liczbę par zgodnych i niezgodnych. Jeśli par zgodnych jest więcej niż niezgodnych, możemy stwierdzić, że zaobserwowaliśmy *pozytywny* związek między zmiennymi – gdy rośnie wartość jednej, rośnie także wartość drugiej. Jeśli par niezgodnych byłoby więcej, mówilibyśmy o związku *negatywnym*, gdyby zaś obie kategorie par były równoliczne, stwierdzilibyśmy, że wiek i wzrost nie są ze sobą powiązane – że są wzajemnie *niezależne*.

Rozpatrzmy przykład z dziedziny nauk społecznych, ilustrujący zastosowanie współczynnika γ. Wyobraźmy sobie, że oczekujemy, że zachodzi pozytywny związek między religijnością a konserwatyzmem politycznym, a zatem jeśli osoba A jest bardziej religijna od osoby B, to można się spodziewać, że jest od niej także bardziej konserwatywna. Wartość γ równa jest proporcji par, dla których to przewidywanie jest poprawne.

Tabela 17.4 zawiera fikcyjne dane dotyczące związku między klasą społeczną a poziomem uprzedzeń. Ogólny charakter związku między tymi zmiennymi jest taki, że im wyższa jest klasa społeczna, tym niższy jest poziom uprzedzeń – istnieje ujemna zależność między obiema zmiennymi.

Tabela 17.4

Hipotetyczne dane dotyczące związku między klasą społeczną a uprzedzeniami

Poziom uprzedzeń	Klasa niższa	Klasa średnia	Klasa wyższa
Niski	200	400	700
Średni	500	900	400
Wysoki	800	300	100

Do obliczenia współczynnika γ potrzebujemy dwóch wartości: (1) liczby par, w których uporządkowanie ze względu na obie zmienne jest takie samo, oraz (2) liczby par, w których uporządkowanie badanych ze względu na obie zmienne jest odwrotne. Pierwszą z tych liczb wyznaczamy w następujący sposób: dla każdej komórki mnożymy jej liczebność przez sumę liczebności komórek znajdujących się „na prawo i w dół" od niej, po czym sumujemy wszystkie uzyskane iloczyny. Dla tabeli 17.4 liczba par zgodnych wynosi zatem $200(900 + 300 + 400 + 100) + 400(400 + 100) + 500(300 + 100) + 900(100) = 340\,000 + 200\,000 + 200\,000 + 90\,000 = 830\,000$.

Aby wyznaczyć liczbę par niezgodnych, mnożymy liczebność każdej komórki przez sumę liczebności komórek znajdujących się „na lewo i w dół" od niej, po czym sumujemy wszystkie uzyskane iloczyny. Dla danych z tabeli 17.4 uzyskujemy w ten sposób $700(500 + 800 + 900 + 300) + 400(800 + 300) + 400(500 + 800) + 900(800) = 1\,750\,000 + 440\,000 + 520\,000 + 720\,000 = 3\,430\,000$ par niezgodnych. Na tej podstawie możemy wyznaczyć wartość γ jako

$$\gamma = \frac{\text{zgodne} - \text{niezgodne}}{\text{zgodne} + \text{niezgodne}}.$$

W naszym przykładzie $\gamma = (830\,000 - 3\,430\,000)/(830\,000 + 3\,430\,000) = -0{,}61$. To, że wartość współczynnika jest ujemna, potwierdza naszą pierwotną obserwację, że związek między klasą społeczną a poziomem uprzedzeń jest ujemny. Wartość 0,61 z kolei oznacza, że przewaga par niezgodnie uporządkowanych nad zgodnie uporządkowanymi odpowiada 61% liczby wszystkich par (zgodnych i niezgodnych).

Jak pamiętamy, wartość λ mogła wahać się od 0 do 1, natomiast wartość γ może zmieniać się od –1 do 1, dzięki czemu informuje nas zarówno o kierunku, jak i o natężeniu związku. Ponieważ zmienne nominalne nie mają struktury porządkowej, nie ma sensu w ich przypadku mówić o kierunku zależności (ujemna wartość λ oznaczałaby, że znając wartość drugiej zmien-

nej, popełniłeś więcej błędów w przewidywaniu wartości pierwszej zmiennej, niż w przypadku, gdybyś wartości drugiej zmiennej nie znał – a to jest logicznie niemożliwe).

W tabeli 17.5 przedstawiony jest inny przykład zastosowania współczynnika γ w badaniach społecznych. Aby zbadać, w jakim stopniu wdowy idealizują swoich zmarłych mężów, Helena Znaniecki-Lopata (1981) przeprowadziła sondaż na losowej próbie 301 wdów. Jednym z elementów kwestionariusza było pytanie o ocenę zmarłych mężów na następującej skali dyferencjału semantycznego:

Charakterystyka

kraniec pozytywny								kraniec negatywny
dobry	1	2	3	4	5	6	7	zły
użyteczny	1	2	3	4	5	6	7	bezużyteczny
szczery	1	2	3	4	5	6	7	nieszczery
doskonały	1	2	3	4	5	6	7	niedoskonały
łagodny	1	2	3	4	5	6	7	okrutny
przyjazny	1	2	3	4	5	6	7	nieprzyjazny
ciepły	1	2	3	4	5	6	7	zimny

Zadaniem respondentek było ocenienie zmarłego małżonka poprzez zakreślenie jednej cyfry dla każdej z par przeciwstawnych określeń. Zauważ, że ciąg liczb między każdą parą określeń stanowi zmienną porządkową.

Analizując dane, Znaniecki-Lopata chciała określić, jak poszczególne miary są powiązane ze sobą. Jako miernik tego związku zastosowała współczynnik γ – w tabeli 17.5 przedstawione zostały uzyskane przez nią wyniki.

Tabele typu tabeli 17.5 nazywane są *macierzami korelacji*. Dla każdej pary cech, Znaniecki-Lopata obliczyła współczynnik γ – na przykład γ dla cech „dobry" i „użyteczny" wynosi 0,79. Macierz korelacji jest bardzo wygodnym sposobem prezentacji wzajemnych związków między wieloma zmiennymi, toteż będziesz się z nią w literaturze spotykał bardzo często. W tym przypadku wszystkie zmienne są ze sobą dość silnie powiązane, choć dla poszczególnych par związek może być silniejszy niż dla innych.

Także współczynnik γ nie jest jedyną miarą siły związku, stosowaną dla zmiennych porządkowych – w podręcznikach do statystyki znajdziesz bardziej szczegółowe informacje na ten temat.

Zmienne interwałowe i ilorazowe

Gdy chcemy zbadać związek między zmiennymi, mierzonymi na skali interwałowej lub ilorazowej (takimi jak *wiek, dochód, liczba uzyskanych punktów*), jako miary siły związku możemy użyć współczynnika korelacji liniowej Pearsona (*r*). Wyprowadzenie i sposób obliczania tego współczynnika wykracza już poza zakres tej książki, zatem przedstawię tutaj tylko kilka ogólnych uwag.

Tak jak λ i γ, współczynnik *r* odwołuje się do przewidywania wartości jednej zmiennej na podstawie znajomości wartości innej zmiennej. Jednak w przypadku ciągłych zmiennych interwałowych lub ilorazowych nie ma raczej szans na dokładne przewidzenie wartości zmiennej. Z kolei przewidywanie jedynie uporządkowania jednostek obserwacji ze względu na wartości dwóch zmiennych oznaczałoby ignorowanie części informacji niesionej przez zmienne ilorazowe i interwałowe. Można po-

Tabela 17.5

Współczynniki γ dla związku między poszczególnymi pozycjami dyferencjału semantycznego skali idealizacji

	Użyteczny	Szczery	Doskonały	Łagodny	Przyjazny	Ciepły
Dobry	0,79	0,88	0,80	0,90	0,79	0,83
Użyteczny		0,84	0,71	0,77	0,68	0,72
Szczery			0,83	0,89	0,79	0,82
Doskonały				0,78	0,60	0,73
Łagodny					0,88	0,90
Przyjazny						0,90

Źródło: Helena Znaniecki-Lopata, *Widowhood and Husband Sanctification*, „Journal of Marriage and the Family" maj 1981, s. 439–450.

wiedzieć, że w pewnym sensie r określa, jak blisko powinniśmy trafić, przewidując wartość jednej zmiennej na podstawie znajomości wartości drugiej.

Aby zrozumieć logikę, na której opiera się r, zastanówmy się, czym kierujemy się, dokonując przewidywania danej wartości zmiennej dla poszczególnych przypadków. W przypadku zmiennych nominalnych, jak pamiętamy, najlepiej było zawsze przewidywać wartość modalną. Natomiast w przypadku zmiennych interwałowych lub ilorazowych zminimalizujemy liczbę błędów, jeśli za każdym razem jako przewidywaną wartość zmiennej będziemy podawali średnią. Chociaż nieczęsto, jeśli w ogóle, podamy jej dokładną wartość, to jednak w ten sposób zminimalizujemy liczbę popełnianych błędów[2]. Na przykład, jeśli mamy odgadywać dochód poszczególnych badanych, znając ich wykształcenie, możemy podawać za każdym razem średni dochód wśród ludzi o określonym wykształceniu (tzn. średni dochód ludzi, którzy uczyli się przez 0, 1, 2 lata itd.).

Obliczając λ, braliśmy pod uwagę liczbę błędów, które popełniliśmy, podając za każdym razem modalną zmiennej. W przypadku r do oceny wielkości błędu używamy sumy kwadratów różnic między wartościami zmiennej dla poszczególnych przypadków a jej średnią. Suma ta nazywana jest *wariancją całkowitą*.

Aby zrozumieć to pojęcie, musimy dowiedzieć się więcej o analizie regresji, tak byśmy mogli z tej perspektywy przyjrzeć się pojęciu korelacji.

Analiza regresji

Wielokrotnie w dalszym tekście będę odwoływał się do ogólnej postaci opisu związku między zmiennymi, przedstawianej wzorem $Y=f(X)$, co należy czytać „Y jest funkcją X", co oznacza że zróżnicowanie wartości zmiennej Y można wyjaśnić, odwołując się do zróżnicowania wartości zmiennej X. Formułując to dobitniej, możemy twierdzić, że X jest przyczyną Y, a więc wartość X determinuje wartość Y. **Analiza regresji** jest to metoda badania, jaką konkretnie postać ma funkcja przekształcająca X w Y. Stosuje się różne typy analizy regresji, właściwe dla różnych poziomów złożoności badanej zależności. Zacznijmy nasze rozważania od przypadku najprostszego.

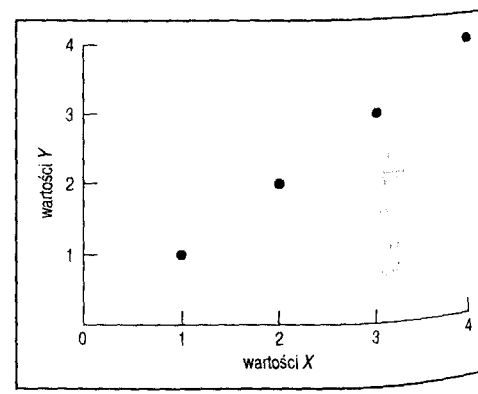
Regresja liniowa

Czym jest model regresji, najłatwiej zobaczyć na przykładzie **analizy regresji liniowej**, czyli ścisłego związku liniowego dwóch zmiennych. Rysunek 17.1 przedstawia w formie graficznej (wykres rozrzutu) wartości zmiennych X i Y, wyznaczonych w jakichś fikcyjnych badaniach. Możemy na nim zobaczyć, że dla każdego z czterech zbadanych przypadków zmienna X przyjmuje taką samą wartość jak Y – dla przypadku, dla którego zmienna X równa jest 1, również zmienna Y przyjmuje wartość 1 itd. Relację między zmiennymi X a Y opisuje równanie $X = Y$, nazywane *równaniem regresji*. Ponieważ wszystkie cztery punkty leżą na jednej prostej, możemy ją wyznaczyć – nazywamy ją *linią regresji*.

Model regresji liniowej ma istotne znaczenie opisowe. Linia regresji może służyć jako graficzne przedstawienie związku między X a Y, a równanie regresji jest efektywną formą syntetycznego opisu tego związku. Modele regresji mają również wartość jako narzędzie przewidywania. W zakresie, w jakim równanie

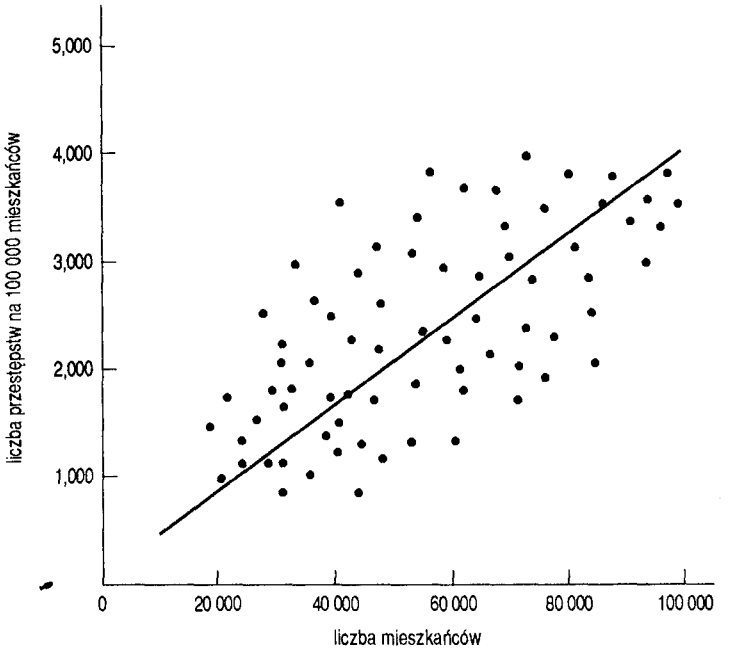
Rysunek 17.1

Prosty wykres rozrzutu wartości zmiennych X i Y



[2] Ściślej rzecz biorąc, zminimalizujemy średni kwadrat błędu przewidywania (przyp. tłum.).

Rysunek 17.2
Hipotetyczny wykres rozrzutu wartości dwóch zmiennych z naniesioną linią regresji



regresji prawidłowo opisuje ogólny związek między dwiema zmiennymi, można go użyć do przewidywania wartości zmiennych dla innych przypadków: na przykład, jeżeli wiemy, że dla jakiegoś przypadku wartość zmiennej X wynosi 3,5, możemy przewidzieć, że wartość zmiennej Y również wynosi 3,5.

Oczywiście w praktyce badania nie ograniczają się zwykle do czterech przypadków, a związki między zmiennymi nie są tak proste, jak ten przedstawiony na rysunku 17.1.

Nieco bardziej realistyczny przykład przedstawiony jest na rysunku 17.2, pokazującym hipotetyczny związek między liczbą mieszkańców a wskaźnikiem przestępczości w małych i średnich miastach. Każdy punkt na wykresie to jedna miejscowość, a jego położenie odpowiada liczbie mieszkańców oraz wskaźnikowi przestępczości (liczba przestępstw na 100 000 mieszkańców). Tak jak w poprzednim przykładzie, wartość Y (poziom przestępczości) jest ogólnie powiązana z wartością X (liczba mieszkańców): w miarę wzrostu X, wzrasta również Y. Zależność ta jednak nie jest nawet w przybliżeniu tak oczywista, jak ta z rysunku 17.1.

Na rysunku 17.2 nie da się wykreślić prostej, która przeszłaby przez wszystkie punkty na wykresie. Możemy jednak wyznaczyć taką linię, która będzie najlepszym możliwym przybliżeniem liniowym poszczególnych punktów, i taka linia jest na nim narysowana.

Jeżeli uczyłeś się kiedyś algebry, to pamiętasz, że każdą prostą na płaszczyźnie można opisać równaniem w postaci $Y = a + bX$, gdzie X i Y to wartości zmiennych. W równaniu tym a równe jest wartości Y, gdy X równe jest 0, natomiast b zależy od nachylenia prostej. Jeśli znamy wartości a i b, możemy wyznaczyć przewidywaną wartość Y dla każdej wartości X.

Możemy teraz stwierdzić bardziej formalnie, że analiza regresji jest techniką, służącą do wyznaczania takiego równania regresji, by opisana przez nie prosta przechodziła najbliżej poszczególnych punktów na wykresie. Równanie regresji z jednej strony daje matematyczny *opis* zależności między zmiennymi, z drugiej

zaś pozwala *wnioskować* o wartości *Y* dla przypadków, dla których znamy *X*. Wracając do przykładu z rysunku 17.2, możemy ocenić poziom przestępczości w różnych miastach, jeśli znamy liczbę ich mieszkańców.

W celu lepszego przewidywania wartości jednej zmiennej na podstawie wartości drugiej zmiennej należy wyznaczyć *linię regresji*, opisaną równaniem regresji w ogólnej postaci $Y' = a + b(X)$, gdzie a i b to odpowiednio wyliczone stałe, X jest znaną wartością jednej zmiennej, Y' zaś jest przewidywaną wartością drugiej zmiennej. Stałe a i b obliczamy tak, by zminimalizować różnice między rzeczywistymi wartościami Y i jej wartościami przewidywanymi na podstawie znanych wartości X, czyli Y'. Suma kwadratów różnic $Y - Y'$ nazywana jest *wariancją niewyjaśnioną*; odpowiada ona błędom przewidywania, które popełniamy pomimo wykorzystania w przewidywaniu znajomości wartości zmiennej X.

Wariancją wyjaśnioną nazywamy różnicę między wariancją całkowitą a wariancją niewyjaśnioną. Analogicznie jak przy obliczaniu współczynnika λ, iloraz wariancji wyjaśnionej i wariancji całkowitej możemy traktować jako miarę *proporcjonalnej redukcji błędu*. Uzyskany współczynnik to r^2, czyli *kwadrat współczynnika korelacji liniowej*. Jeśli zatem $r = 0{,}7$, to $r^2 = 0{,}49$, a więc wyjaśniono około połowy wariancji. W praktyce częściej posługujemy się współczynnikiem r, a nie r^2, ponieważ współczynnik korelacji według momentu iloczynowego przyjmuje, w zależności od kierunku zależności między obiema zmiennymi, wartości dodatnie lub ujemne (gdybyśmy obliczali tylko r^2, po czym wyciągali z niego pierwiastek, zawsze uzyskiwalibyśmy liczbę dodatnią). Sposób wyznaczania r można znaleźć w każdym podręczniku do statystyki, zresztą większość czytelników tej książki ma zapewne dostęp do programów komputerowych z wbudowaną tą funkcją.

Niestety (a może na szczęście), rzeczywistość społeczna jest często zbyt skomplikowana, by realne zjawiska dawały się dobrze opisać za pomocą prostych modeli regresji liniowej. Jak pamiętamy z rozdziału 15, do analizowania trzech lub więcej zmiennych można używać tabel rozkładów łącznych wielu zmiennych, ale wraz ze wzrostem liczby zmiennych stają się one coraz bardziej skomplikowane i coraz mniej czytelne. Właśnie w takich przypadkach użyteczną alternatywą może stać się model regresji.

Regresja wielokrotna

Bardzo często się zdarza, że na zmienną zależną oddziałuje jednocześnie wiele zmiennych niezależnych. W takich sytuacjach zastosowanie znajduje **analiza regresji wielokrotnej**. Przyjrzyjmy się badaniom przeprowadzonym przez Beverly Yerg (1981), dotyczącym efektywności nauczania w wychowaniu fizycznym. Yerg sformułowała swoje oczekiwania w postaci następującego równania regresji wielokrotnej:

$$F = b_0 + b_1 I + b_2 X_1 + b_3 X_2 + b_4 X_3 + b_5 X_4 + e,$$

gdzie:

- F – końcowy poziom osiągnięć ucznia
- I – początkowy poziom osiągnięć ucznia
- X_1 – ocena sposobu prowadzenia zajęć przez nauczyciela
- X_2 – poziom kwalifikacji fachowych nauczyciela
- X_3 – ocena wartości informacji, które uczeń uzyskuje od nauczyciela po wykonaniu zadania
- X_4 – ocena jasności i precyzji, z jaką uczniowi przedstawiane są jego zadania
- b – współczynnik regresji
- e – residuum

(na podstawie Yerg 1981, s. 42)

Zauważ, że zamiast jednej zmiennej X z równania regresji liniowej mamy ich tutaj więcej, więcej jest też współczynników b. Występujący w równaniu współczynnik b_0 pełni analogiczną funkcję, jak współczynnik a z równania regresji liniowej. Wreszcie na końcu równania znajduje się residuum e, reprezentujące tę część wariancji Y, która nie jest wyjaśniana przez uwzględnione w równaniu zmienne X.

Wychodząc od powyższego równania, Yerg obliczyła wartości poszczególnych b, określające względny udział poszczególnych czynników w kształtowaniu ostatecznego poziomu osiągnięć badanych uczniów. Następnie obli-

Miary związku a poziom pomiaru

Peter Nardi

Pitzer College

Tak jak zwykle, w kolumnach odwołujemy się do zmiennych zależnych, w wierszach zaś do zmiennych niezależnych. Zwróć przy tym uwagę, że typy skal pomiaru same w sobie tworzą skalę porządkową.

Jeśli chcesz wykorzystywać tabele do analizy zmiennej interwałowej i ilorazowej, musisz je uprzednio zrekodować na zmienną porządkową.

		zmienna niezależna		
		nominalna	**porządkowa**	**interwałowa/ilorazowa**
zmienna zależna	**nominalna**	*tabele krzyżowe* χ^2 λ	*tabele krzyżowe* χ^2 λ	
	porządkowa	*tabele krzyżowe* χ^2 λ	*tabele krzyżowe* χ^2 λ γ τ Kendalla *d* Somersa	
	interwałowa/ilorazowa	*średnie* test *t* ANOVA (analiza wariancji)	*średnie* test *t* ANOVA (analiza wariancji)	*korelacje* *r* Pearsona regresja (*R*)

czyła współczynnik korelacji wielokrotnej, określający, w jakim stopniu wszystkie sześć uwzględnionych w równaniu zmiennych niezależnych determinuje ostateczny wynik ucznia. Współczynnik ten ma taką samą interpretację jak zwykły współczynnik korelacji liniowej, o którym mówiliśmy wcześniej; w przypadku regresji wielokrotnej oznacza go się dużą literą *R*. Wartość *R* wyznaczona przez Yerg wynosiła 0,877, co oznacza, że 77% (bo $0{,}877^2 = 0{,}77$) wariancji zmiennej *Y* wyjaśniane jest przez łączne działanie sześciu uwzględnionych zmiennych.

Regresja cząstkowa

Omawiając w rozdziale 16 model analizy rozbudowanej, zwracaliśmy szczególną uwagę na zależności między dwiema zmiennymi przy ustalonej wartości trzeciej, kontrolowanej zmiennej. Możemy na przykład badać wpływ wykształcenia na poziom uprzedzeń przy stałej wartości wieku – w ten sposób uzyskując informacje o niezależnym od wieku wpływie wykształcenia. W tym celu analizowalibyśmy tabele zawierające dane o wykształceniu i uprzedzeniach osobno dla różnych grup wiekowych.

Analiza regresji cząstkowej opiera się na tej samej logice – równanie opisujące związek między zmiennymi wyznaczane jest przy założeniu stałej wartości zmiennej kontrolnej. Podobnie jak w przypadku modelu analizy rozbudowanej, dla wyjaśnienia ogólnej zależności, wyniki uzyskane w analizie regresji cząstkowej można następnie porównać z zależnością między dwiema zmiennymi, uzyskaną bez stosowania zmiennej kontrolnej.

Regresja krzywoliniowa

Jak dotąd zajmowaliśmy się takimi związkami między dwiema zmiennymi, które można przedstawić w postaci linii prostej. Model re-

gresji jest jednak ogólniejszy, niż wynikałoby to z naszego dotychczasowego omówienia.

Być może miałeś już do czynienia z funkcjami krzywoliniowymi, które, podobnie jak zwykłe funkcje liniowe, można przedstawiać w postaci odpowiedniego równania. Na przykład równanie $X^2 + Y^2 = 25$ opisuje okrąg o promieniu równym 5. Gdy w równaniu pojawiają się zmienne podniesione do potęgi innej niż 1, na wykresie zamiast linii prostej uzyskamy krzywą. Ponieważ nie wszystkie zachodzące w realnym świecie zależności między zmiennymi mają charakter liniowy, w wielu przypadkach **regresja krzywoliniowa** może stanowić znacznie lepszy opis zależności empirycznych niż modele liniowe.

Jak pamiętamy, linia regresji pełni podwójną funkcję: z jednej strony, stanowi opis zbioru empirycznych obserwacji, z drugiej dostarcza ogólnego modelu dla wnioskowania o zależności między dwiema zmiennymi w populacji generalnej, reprezentowanej przez nasze obserwacje. Formułując odpowiednio złożone równanie, możemy uzyskać bardzo skomplikowaną linię, która faktycznie przejdzie przez wszystkie punkty i w tym sensie będzie stanowić doskonały opis obserwacji empirycznych. Mimo to nie można będzie mieć żadnej pewności, że linia ta pozwalałaby przewidywać trafnie nowe obserwacje, bądź też, żeby w jakikolwiek znaczący sposób opisywała charakter zależności między dwiema analizowanymi zmiennymi w ogóle. Takie rozwiązanie nie miałoby zatem żadnej lub niemal żadnej wartości dla wnioskowania.

Zajmowaliśmy się już wcześniej problemem zrównoważenia szczegółowości i czytelności przy redukcji danych. Niewątpliwie badacze dążą do skonstruowania jak najwierniejszej, ale także jak najprostszej reprezentacji danych empirycznych. Odnosi się to również do analizy regresji. Należy stosować modele możliwie najprostsze (stąd popularność regresji liniowej), które jak najlepiej reprezentują dane empiryczne. Regresja krzywoliniowa stanowi tutaj dla badacza dodatkową możliwość, ale nie zawsze jest rozwiązaniem idealnym – bo też takie w ogóle nie istnieją.

Ograniczenia analizy regresji

Wykorzystanie analizy regresji we wnioskowaniu statystycznym wymaga przyjęcia takich samych założeń, jak przy analizie korelacji: zakładamy, że mamy do czynienia z prostą próbą losową, że dane nie są obciążone błędami o charakterze nielosowym, oraz że analizowane zmienne są interwałowe i ciągłe. Analizowane w ramach badań naukowych dane rzadko spełniają całkowicie te założenia, dlatego do wyników analizy regresji zawsze należy podchodzić z pewną ostrożnością.

Jeśli wyznaczyliśmy równanie regresji (prosto- lub krzywoliniowej), możemy je wykorzystać do *interpolacji* (przewidywania przypadków leżących między dotychczasowymi obserwacjami), znacznie mniej zaufania powinniśmy natomiast mieć do wyników uzyskanych poprzez *ekstrapolację* (przewidywanie wartości zmiennej zależnej dla przypadków poza dotychczas obserwowanym zakresem wartości). Świadomość ograniczenia, dotyczącego ekstrapolacji jest ważna z dwóch względów. Po pierwsze, dokonując jej możemy dojść do absurdalnych wniosków: na przykład odwołując się do równania regresji, dotyczącego poziomu przestępczości ze względu na wielkość miejscowości, oczekiwalibyśmy, że w małych miejscowościach, liczących 1000 mieszkańców, wskaźnik ten wyniósłby 123 przestępstwa na 100 000 mieszkańców rocznie. Nieadekwatność przewidywania w takim przypadku nie dyskwalifikuje analizy regresji – potwierdza jednak zasadę, że uzyskane równanie opisuje zależność jedynie w określonym zakresie liczby ludności. Po drugie, niekiedy badacze formułują wnioski na podstawie ekstrapolacji – w takich sytuacjach istnieją słuszne podstawy, by je kwestionować.

Przedstawiłem pokrótce kilka technik pomiaru siły związku między zmiennymi, odpowiednich dla różnych poziomów pomiaru. Sprawa nieco się komplikuje, gdy chcemy zbadać zależność między dwiema zmiennymi mierzonymi na różnych skalach. Nie będziemy w tej książce zajmowali się tą kwestią, jednak pewne dodatkowe wskazania, co robić w takich sytuacjach, możesz znaleźć w ramce „Miary związku a poziom pomiaru", autorstwa Petera Nardiego.

Inne techniki wielozmiennowe

W książce tej zajmujemy się na ogół dosyć elementarnymi operacjami na danych, takimi jak tworzenie tabel krzyżowych i procentowanie. W tej formie prezentowałem model analizy rozbudowanej i w zasadzie niewiele poza ten zakres wykraczały techniki statystyczne prezentowane dotychczas w tym rozdziale.

Ten podrozdział poświęcę skrótowemu przedstawieniu, właśnie z perspektywy logicznego modelu analizy rozbudowanej, trzech innych technik analizy wielozmiennowej. Nie oczekuj, że po ich lekturze będziesz umiał je stosować – moją intencją jest, byś wiedział, o co chodzi, jeżeli spotkasz się z nimi w czasie lektury raportów badawczych. Trzy techniki, z którymi się zapoznamy – analiza ścieżkowa, analiza szeregów czasowych i analiza czynnikowa – to jedynie kilka spośród wielu technik analizy wielozmiennowej, stosowanych w naukach społecznych.

Analiza ścieżkowa

Analiza ścieżkowa służy do formułowania *modeli przyczynowych*, pozwalających zrozumieć zależności między zmiennymi. Oparta jest na analizie regresji, ale jej szczególną własnością jest to, że jej wyniki prezentowane są w postaci graficznych schematów, ukazujących relacje między wieloma zmiennymi lepiej, niż mogą to uczynić inne narzędzia. Analiza ścieżkowa oparta jest na założeniu, że wartość zmiennej jest zależna przyczynowo od wartości innej zmiennej – dlatego też kluczową rzeczą jest rozróżnienie zmiennych zależnych i niezależnych. Dotyczy to oczywiście także innych typów analiz, ale wyniki analizy ścieżkowej, dzięki szczególnej formie ich prezentacji, są wyjątkowo łatwe w interpretacji.

Wróćmy na moment do sposobu, w jaki przedstawiłem w rozdziale 16 (rysunek 16.1) model analizy rozbudowanej. Logikę jego interpretacji możemy przedstawić za pomocą następującego schematu:

zmienna niezależna	→	zmienna pośrednicząca	→	zmienna zależna

Tak więc zmienna niezależna wpływa na zmienną pośredniczącą, a ta z kolei wpływa na zmienną zależną. Analiza ścieżkowa prowadzi do podobnych schematów związków między zmiennymi, w których jednak uwzględniana jest zwykle znacznie większa liczba zmiennych niż w przedstawionym schemacie.

Oprócz graficznego przedstawienia sieci zależności między zmiennymi analiza ścieżkowa umożliwia także ukazanie siły tych zależności. Obliczane technikami analizy regresji tzw. *współczynniki ścieżkowe* odpowiadają zależnościom cząstkowym w modelu analizy rozbudowanej i mierzą siłę związków między poszczególnymi parami zmiennych przy założeniu stałych wartości pozostałych zmiennych z modelu.

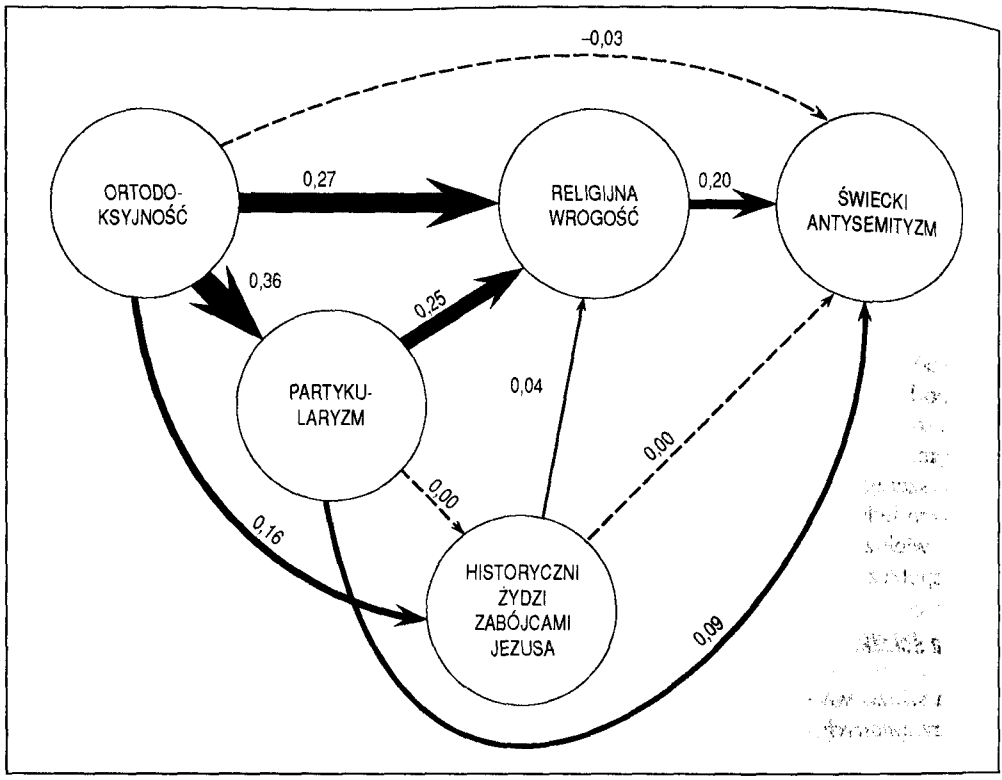
Na przykład przedmiotem analizy przedstawionej na rysunku 17.3 były religijne źródła antysemityzmu wśród chrześcijan. Zmienne uwzględnione w modelu to (na schemacie od lewej do prawej) odpowiednio: (1) ortodoksyjność – stopień, w jakim badany akceptuje tradycyjne poglądy religijne na temat Boga, Jezusa, cudów biblijnych itd.; (2) partykularyzm – przekonanie, że własna religia jest jedyną prawdziwą; (3) akceptacja poglądu, że Żydzi ukrzyżowali Jezusa; (4) religijnie motywowana wrogość wobec współczesnych Żydów – na przykład przekonanie, że zostali ukarani przez Boga, bądź że będą potępieni, dopóki nie nawrócą się na chrześcijaństwo, oraz (5) świecki antysemityzm, przejawiający się na przykład w przekonaniu, że Żydzi są nieuczciwi w biznesie, nie są patriotami itp.

Badacze zaczęli od przyjęcia hipotezy, że świecki antysemityzm generowany jest w wyniku następujących relacji między tymi pięcioma zmiennymi: ortodoksja jest przyczyną partykularyzmu, partykularyzm – traktowania historycznych Żydów jako tych, którzy ukrzyżowali Jezusa, co prowadzi do religijnej wrogości wobec współczesnych Żydów, a ta do świeckiego antysemityzmu.

Model ścieżkowy pokazuje jednak, że zależności te przebiegają inaczej. Na przykład badacze stwierdzili, że tradycyjne poglądy na rolę Żydów w ukrzyżowaniu Jezusa nie mają znaczenia w procesie prowadzącym do antyse-

Rysunek 17.3

Graficzne przedstawienie religijnych źródeł antysemityzmu



Źródło: Rodney Stark, Bruce D. Foster, Charles Y. Glock, Harold E. Quinley, *Wayward Shepherds – Prejudice and the Protestant Clergy*. Copyright © 1971 by Anti-Defamation League of B'nai Brith. Przedruk za pozwoleniem Harper & Row, Publishers, Inc.

mityzmu. Ponadto, jakkolwiek partykularyzm jest częścią procesu prowadzącego do świeckiego antysemityzmu, z diagramu widać, że antysemityzm wynika bardziej bezpośrednio z ortodoksji i religijnej wrogości. Ortodoksyjność prowadzi do wrogości religijnej nawet bez partkularyzmu, a wrogość religijna prowadzi następnie do wrogości świeckiej.

Należy jednak pamiętać, że analiza ścieżkowa, chociaż stanowi doskonałe narzędzie do analizowania łańcuchów zależności przyczynowych oraz struktury powiązań między zmiennymi, sama z siebie nie mówi o kierunkach zależności przyczynowych. Diagram ścieżkowy nie został też wygenerowany przez komputer. To badacz przyjął jakąś strukturę zależności między zmiennymi, a za pomocą analizy ścieżkowej wyznaczył jedynie odpowiednie współczynniki ścieżkowe, znajdujące zastosowanie do danej struktury.

Analiza szeregów czasowych

Różne warianty analizy regresji stosowane są do badania *szeregów czasowych*, czyli danych opisujących zmiany wartości jednej lub wielu zmiennych w czasie. Na pewno wiesz, że w USA poziom przestępczości w ciągu minionych lat generalnie wzrastał. **Analiza szeregów czasowych** poziomu przestępczości pozwala wyznaczyć – w kategoriach analizy regresji – długoterminowy trend oraz weryfikować możliwe jego wyjaśnienia, takie jak wzrost liczby ludności lub wahania sytuacji gospodarczej,

a także pozwala formułować prognozy dotyczące poziomu przestępczości w przyszłości.

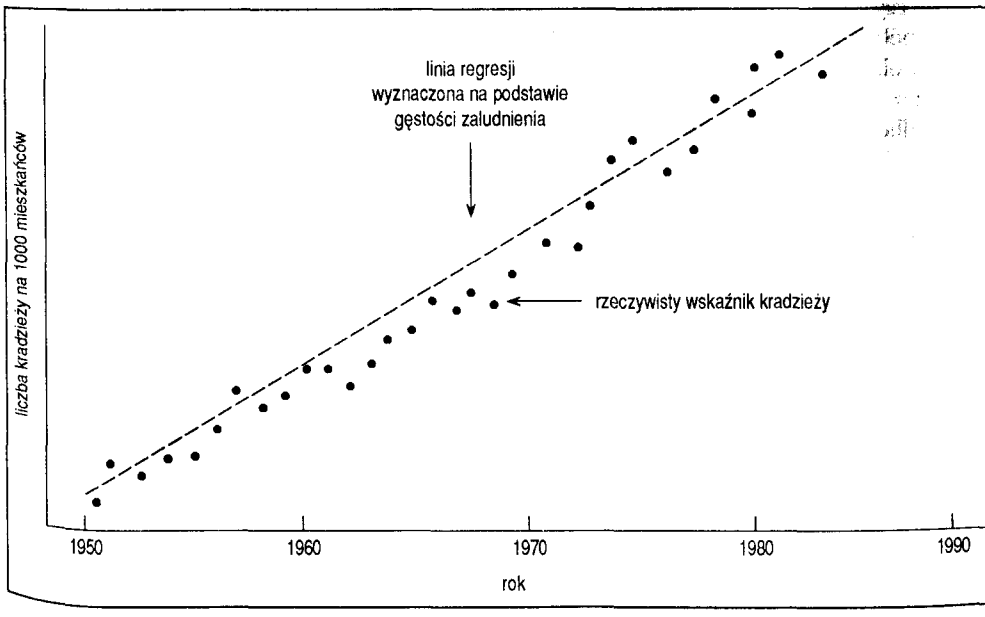
Na rysunku 17.4 przedstawiony jest prosty wykres ilustrujący zmiany wskaźnika kradzieży w czasie w jakimś fikcyjnym mieście. Każdy punkt na wykresie odpowiada liczbie przypadków kradzieży zgłoszonych policji w danym roku.

Załóżmy, że spodziewamy się, że poziom kradzieży jest częściowo funkcją przeludnienia. Może się to opierać na przekonaniu, że nadmierne zagęszczenie ludzi prowadzi do wzrostu poziomu stresu i frustracji, to zaś prowadzi do zwiększenia się częstości różnorodnych przestępstw. Działając podobnie, jak w omawianej uprzednio analizie regresyjnej, możemy wyznaczyć równanie regresji, opisujące zależność między poziomem kradzieży a gęstością zaludnienia, odwołując się do danych o wartościach tych zmiennych, traktując kolejne lata jako jednostki analizy. Dysponując optymalnym równaniem regresji, możemy następnie obliczyć wskaźnik kradzieży dla każdego roku na podstawie ówczesnego współczynnika gęstości zaludnienia. Dla uproszczenia przyjmijmy, że liczba mieszkańców naszego miasta (a zatem i gęstość zaludnienia) w analizowanym okresie wzrastała jednostajnie – przy takim założeniu będziemy przewidywali, że także wskaźnik kradzieży będzie jednostajnie wzrastał. Nasze przewidywania zostały przedstawione na rysunku 17.4 jako przerywana linia regresji.

Zależności opisywane przez szeregi czasowe mogą mieć charakter znacznie bardziej złożony niż prosty przypadek omawiany przed chwilą. Na przykład, możemy rozpatrywać model z większą liczbą zmiennych niezależnych – powiedzmy, że na liczbę kradzieży wpływa także poziom bezrobocia. Moglibyśmy wyznaczyć równanie regresji do przewidywania poziomu kradzieży na podstawie dwóch zmiennych przyczynowych – i mogłoby się okazać, że przewidywania nie układają się już na prostej, ponieważ poziom bezrobocia – inaczej niż gęstość zaludnienia, o której założyliśmy, że w naszym mieście jednostajnie wzrastała – okresowo wzrasta i maleje. Tak więc przewidywany przez nas poziom kradzieży również będzie wahał się w górę i w dół.

Rysunek 17.4

Zmiany w czasie wskaźnika kradzieży w hipotetycznym mieście



Kontynuując analizę zależności między poziomem bezrobocia a liczbą kradzieży, możemy oczekiwać, że ludzie nie będą kraść bezpośrednio po tym, jak stracą pracę – zapewne najpierw będą żyli z oszczędności, potem – ciągle mając nadzieję na znalezienie pracy – będą pożyczali pieniądze od przyjaciół; kraść zaczną dopiero wtedy, gdy nie będą mieli już innych środków utrzymania.

Do rozwiązania takiego bardziej złożonego problemu można zastosować *analizę regresji z przesunięciem czasowym*. Na przykład moglibyśmy przewidywać poziom kradzieży w danym roku za pomocą równania regresji, w którym jedną ze zmiennych niezależnych byłby poziom bezrobocia w poprzednim roku albo średni poziom bezrobocia w ciągu dwóch poprzednich lat; można tu brać pod uwagę bardzo wiele różnych możliwości.

Zwróć uwagę, że w bardzo wielu przypadkach, analizując relacje przyczynowe, należałoby uwzględnić przesunięcie w czasie. Na przykład liczba mieszkańców wielu ubogich krajów utrzymuje się na mniej więcej stałym poziomie, przy wysokim współczynniku urodzeń i jednocześnie wysokiej śmiertelności. Zaobserwowano, że z reguły, gdy współczynnik śmiertelności ulegał obniżeniu, na przykład w wyniku rozwoju opieki medycznej, poprawy warunków sanitarnych czy zwiększenia efektywności rolnictwa, współczynnik urodzeń spadał dopiero po pewnym okresie, a w tym czasie kraj przechodził przez okres gwałtownego wzrostu zaludnienia. Albo zupełnie inny przykład: wzmożenie policyjnych kontroli samochodów przekraczających dozwoloną prędkość powinno spowodować zmniejszenie ich przeciętnej prędkości, ale przyczyna spowoduje skutek dopiero po pewnym czasie – dniach, tygodniach, może miesiącach – dopiero wtedy, gdy kierowcy zaczną sobie zdawać sprawę, że policja rzeczywiście poważnie wzięła się do rzeczy.

W tego rodzaju sytuacjach równania regresji mogą przybierać bardzo różne formy. W każdym przypadku kryterium sukcesu lub porażki modelu jest to, na ile badacz mógł wyjaśnić rzeczywiste wartości zmiennej zależnej, odwołując się do zróżnicowania zmiennych niezależnych.

Analiza czynnikowa

Przykładem zupełnie innego, specyficznego podejścia do problemów analizy wielozmiennowej jest analiza czynnikowa. Ponieważ jej podstawy statystyczne są dość złożone, a przy tym zupełnie inne niż w przypadku technik, o których już mówiliśmy, przedstawię tylko jej ogólne założenia.

Analiza czynnikowa to złożona technika algebraicznego wyszukiwania struktur, stojących za zmiennością wielu zmiennych. Struktury te są opisywane w postaci sztucznie stworzonych wymiarów (*czynników*), silnie skorelowanych z analizowanymi zmiennymi, natomiast niezależnych od siebie nawzajem. Procedura obliczeniowa prowadząca do wyznaczenia tych czynników musi być wykonywana z pomocą komputera.

Wyobraźmy sobie, że zbiór danych statystycznych zawiera wartości kilku wskaźników mierzących poziom uprzedzeń. Każdy ze wskaźników jest z uprzedzeniami jakoś związany, żaden z nich jednak sam z siebie nie mierzy ich precyzyjnie. Możemy się przy tym spodziewać, że poszczególne wskaźniki będą ze sobą silnie skorelowane empirycznie. Stosując analizę czynnikową, stworzylibyśmy sztucznie jeden wymiar (czynnik), silnie skorelowany z poszczególnymi wskaźnikami oraz określilibyśmy pozycję na tym wymiarze (wartość czynnika) dla każdego z badanych; wartość ta powinna pozwolić na trafne przewidywanie zaobserwowanych wartości poszczególnych wskaźników.

Gdyby w tych samych badaniach pojawiły się także wskaźniki odnoszące się do zdolności matematycznych badanych osób, to zapewne dałoby się sztucznie wyznaczyć inny czynnik, silnie skorelowany z poszczególnymi wskaźnikami.

Wydruk generowany przez program robiący analizę czynnikową składa się z wielu kolumn, odpowiadających poszczególnym czynnikom (sztucznie stworzonym wymiarom), wyznaczonym na podstawie zaobserwowanych zależności między zmiennymi, oraz zawierających współczynniki korelacji między czynnikami a poszczególnymi zmiennymi – tzw. *ładunki czynnikowe*.

Tak więc analiza czynnikowa danych z tych przykładowych badań doprowadziłaby zapew-

ne do wyznaczenia dwóch czynników, jednego odpowiadającego mniej więcej poziomowi uprzedzeń i drugiego, odpowiadającego poziomowi zdolności matematycznych. Dane z pytań, mierzących uprzedzenia, miałyby wysokie ładunki dotyczące czynnika uprzedzeń (wysokie współczynniki korelacji z tym czynnikiem), oraz niskie ładunki dotyczące czynnika zdolności matematycznych. W przypadku zmiennych związanych ze zdolnościami matematycznymi struktura danych będzie odwrotna.

W praktyce analizę czynnikową wykonuje się nieco inaczej. Po podaniu programowi, które zmienne mają być analizowane, otrzymujemy wydruk zawierający wiele czynników i odpowiednie ładunki czynnikowe. Zadaniem badacza jest interpretacja uzyskanych czynników na podstawie zmiennych, z którymi poszczególne czynniki są wysoko skorelowane. Komputer, wyznaczając czynniki, nie uwzględnia w żaden sposób sensu poszczególnych zmiennych, a jedynie empiryczne relacje między nimi. Czynniki wyznaczane są na podstawie dwóch kryteriów: (1) czynnik powinien wyjaśniać możliwie największą część łącznej wariancji analizowanych zmiennych, (2) czynniki powinny być wzajemnie niezależne.

Przyjrzyjmy się przykładowi wykorzystania analizy czynnikowej. Wielu badaczy zajmuje się problemem przestępczości, jeśli jednak przyjrzymy mu się dokładniej, to zauważymy, że istnieją różne kategorie przestępców. W swoich badaniach uczniów szkół średnich pewnego małego miasta stanu Wyoming, Morris For-

Tabela 17.6

Analiza czynnikowa: wykroczenia (biali)

Wykroczenie	Wykroczenia przeciwko mieniu	Niezdyscyplinowanie	Narkotyki i wagarowanie	Bójki
	czynnik 1	czynnik 2	czynnik 3	czynnik 4
Tłuczenie latarni ulicznych itp.	0,669	0,126	0,119	0,167
Wybijanie szyb	0,637	0,093	0,077	0,215
Wywracanie płotów itp.	0,621	0,186	0,186	0,186
Kradzież rzeczy o wartości 2-50 USD	0,616	0,187	0,233	0,068
Spuszczanie powietrza z kół	0,587	0,243	0,054	0,156
Kradzież rzeczy o wartości ponad 50 USD	0,548	–0,017	0,276	0,034
Rzucanie jajkami, śmieciami itp.	0,526	0,339	–0,023	0,266
Kradzież rzeczy o wartości do 2 USD	0,486	0,393	0,143	0,077
Kradzież rzeczy w szkole (np. z ławek)	0,464	0,232	–0,002	0,027
Użycie samochodu bez zgody właściciela	0,461	0,172	0,080	0,040
Malowanie grafitti	0,451	0,237	0,071	0,250
Nieposłuszeństwo wobec rodziców	0,054	0,642	0,209	0,039
Robienie napisów na ławkach, ścianach itp.	0,236	0,550	–0,061	0,021
Przeklinanie	0,134	0,537	0,045	0,100
Nieposłuszeństwo wobec nauczycieli i pracowników szkoły	0,240	0,497	0,223	0,195
Bezpośrednie sprzeciwianie się rodzicom	0,232	0,458	0,305	0,058
Wykonywanie anonimowych telefonów	0,373	0,446	0,029	0,135
Palenie marihuany	0,054	0,064	0,755	–0,028
Używanie innych narkotyków	0,137	0,016	0,669	0,004
Fałszowanie pisemnych usprawiedliwień	0,246	0,249	0,395	0,189
Picie alkoholu pod nieobecność rodziców	0,049	0,247	0,358	0,175
Wagarowanie	0,101	0,252	0,319	0,181
Pobicie kogoś	0,309	0,088	0,181	0,843
Udział w bójce	0,242	0,266	0,070	0,602
Procent wariancji	67,2%	13,4%	10,9%	8,4%

Źródło: Morris A. Forslund *Patterns of Delinquency Involvement: An Empirical Typology*, referat przedstawiony na Annual Meeting of the Western Assosation of Sociologists and Anthropologists, Lethbridge, Alberta, 8 lutego 1980; tabela dostosowana na podstawie strony 10.

slund (1980) starał się stworzyć typologię przestępczości. W swoim kwestionariuszu Forslund pytał uczniów, czy kiedykolwiek dopuścili się różnych kategorii działań tego rodzaju, po czym otrzymane odpowiedzi poddał analizie czynnikowej – jej wyniki przedstawione są w tabeli 17.6.

W tabeli różne typy wykroczeń wypisane są z lewej strony, natomiast liczby wewnątrz tabeli to ładunki czynnikowe dla czterech wygenerowanych czynników. Jak możesz zauważyć, Forslund po zbadaniu poszczególnych wymiarów (czynników) nadał każdemu z nich nazwę; w tabeli zaznaczyłem dla każdego z nich te pozycje, które zadecydowały o jego interpretacji. Forslund w ten sposób podsumował wyniki swoich badań:

> „Wyniki uzyskane na całej próbie pozwalają na wyodrębnienie czterech, zasadniczo różnych wzorców wykroczeń. W kolejności wielkości wyjaśnianej wariancji zostały one oznaczone jako: (1) wykroczenia przeciwko mieniu, wliczając w to zarówno akty wandalizmu, jak i kradzieże, (2) niezdyscyplinowanie; (3) narkotyki i wagarowanie; (4) bójki. Interesujący, a być może także zaskakujący jest fakt, że akty wandalizmu i kradzieże wchodzą do tego samego czynnika – wynika z tego, że ci uczniowie, którzy dokonywali wykroczeń przeciwko mieniu, dopuszczali się zarówno aktów wandalizmu, jak i kradzieży. Warto także zauważyć, że do jednego czynnika wchodzą używanie narkotyków, picie alkoholu i wagarowanie".
>
> (1980, s. 4)

Po wyznaczeniu struktury dla całej próby Forslund powtórzył analizę czynnikową osobno dla chłopców i dziewcząt – w obu przypadkach uzyskując zasadniczo taką samą strukturę.

Jak z powyższego przykładu widać, analiza czynnikowa stanowi efektywne narzędzie do wyszukiwania istotnych struktur organizujących większą liczbę zmiennych. Stosując analizę czynnikową, nie trzeba przy tym wykonywać niekończących się analiz korelacyjnych – liniowych, wielokrotnych, cząstkowych – aby odkryć strukturę zbioru zmiennych. Przy okazji jest to dobry przykład efektywnego wykorzystania komputera.

Dodatkową korzyścią z analizy czynnikowej jest możliwość prezentacji danych w postaci łatwej do interpretacji dla badacza lub czytelnika. Dla każdego z czynników łatwo jest wskazać zmienne z wysokimi ładunkami czynnikowymi i w ten sposób wykrywać całe grupy zmiennych powiązanych ze sobą. Podobnie czytelnik łatwo może się zorientować, z którymi czynnikami jest, a z którymi nie jest powiązana dana zmienna.

Ale analiza czynnikowa ma także swoje wady. Po pierwsze, jak już wspominałem, przy generowaniu czynników nie jest w żaden sposób uwzględniany merytoryczny sens poszczególnych zmiennych. Robiąc analizę czynnikową, często napotykamy czynniki z bardzo wysokimi ładunkami czynnikowymi dla zmiennych, nie mających ze sobą nic wspólnego. Na przykład poziom uprzedzeń i religijność mogą mieć wysokie dodatnie ładunki czynnikowe, a poziom wykształcenia wysoki ujemny ładunek czynnikowy dla jednego czynnika. Niewątpliwie, te trzy zmienne mogą być silnie skorelowane, ale jakiemu realnemu zjawisku odpowiadać ma taki czynnik? Często niestety się zdarza, że niedoświadczeni badacze odkrywają w ten sposób „religijno-uprzedzeniowy brak wykształcenia" albo coś równie nonsensownego.

Z drugiej strony analiza czynnikowa krytykowana jest także na gruncie filozoficznym. Jak pamiętamy, poprawna hipoteza musi być falsyfikowalna – jeśli badacz nie potrafi wskazać, jakie przyszłe wyniki mogłyby wykazać jej fałszywość, w istocie jest ona tautologiczna albo bezużyteczna. W pewnym sensie jest to problem analizy czynnikowej: niezależnie od tego, co zawierają dane, analiza czynnikowa zawsze daje rozwiązanie w postaci czynników. Tak więc jeśli badacz stawia pytanie: „Czy za moimi danymi stoi jakaś struktura?", zawsze otrzyma odpowiedź: „tak". Fakt ten trzeba brać pod uwagę, oceniając wartość wyników analizy czynnikowej. Z samego faktu, że otrzymana została struktura czynnikowa, nie wynika jeszcze, że ma ona jakiekolwiek znaczenie.

Moja osobista opinia o analizie czynnikowej jest taka sama, jak o innych złożonych technikach analitycznych. Może być ona bardzo użyteczna dla badań społecznych i warto ją sto-

sować wszędzie tam, gdzie ułatwia badaczowi interpretowanie danych. Trzeba jednak pamiętać, że jest to jedynie narzędzie, a nie magiczny środek na wszystkie problemy badawcze.

Na zakończenie powtórzę jeszcze raz – techniki analityczne, o których w skrócie mówiliśmy, to tylko kilka spośród wielu, stosowanych powszechnie w naukach społecznych. W toku dalszych studiów będziesz zapewne miał jeszcze okazję bardziej szczegółowego ich poznania.

Wnioskowanie statystyczne

Większość naukowych badań społecznych zakłada analizę danych zebranych na próbie wylosowanej z większej populacji. Na próbie ludzi mogą być realizowane wywiady kwestionariuszowe w badaniach sondażowych, kodowaniu i analizie może zostać poddana próba akt spraw rozwodowych, na próbie gazet można przeprowadzać badania metodą analizy treści. Badań na próbie raczej nie przeprowadza się po to, by wypowiadać się o próbie jako takiej – celem badacza jest raczej sformułowanie wniosków dotyczących całej populacji, z której próba została dobrana. Wiąże się to z koniecznością interpretowania wyników analiz jedno- i wielozmiennowych za pomocą wnioskowania statystycznego na temat danej populacji. W tym podrozdziale zajmiemy się **wnioskowaniem statystycznym** – narzędziami statystycznymi stosowanymi do uogólniania obserwacji z próby na szerszą populację. Zaczniemy od analiz jednozmiennowych, po nich zaś przejdziemy do technik związanych z analizami wielozmiennowymi.

Wnioskowanie jednozmiennowe

Na początku rozdziału 15 zajmowaliśmy się sposobami prezentacji jednej zmiennej. Używaliśmy do tego różnorodnych miar, mających syntetycznie opisywać przebadaną próbę. Teraz zastosujemy analogiczne miary do formułowania szerszych twierdzeń, dotyczących całej populacji. Zajmiemy się tutaj dwiema miarami: częstościami oraz średnią.

Jeśli 50% badanych w próbie stwierdziło, że w minionym roku przechodziło przeziębienie, to także na 50% będziemy szacowali odsetek przechodzących przeziębienie w całej populacji, z której dobrana była próba (oczywiście przyjmujemy założenie, że mamy do czynienia z prostą próbą losową). Nie powinniśmy się spodziewać, że rzeczywiście dokładnie 50% populacji przeziębiło się w zeszłym roku, ale, jeżeli próba została dobrana zgodnie z odpowiednim schematem doboru, możemy oszacować oczekiwany zakres błędu, który moglibyśmy popełnić uogólniając nasz wynik z próby na całą populację.

Gdy w rozdziale 7 zajmowaliśmy się teorią doboru próby, podałem sposób dokonywania takiego oszacowania, który teraz tylko przypomnę. W przypadku odsetków *błąd standardowy* wyznacza się za pomocą wzoru

$$\sqrt{\frac{p \times q}{n}},$$

gdzie p jest proporcją, $q = 1-p$, n zaś jest wielkością próby. Jak wspominałem w rozdziale 7, wartość ta jest bardzo istotna przy estymowaniu błędu z próby. Na 68% możemy być pewni, że rzeczywista wielkość w populacji nie różni się od wyniku w próbie o więcej niż plus minus jeden błąd standardowy, na 95%, że nie różni się bardziej niż o plus minus dwa błędy standardowe oraz na 99,9%, że nie różni się o więcej niż plus minus trzy błędy standardowe.

Tak więc każde określenie błędu z próby (błędu losowego) musi zawierać dwie liczby: *poziom ufności* (na przykład 95%) oraz *przedział ufności* (na przykład 62,5%). Jeśli 50% badanych w 1600-osobowej próbie zadeklaruje, że przechodziło w ubiegłym roku przeziębienie, możemy powiedzieć, że z 95-procentową pewnością odsetek osób przechodzących przeziębienie w ubiegłym roku w całej populacji mieści się w przedziale od 47,5% do 52,5%.

W powyższym przykładzie przeszliśmy od prostego opisywania próby do oszacowywania parametru zmiennej w całej populacji (czyli wnioskowanie statystyczne). Przeprowadzając takie wnioskowania, musimy jednak zachować w kilku punktach szczególną ostrożność.

Po pierwsze, możemy wnioskować tylko o tej populacji, z której nasza próba była dobrana. Na przykład na podstawie próby wylosowanej z książki telefonicznej nie możemy formułować wniosków o populacji wszystkich mieszkańców miasta, a jedynie o populacji posiadaczy telefonów, którzy nie zastrzegli swoich numerów.

Po drugie, przy wnioskowaniu statystycznym przyjmuje się pewne założenia. Zakłada się, że próba jest próbą losową prostą – co w praktyce badawczej prawie nigdy się nie zdarza. Co więcej, zakłada się przy tym, że było to – co również prawie nigdy się nie zdarza – losowanie ze zwracaniem; to jednak ma akurat niewielkie znaczenie. Chociaż częściej niż losowanie proste stosuje się losowanie systematyczne, jeśli zostało ono przeprowadzone poprawnie, to nie powinno to prowadzić do istotnych problemów. Warstwowy schemat losowania – poprawiający reprezentatywność – tym bardziej problemu nie stanowi. Inaczej jest jednak z próbą grupową – w jej przypadku oszacowanie błędu z próby może być zaniżone. Oczywiście w przypadku „próby ulicznej" w ogóle nie możemy mówić o wnioskowaniu statystycznym. Trzeba jeszcze wziąć pod uwagę, że wyznaczając błąd standardowy zakładamy, że próba została zrealizowana w 100%, tzn. że wszystkie osoby wylosowane do próby wzięły udział w badaniach. Im niższy jest poziom realizacji próby, tym poważniejszy staje się problem.

Po trzecie, wnioskowanie statystyczne uwzględnia tylko błędy losowe (z próby), a nie **błędy systematyczne**, takie jak błędy kodowania lub wynikające z niezrozumienia pytań przez część badanych. Tak więc, choć w pełni prawidłowe byłoby zdanie, że od 47,5 do 52,5% populacji (95% ufności) *stwierdziło,* że przechodziło przeziębienie w zeszłym roku, nie możemy na tym samym poziomie ufności twierdzić, że odsetek tych, którzy rzeczywiście w zeszłym roku *byli* przeziębieni, mieści się w tym przedziale. Musimy liczyć się z tym, że jeśli stosowaliśmy właściwy schemat losowania próby, błędy systematyczne będą zapewne znacznie większe niż błędy losowe, tak więc uogólniając wyniki z próby na populację, musimy rzeczywiście zachować szczególną ostrożność.

Testy istotności statystycznej

Nie ma „naukowej" odpowiedzi na pytanie, czy obserwowany związek między dwiema zmiennymi jest istotny, silny, ważny, interesujący bądź w ogóle wart podania do publicznej wiadomości. W pewnym sensie ostatecznym „testem istotności" jest to, czy badacz potrafi przekonać swoich czytelników – obecnych lub przyszłych – że znaleziona przez niego zależność rzeczywiście ma jakieś znaczenie. Z drugiej jednak strony, istnieje cały dział wnioskowania statystycznego, zajmujący się *parametrycznymi testami istotności.* Słowo „parametryczne" użyte w nazwie odnosi się do tego, że stosując je przyjmuje się pewne założenia co do parametrów opisujących populację, z której została wylosowana próba. Stosując testy istotności, możemy określić **istotność statystyczną** związku między zmiennymi. Z „istotności statystycznej" nie wynika w żaden sposób „istotność" czy „ważność" w jakimś szerszym sensie – odnosi się ona jedynie do prawdopodobieństwa, że zależności, zaobserwowane w próbie, mogłyby być spowodowane w rzeczywistości jedynie przez błąd z próby.

Chociaż literatura naukowa powołuje się na **testy istotności statystycznej** bardzo często, ich złożona logika często jest niewłaściwie rozumiana. Testy istotności odwołują się do teorii doboru próby, o której już była w tej książce mowa. Aby zrozumieć ich konstrukcję, wróćmy jeszcze na chwilę do koncepcji błędu z próby w odniesieniu do jednej zmiennej.

Jak pamiętamy, wartość statystyki w próbie jest z reguły najlepszym estymatorem odpowiedniego parametru w populacji, ale rzadko kiedy są one sobie dokładnie równe. Tak więc naszą estymację przedstawialiśmy podając prawdopodobieństwo, że parametr w populacji należy do określonego przedziału (przedziału ufności). Stopień niepewności co do dokładnej wartości parametru w ramach tego przedziału wiąże się z wielkością błędu z próby, dlatego bardzo mało prawdopodobne jest, aby parametr w populacji nie mieścił się w podanym przedziale ufności *wyłącznie* ze względu na błąd z próby. Tak więc, jeśli na poziomie ufności 99,9% stwierdzimy, że parametr mieści się w przedziale 45–55%, to wynika z tego, że jest niesłychanie mało praw-

dopodobne, że jego wartość w rzeczywistości wynosi 90%, jeżeli błąd estymacji miałby mieć *jedynie* losowy charakter. Na takiej podstawowej logice opierają się testy istotności statystycznej.

Logika istotności statystycznej

Spróbuję przedstawić logikę istotności statystycznej za pomocą serii diagramów, ilustrujących dobór próby z populacji. Uwzględnimy przy tym następujące elementy:

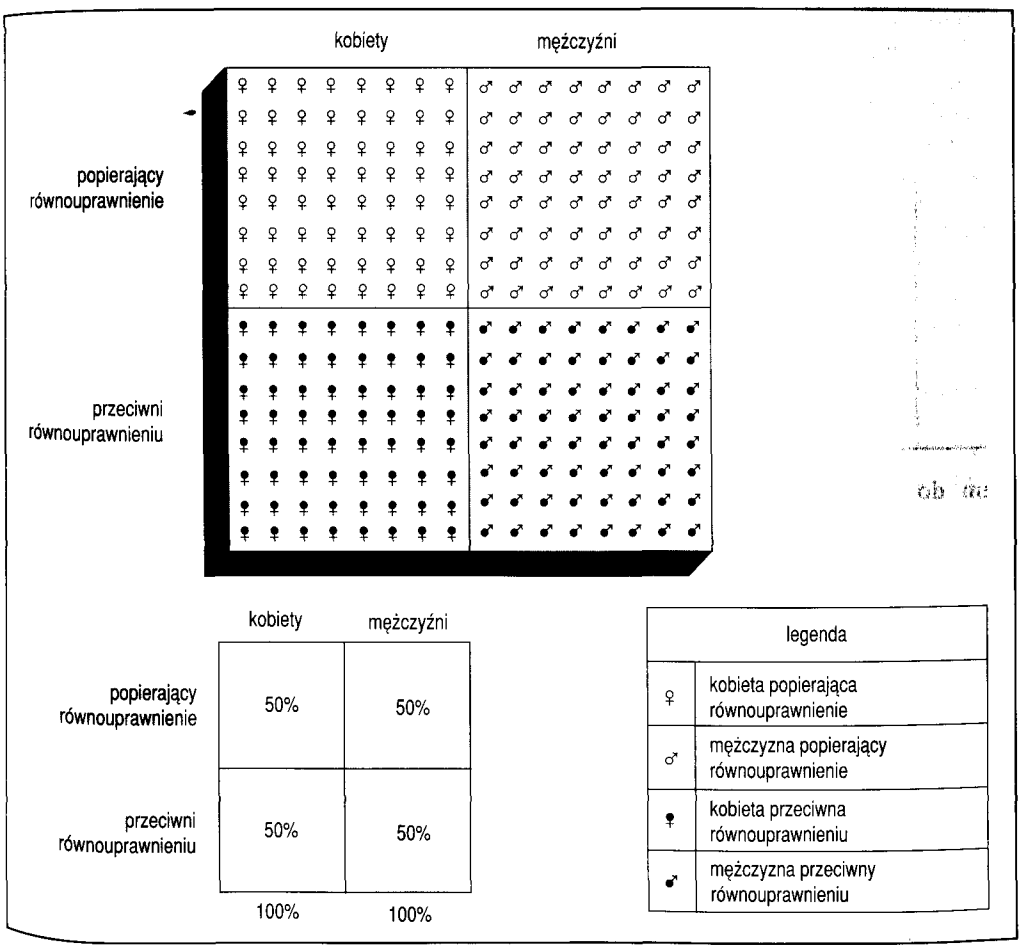
1. Zakładamy, że w badanej populacji dwie analizowane zmienne są od siebie niezależne.
2. Zakładamy, że próba wylosowana z populacji zgodnie ze standardowymi procedurami jest reprezentatywna.
3. Znamy łączny rozkład w próbie dwóch analizowanych zmiennych.

Rysunek 17.5 przedstawia hipotetyczną populację 256 osób, składającą się po połowie z kobiet i mężczyzn. Na rysunku zaznaczone jest także, jaki jest stosunek każdej osoby do równouprawnienia płci (osoby popierające równouprawnienie reprezentowane są przez puste kółka, a nie popierające równouprawnienia przez wypełnione kółka).

Problem, który będziemy analizować, to to, czy istnieje związek między płcią a stosun-

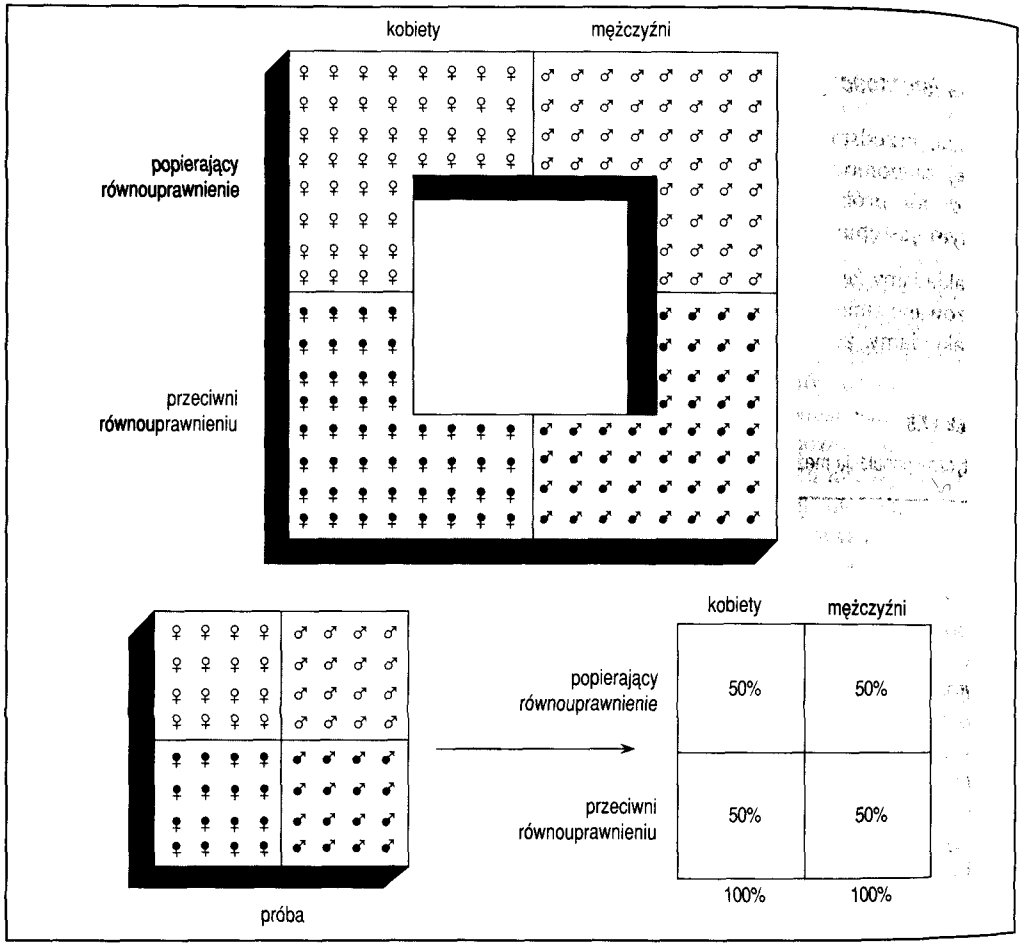
Rysunek 17.5

Hipotetyczna populacja mężczyzn i kobiet popierających równouprawnienie płci lub przeciwnych mu



Rysunek 17.6

Próba reprezentatywna



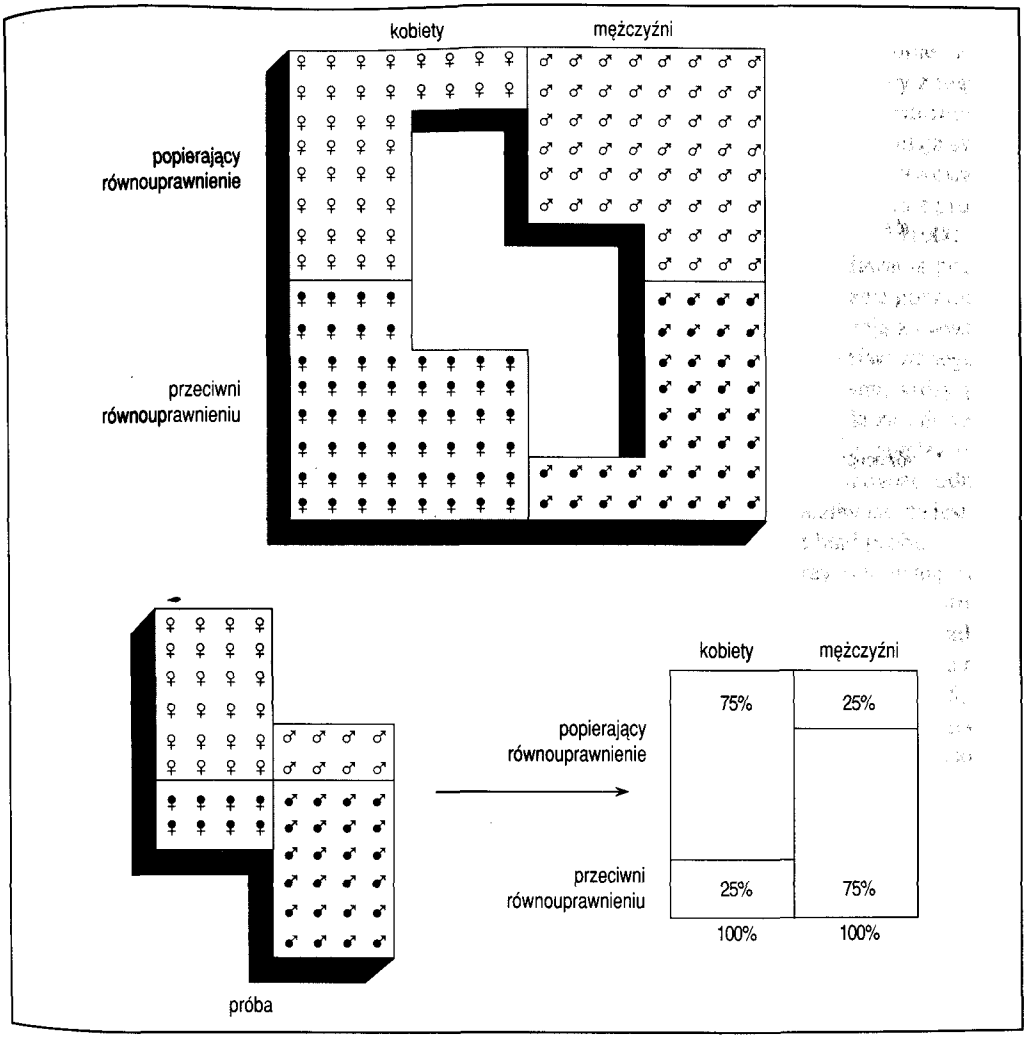
kiem do równouprawnienia. Ściślej rzecz biorąc, będziemy chcieli sprawdzić, czy kobiety, skoro powinny na tym więcej zyskać, popierają równouprawnienie częściej niż mężczyźni. Rzuć okiem na rysunek 17.5, żeby znaleźć odpowiedź na powyższe pytanie.

Jak wynika z rysunku, między płcią a stosunkiem do równouprawnienia żaden związek nie zachodzi. Dokładnie połowa osób z każdej grupy popiera równouprawnienie, a połowa jest mu przeciwna. Nawiązując do tego, co mówiliśmy o proporcjonalnej redukcji błędu, w naszym przypadku znajomość płci nie zredukowałaby liczby błędów przy przewidywaniu stosunku do równouprawnienia. Tabela na dole rysunku 17.5 w innej formie pokazuje to samo co rysunek.

Na rysunku 17.6 przedstawiona jest próba, licząca 1/4 całej populacji. Posługując się dalej „graficzną" terminologią, wybranie małego kwadratu ze środka figury, reprezentującej całą populację, przedstawia wyłonienie reprezentatywnej próby. Jak widzimy, w próbie mamy po 16 elementów należących do każdego z typów; połowa osób w próbie to kobiety, połowa to mężczyźni; po połowie osób z każdej z płci to zwolennicy równouprawnienia, a druga połowa to jego przeciwnicy.

Na podstawie badania próby takiej, jak ta z rysunku 17.6, będziemy mogli formułować po-

Rysunek 17.7
Próba niereprezentatywna



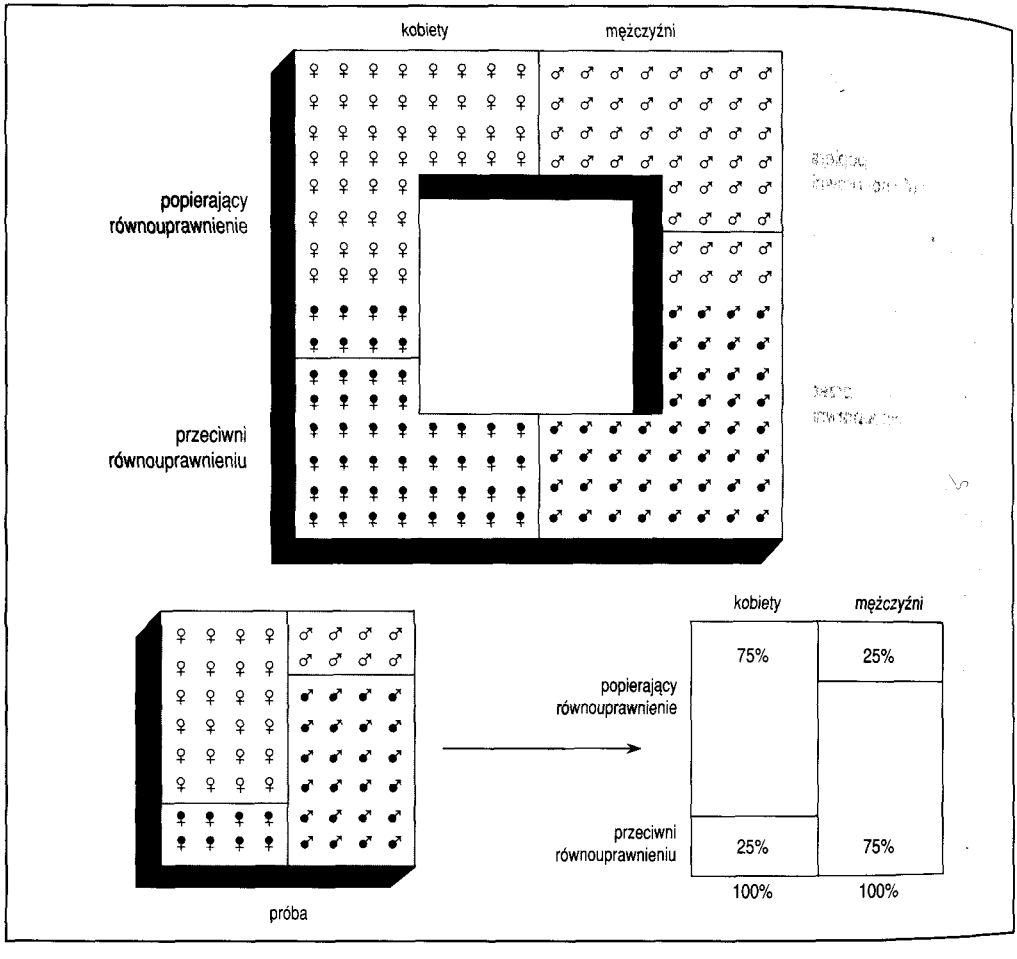
prawne wnioski, odnoszące się do zależności między płcią a stosunkiem do równouprawnienia w całej populacji w ogóle. Opierając się na logice doboru próby, którą znasz z rozdziału 7, będziemy mogli stwierdzić, że w próbie nie zachodzi związek między płcią a stosunkiem do równouprawnienia, a zatem, wobec faktu, że wyłoniliśmy próbę w sposób zgodny z zasadami, możemy wnioskować, że związek między tymi zmiennymi nie zachodzi także w populacji.

Oczywiście w praktyce próba rzadko kiedy stanowi tak idealne odzwierciedlenie populacji. Nie byłoby nic dziwnego, gdybyśmy z populacji, w której nie ma zależności między zmiennymi, przypadkowo wylosowali dodatkowo jednego czy dwóch mężczyzn przeciwnych równouprawnieniu, albo kilka więcej kobiet – zwolenniczek równości płci. Takie niewielkie różnice są w próbach losowych zupełnie normalne, jak widzieliśmy w rozdziale 7.

Na rysunku 17.7 mamy jednak do czynienia z próbą, która odbiega od populacji znacznie bardziej, niż można by zaakceptować: zawiera zdecydowanie za dużo kobiet – zwolenniczek

Rysunek 17.8

Reprezentatywna próba wylosowana z populacji, w której zmienne są zależne



równouprawnienia i przeciwnych mu mężczyzn. Jak widać z tabeli, aż trzy czwarte kobiet i zaledwie jedna czwarta mężczyzn w próbie popiera równouprawnienie. Gdybyśmy wylosowali taką próbę z populacji, w której te dwie zmienne są niezależne, to skończyłoby się to sformułowaniem przez nas wniosków, mających bardzo niewiele wspólnego z rzeczywistością.

Jak pamiętasz, jest bardzo mało prawdopodobne, by prawidłowo wylosowana próba była tak różna od swojej populacji, jak próba z rysunku 17.7. Gdybyśmy rzeczywiście wylosowali taką próbę, to wyjaśnienia nie szukalibyśmy w błędach losowych, ale gdzie indziej, na przykład przyjmując, że sytuacja jest taka, jak przedstawiono na rysunku 17.8.

Jak widzimy, w próbie pokazanej na rysunku 17.8 również mamy do czynienia z silną zależnością między płcią a stosunkiem do równouprawnienia, ale powodem nie jest niereprezentatywność próby – nasza próba po prostu została wylosowana z populacji, w której faktycznie między tymi zmiennymi zachodzi silny związek. Na tym rysunku przedstawiona jest populacja, w której kobiety są znacznie częściej zwolenniczkami równouprawnienia płci niż mężczyźni; próba zaś odzwierciedla sytuację w populacji.

W praktyce oczywiście nigdy nie wiemy, co naprawdę dzieje się w populacji generalnej –

gdybyśmy wiedzieli, to nie mielibyśmy po co robić badań na próbie. Tak więc, jeśli wylosowalibyśmy próbę, w której zależność między zmiennymi byłaby *tak silna* jak w próbach z rysunków 17.7 i 17.8, musielibyśmy zdecydować, czy uznajemy, że zależność ta odpowiada zależności *zachodzącej* w populacji, czy też jest jedynie wynikiem błędu z próby.

Stosując testy istotności statystycznej, odwołujemy się do następującego rozumowania: jeśli stwierdzamy rozbieżność między założoną uprzednio niezależnością dwóch zmiennych w populacji a ich zaobserwowanym rozkładem w próbie, możemy ją wytłumaczyć na dwa sposoby. Albo (1) wynika ona z niereprezentatywności próby, albo (2) założenie o niezależności zmiennych w populacji można odrzucić. Teoria doboru losowego oraz opracowane w jej ramach miary pozwalają ocenić prawdopodobieństwo *uzyskania próby* o określonym stopniu niereprezentatywności (wyrażonej w kategoriach błędu z próby). Najprościej mówiąc, prawdopodobieństwo, że próba jest w niewielkim stopniu niereprezentatywna, jest *duże*, natomiast prawdopodobieństwo, że próba jest niereprezentatywna w dużym stopniu, jest *małe*.

Istotność statystyczna obserwowanego w próbie związku między zmiennymi jest zatem zawsze wyrażana w kategoriach prawdopodobieństwa. „Istotny na poziomie 0,05 ($p \leq 0,05$)" znaczy tyle, że związek zaobserwowany w próbie jest na tyle silny, iż prawdopodobieństwo, że został on w próbie wywołany jedynie przez błąd losowy, jest mniejsze niż 5%. Albo ujmując rzecz inaczej: jeśli w populacji obie badane zmienne są niezależne i z populacji tej wyłonilibyśmy 100 prób losowych, to w nie więcej niż pięciu zaobserwowalibyśmy zależność między zmiennymi tak silną jak ta, którą obserwujemy.

Tak więc **poziom istotności** wyznaczony za pomocą testu istotności można interpretować jako prawdopodobieństwo, że wykryty w próbie związek między zmiennymi powodowany jest *wyłącznie* przez błąd z próby. Podobnie jak przedziały ufności, poziomy istotności wyznaczane są na podstawie modelu logicznego, w którym rozpatruje się wiele prób losowanych z tej samej populacji. W naszym przypadku zakładamy, że w populacji tej dwie zmienne są od siebie niezależne, po czym obliczamy, w jakim odsetku prób wylosowanych z takiej populacji zaobserwowalibyśmy związek między zmiennymi co najmniej tak silny, jak ten, który stwierdziliśmy w próbie. Przy prezentacji wyników badań naukowcy z reguły posługują się trzema poziomami istotności: 0,05, 0,01 oraz 0,001, które odpowiadają sytuacjom, w których ryzyko, że zaobserwowana zależność jest wynikiem jedynie błędu z próby, wynosi odpowiednio 5/100, 1/100 i 1/1000.

Istnieją dwie konwencje podawania poziomów istotności. Niektórzy badacze z góry określają, jaki poziom istotności uważają za wystarczający, i każdy przypadek zaobserwowanego w próbie związku między zmiennymi, który jest istotny na takim poziomie, uznają za odpowiadający rzeczywistemu związkowi między tymi zmiennymi w populacji; inaczej mówiąc, odrzucają w takich przypadkach możliwość, że jest on spowodowany jedynie przez błąd próby.

Są także badacze, którzy odrzucają zwyczaj stosowania wyłącznie poziomów istotności takich jak 0,05, 0,01 czy 0,001 i za każdym razem, gdy mierzą siłę związku, podają dla niej dokładnie wyznaczony poziom istotności, na przykład 0,023 (czyli że prawdopodobieństwo, że zaobserwowany związek ma jedynie losowy charakter, wynosi 23/1000).

Test chi-kwadrat

Test chi-kwadrat (χ^2) jest w naukach społecznych jednym z częściej używanych testów istotności. Wykonując go, odwołujemy się do pojęcia **hipotezy zerowej**, zakładającej, że w populacji dwie zmienne są niezależne (patrz rozdział 2). Uwzględniając empiryczne rozkłady liczebności wartości obu – traktowanych osobno – zmiennych, obliczamy oczekiwany rozkład liczebności dla tych zmiennych, wynikający z hipotezy o ich niezależności. W ten sposób dla każdej komórki rozkładu łącznego w tabeli krzyżowej możemy określić *liczebność oczekiwaną*, którą możemy porównać z liczebnością empiryczną, zaobserwowaną w próbie. Następnie określamy prawdopodobieństwo, z jakim różnice między liczebnościami oczekiwanymi a zaobserwowanymi mogły wyniknąć jedynie z błędów losowych.

Tabela 17.7

Przykład zastosowania testu chi-kwadrat

I. Liczebności oczekiwane	mężczyźni	kobiety	razem
brał(a) udział w nabożeństwie	28	42	70
nie brał(a) udziału w nabożeństwie	12	18	30
łącznie	40	60	100
II. Liczebności zaobserwowane	mężczyźni	kobiety	razem
brał(a) udział w nabożeństwie	20	50	70
nie brał(a) udziału w nabożeństwie	20	10	30
łącznie	40	60	100
III. (Zaobserwowane – oczekiwane)^2/ oczekiwane	mężczyźni	kobiety	
brał(a) udział w nabożeństwie	2,29	1,52	$\chi^2 = 12{,}70$
nie brał(a) udziału w nabożeństwie	5,33	3,56	$p < 0{,}001$

Przyjrzyjmy się tej procedurze na przykładzie. Załóżmy, że badamy zależność między płcią a uczestnictwem w praktykach religijnych członków jakiegoś wyznania i w tym celu losujemy 100-osobową próbę. Wyobraźmy sobie, że do naszej próby weszło 40 mężczyzn i 60 kobiet, oraz że 70% badanych zadeklarowało, że w ciągu poprzedniego tygodnia brało udział w nabożeństwie, 30% zaś zadeklarowało, że w tym okresie nie uczestniczyło w praktykach religijnych.

Jeśli między płcią a praktykami religijnymi nie zachodzi związek, zarówno wśród mężczyzn, jak i wśród kobiet w próbie odsetek uczestniczących w nabożeństwie powinien wynosić po 70%, przy 30-procentowym udziale w obu grupach osób, które w nabożeństwie nie brały udziału. W tabeli 17.7 (część I) zapisaliśmy, że zgodnie z modelem w nabożeństwie powinno brać udział 28 mężczyzn i 42 kobiety, 12 mężczyzn i 18 kobiet zaś w nabożeństwie by nie uczestniczyło.

W części II tabeli 17.7 podane są empiryczne liczebności w 100-osobowej próbie: 20 mężczyzn zadeklarowało, że wzięło udział w nabożeństwie, 20 zaś zadeklarowało, że w nabożeństwie nie uczestniczyło, natomiast spośród kobiet 50 stwierdziło, że uczestniczyło, 10 zaś, że nie. Porównując liczebności oczekiwane i zaobserwowane (dane z części I i II), możemy stwierdzić, że w nabożeństwie wzięło udział nieco mniej mężczyzn i nieco więcej kobiet niż się spodziewaliśmy.

Wartość statystyki chi-kwadrat oblicza się w następujący sposób: dla każdej komórki tabeli (1) odejmujemy wartość oczekiwaną od wartości zaobserwowanej, (2) uzyskaną wartość podnosimy do kwadratu, po czym (3) kwadrat różnicy dzielimy przez liczebność oczekiwaną. Następnie wartości uzyskane dla poszczególnych komórek dodajemy do siebie – i w ten sposób wyznaczamy wartość chi-kwadrat – w naszym przykładzie (patrz część III tabeli 17.7) wynosi ona 12,70.

Wartość chi-kwadrat jest miarą różnicy między zaobserwowanym rozkładem łącznym analizowanych zmiennych w próbie a rozkładem, który wynikałby z założenia, że zmienne te są od siebie niezależne. Sam fakt, że taka różnica istnieje, nie oznacza jeszcze oczywiście, że zmienne są ze sobą powiązane – różnica mogła przecież powstać jako wynik błędu z próby nawet w sytuacji, gdy zmienne w populacji są całkowicie od siebie niezależne. Mając wartość chi-kwadrat, możemy jednak określić prawdopodobieństwo, że za zaobserwowane różnice odpowiedzialne są tylko błędy losowe.

Stopnie swobody

Aby określić poziom istotności zaobserwowanej zależności między zmiennymi, musimy posłużyć się standardową tablicą wartości chi-kwadrat. Aby móc jej użyć, musimy najpierw obliczyć liczbę *stopni swobody*, określającą liczbę możliwych źródeł zmienności w ramach modelu statystycznego.

Wyobraźmy sobie, że chcę, abyś podał trzy liczby, których średnia wynosi 11. Liczba możliwych odpowiedzi jest nieskończona: (11, 11, 11), (10, 11, 12), (–11, 11, 33) itd. Gdybym dodatkowo zażądał, aby jedna z nich wynosiła 7, to wciąż pozostałe dwie liczby można by podać na nieskończenie wiele sposobów; gdybym jednak zażądał, by w trójce liczb znalazło się także 10, to trzecią liczbę można by wskazać tylko na jeden sposób, ponieważ suma trzech liczb, których średnia wynosi 11, musi wynosić 33, a skoro suma dwóch z nich wynosi 17, to trzecią musi być 16. W sytuacji wyboru trzech liczb o średniej 11 mamy do czynienia z dwoma stopniami swobody: możemy dowolnie wybrać dwie liczby, ale gdy to już zrobimy, determinujemy tym samym wartość trzeciej.

Bardziej ogólnie, jeśli analizujemy problem średniej N liczb, liczba stopni swobody wynosi $N - 1$; tak więc jeśli mamy znaleźć 23 liczby, których średnia jest dana, 22 z nich możemy wybrać dowolnie, wyznaczając w ten sposób jednoznacznie wartość 23 liczby.

Podobną logiką kieruje się określanie liczby stopni swobody dla tabel dwuzmiennowych, takich jak te analizowane za pomocą testu chi-kwadrat. Rozważmy tabelę opisującą zależność między dwiema zmiennymi dwuwartościowymi – *płcią* (kobieta/mężczyzna) i *postawami wobec aborcji* (dopuszcza/nie dopuszcza). Zauważmy, że w tabeli podane są liczebności brzegowe obu zmiennych.

Postawy wobec aborcji	Mężczyźni	Kobiety	Łącznie
Dopuszcza			500
Nie dopuszcza			500
Razem	500	500	1000

Choć jako liczebności brzegowych użyliśmy dla prostoty okrągłych liczb, to wewnątrz tabeli możemy mieć do czynienia z bardzo różnymi układami wartości. Być może 500 mężczyzn aprobuje aborcję, a 500 kobiet aborcji nie akceptuje; może jest dokładnie na odwrót. Może w każdą komórkę tabeli należy wpisać 250 – oczywiście możliwości jest nieporównanie więcej.

Ale postawmy pytanie: ile liczb możemy dowolnie wpisać w tabelę do momentu, gdy wartości wszystkich pozostałych komórek będą zdeterminowane przez wartości brzegowe? Odpowiedź jest prosta – tylko jedną. Na przykład, jeśli wiemy, że 300 mężczyzn aborcję akceptuje, to wynika z tego, że aborcji nie akceptuje 200 mężczyzn, a wśród kobiet wartości te muszą układać się na odwrót. Tak więc w przypadku tej tabeli możemy mówić o jednym stopniu swobody.

A co w przypadku tabeli trzy na trzy (z podanymi liczebnościami brzegowymi)? Czy umiesz powiedzieć, ile ma ona stopni swobody?

W przypadku testu chi-kwadrat liczbę stopni swobody wyznacza się mnożąc liczbę wierszy tabeli pomniejszoną o 1 przez liczbę kolumn tabeli pomniejszoną o 1, co można oddać wzorem $(w\text{-}1)(k\text{-}1)$. Tak więc w przypadku tabeli 3×3, mamy cztery stopnie swobody, gdyż $(3 - 1)(3 - 1) = 2 \times 2 = 4$.

W przykładzie badań dotyczących związku między płcią a praktykami religijnymi mieliśmy do czynienia z tabelą o dwóch wierszach i dwóch kolumnach (nie licząc wartości ogółem), a zatem liczba stopni swobody wynosi 1. Zaglądając do tabeli wartości chi-kwadrat (załącznik F) możemy sprawdzić, że losując próby z populacji, w której zmienne są od siebie niezależne, w 10% przypadków możemy się spodziewać chi-kwadrat większego lub równego 2,7, natomiast tylko w 1% prób chi-kwadrat przekroczy 6,6, a w zaledwie w pół procencie (0,005) 7,9. Im większa jest wartość chi-kwadrat, tym mniej jest prawdopodobne, że zależność jest spowodowana jedynie przez błąd z próby.

W naszym przykładzie wyliczona wartość chi-kwadrat wynosiła 12,70. Gdyby w populacji członków badanego wyznania nie było związku między praktykami religijnymi a płcią, to taka wartość chi-kwadrat trafiałaby się w mniej niż jednej na tysiąc (a więc mniej 0,1%) prób; innymi słowy prawdopodobieństwo uzyskania takiej lub większej wartości chi-kwadrat w próbie losowej, w sytuacji niezależności zmiennych w populacji, jest mniejsze niż 0,001. Opisując ten wynik, mówimy, że zaobserwowana przez nas w próbie zależność jest istotna na poziomie 0,001. Ponieważ prawdopodobieństwo, że jest ona spowodowana jedynie przez błąd z próby, jest tak małe, możemy odrzucić hipotezę zerową i uznać, że w całej populacji członków ba-

danego wyznania zachodzi związek między płcią a uczestnictwem w praktykach religijnych.

Dla większości miar związku procedurę testowania i wyznaczania poziomu istotności przeprowadza się w podobny sposób. Standardowe tablice pozwalają nam określić, czy dana zależność jest istotna statystycznie i na jakim poziomie. Szczegółowy sposób wykorzystania tablic statystycznych można znaleźć w każdym podręczniku statystyki.

Parę słów ostrzeżenia

Testy istotności są obiektywnym narzędziem, pozwalającym ocenić istotność statystyczną zależności między zmiennymi. Dzięki nim możemy osądzić, czy zaobserwowana przez nas zależność odpowiada rzeczywistej zależności zachodzącej między tymi zmiennymi w badanej populacji. Jednak badacze stosujący testy istotności oraz czytelnicy natykający się na nie w lekturze powinni być świadomi zagrożeń związanych z ich interpretacją.

Po pierwsze, w ninieszym rozdziale mówiliśmy o testach istotności statystycznej – nie ma jednak testów obiektywnej istotności *merytorycznej*. Tak więc mamy być przekonani, że zaobserwowana zależność nie została wygenerowana jedynie przez błąd z próby – ale może się zdarzyć, że ich związek jest bardzo słaby. Pamiętajmy, że błąd próby jest malejącą funkcją wielkości próby – im większa jest próba, tym mniejszego błędu z próby możemy się spodziewać. Zatem współczynnik korelacji o wartości, powiedzmy, 0,1, może być istotny (na określonym poziomie), jeśli nasza próba była duża, ale taka sama korelacja między tymi dwiema zmiennymi, wykryta w małej próbie może już nie być istotna statystycznie. Jest to zresztą całkowicie zrozumiałe w świetle teorii testów istotności: im większa jest próba, tym mniejsze prawdopodobieństwo, że zaobserwowana korelacja była wynikiem jedynie błędu z próby jest znacznie mniejsze. Ale niezależnie od wielkości próby korelację na tym poziomie można by uznać za nieistotną merytorycznie.

Różnicę między istotnością statystyczną a merytoryczną najłatwiej zilustrować przykładem, w którym możemy mieć *absolutną pewność*, że obserwowana zależność nie jest skutkiem błędu z próby. Ma to miejsce wtedy, gdy badamy całą populację. Wyobraźmy sobie, że ustaliliśmy daty urodzenia wszystkich urzędników państwowych w USA i w Rosji, przy czym okazało się, że średnia wieku urzędnika amerykańskiego wynosi 45 lat, a rosyjskiego 46. Ponieważ znamy wiek wszystkich urzędników, nie można mówić o jakimkolwiek błędzie z próby, wiemy zatem z całą pewnością, że średnio urzędnicy rosyjscy są starsi od urzędników amerykańskich. Mimo to uznamy, że różnica ta nie ma żadnego znaczenia merytorycznego – w praktyce możemy uznać, że urzędnicy w obu krajach są średnio w tym samym wieku.

Po drugie, aby nie zwiódł cię ten fikcyjny przykład, zauważ, że nie można wyznaczyć poziomu istotności statystycznej w przypadku zależności obserwowanej w całej populacji. Testy istotności odwołują się do prawdopodobieństwa, że zaobserwowany związek jest skutkiem wyłącznie błędu z próby, ale tam, gdzie nie było próby, nie może być błędu z próby.

Po trzecie, musimy także pamiętać, że testy istotności statystycznej odwołują się do tych samych założeń dotyczących doboru próby, które przyjmowaliśmy obliczając przedziały ufności. W im mniejszym stopniu są one spełnione, w tym mniejszym stopniu testy istotności są wiarygodne.

Jedynym omawianym przez nas testem istotności był test chi-kwadrat; w naukach społecznych stosuje się też oczywiście wiele innych testów. W toku dalszych studiów trafisz jeszcze przynajmniej na analizę wariancji i *t*-test.

Tak jak i wobec innych zagadnień omawianych w tej książce, także w tym przypadku mam swoje własne uprzedzenia: nie lubię testów istotności. Nie mam nic przeciwko logice istotności statystycznej – jest ona bez zarzutu, uważam jednak, że w wielu przypadkach testy istotności raczej wprowadzają w błąd, niż coś wyjaśniają. Wynika to, moim zdaniem, z następujących przyczyn:

1. Testy istotności opierają się na założeniach dotyczących doboru próby, które przez praktycznie stosowane schematy losowania właściwie nigdy nie są spełnione.

2. Testy istotności zakładają brak błędów systematycznych, co w warunkach badań empirycznych jest założeniem zawsze co najmniej wątpliwym.
3. W praktyce zbyt często są stosowane do oceny miar siły związku, które zostały obliczone z naruszeniem założeń, na których miary te są oparte (na przykład współczynników korelacji według momentu iloczynowego, obliczonych dla zmiennych porządkowych).
4. Poziom istotności statystycznej często jest interpretowany błędnie jako „siła związku" bądź też jako ocena jego znaczenia merytorycznego.

Problemy te podkreślają także autorzy analizy polityki wydawniczej, prowadzonej przez dziewięć czasopism psychologicznych i trzy czasopisma medyczne (Sterling, Rosenbaum, Weinkam 1995), którzy stwierdzili, że redakcje bardzo niechętnie publikowały artykuły, których autorzy nie powoływali się na statystycznie istotne korelacje między zmiennymi. Przykładem może być cytowany w pracy list z odmową publikacji:

> „Niestety, nie możemy opublikować Państwa pracy. Jest ona doskonale napisana i dobrze udokumentowana, ale negatywny charakter wyników powoduje, że jej znaczenie naukowe jest bardzo niewielkie. Zachęcamy do kontynuowania prac badawczych; z przyjemnością rozważymy publikację następnych artykułów, które nam Państwo w przyszłości nadeślą".
>
> (Sterling i in. 1995, s. 109)

Załóżmy, że badacz perfekcyjnie zrealizował projekt badawczy, mający sprawdzić, czy *X* powoduje *Y*. Jego wyniki wskazują, że nie istnieje istotna statystycznie korelacja. To jednak jest pewna informacja – jeśli chcemy się dowiedzieć, jakie są przyczyny raka, wojen, przestępczości nieletnich, to przyda nam się wiedza, że coś, co mogłoby być przyczyną, jednak nią *nie jest*. Wiedząc to, będziemy mogli zawęzić obszar dalszych badań.

Ale, jak widzieliśmy, takie badania mogłyby być odrzucone przez czasopisma naukowe – w efekcie inni badacze nadal będą poszukiwali związku przyczynowego między *X* a *Y*, nie zdając sobie sprawy, że został on w już przeprowadzonych badaniach wykluczony. W ten sposób niepotrzebnie zostanie zrealizowanych wiele badań, których wyniki nigdy nie będą opublikowane i w efekcie nie przyczynią się do pogłębienia wiedzy o zależności przyczynowej *X* a *Y*.

Jednak jak wynika z rachunku prawdopodobieństwa, jeżeli badań takich będzie odpowiednio wiele, to jest spora szansa, że w jednym z nich między tymi zmiennymi zostanie wykryta istotna statystycznie korelacja między *X* a *Y*. Nawet jeśli w populacji w ogóle nie występuje zależność między tymi zmiennymi, to w 5 na 100 badań możemy się spodziewać wykrycia związku istotnego na poziomie 0,05 – bo też dokładnie taka jest treść pojęcia „poziom istotności 0,05". Jeśli więc przeprowadzonych zostanie 100 badań, to autorzy pięciu z nich znajdą zależność przyczynową tam, gdzie jej nie ma – i wyniki właśnie tych pięciu będą miały największe szanse na publikację!

Tak więc zbytnie zaufanie do statystycznych testów istotności prowadzić może do poważnych problemów. Mimo to uważam, że – może paradoksalnie – testy istotności *mogą* być dla badacza pomocnym narzędziem, ułatwiającym zrozumienie danych. Chociaż wiele moich uwag może sugerować skrajnie konserwatywny stosunek do testów istotności – że należy je stosować tylko wtedy, gdy wszystkie ich założenia są spełnione – w istocie moje nastawienie do nich jest wręcz przeciwne.

Zachęcam cię do stosowania wszelkich technik statystycznych, wszystkich miar związku czy testów istotności, jeżeli tylko ułatwi ci to zrozumienie danych – nawet jeśli miałoby to oznaczać obliczanie współczynnika korelacji według momentu iloczynowgo dla zmiennych nominalnych czy badanie istotności statystycznej związku zaobserwowanego w zupełnie nielosowej próbie. Mówię to w duchu Hanan Selvin, jednej z pionierek modelu analizy rozbudowanej, która działania takie określała jako *data-dredging*[3]. Dozwolone jest wszystko, co pozwala lepiej zrozumieć dane empiryczne i badaną rzeczywistość społeczną.

[3] W wolnym tłumaczeniu „przeorywanie danych" (przyp. tłum.).

Ceną za taką radykalną wolność musi być jednak rezygnacja ze ścisłej interpretacji statystycznej uzyskanych wyników. Nie będziesz mógł dowodzić zasadniczego znaczenia uzyskanych wyników jedynie powołując się na korelację istotną na poziomie 0,05. Niezależnie od tego, jakimi metodami uda się dokonać zasadniczego odkrycia, dane muszą zostać zaprezentowane w sposób zgodny z przyjętymi normami, a ich znaczenie musi dać się logicznie uzasadnić.

Główne zagadnienia

- Statystyka opisowa służy do syntetycznej prezentacji wyników badań; w jej zakres wchodzą rozkłady jednej zmiennej, a także syntetyczne miary związku między zmiennymi.
- Mierniki statystyczne służące do opisu zależności między zmiennymi nazywane są *miarami siły związku*.
- Szereg miar siły związku oparty jest na modelu proporcjonalnej redukcji błędu (PRE). Model ten opiera się na porównaniu (1) liczby błędów przewidywania wartości jednej zmiennej dla każdego z badanych przypadków przy założeniu, że znamy tylko rozkład wartości tej zmiennej, oraz (2) liczby błędów przewidywania wartości jednej zmiennej w sytuacji, gdy znamy łączny rozkład obu zmiennych, a zgadując wartość jednej ze zmiennych, znamy wartość drugiej zmiennej dla danego przypadku. Do miar opartych na PRE należą λ (lambda), stosowana do mierzenia siły związku między dwiema zmiennymi nominalnymi, γ (gamma), stosowana do pomiaru związku między dwiema zmiennymi porządkowymi, oraz współczynnik korelacji według momentu iloczynowego Pearsona r, stosowany do analizy związku między dwiema zmiennymi interwałowymi lub ilorazowymi.
- Analiza regresji służy do przedstawiania zależności między zmiennymi w formie równania pozwalającego wyznaczyć przewidywaną wartość zmiennej zależnej, jeśli znamy wartości jednej lub wielu zmiennych niezależnych.
- Równania regresji odpowiadają liniom regresji, poprowadzonym tak, by z jak najmniejszymi błędami oddawał faktyczne położenie punktów na wykresie rozrzutu.
- Wyróżniamy regresję liniową, regresję wielokrotną, regresje cząstkową, regresję krzywoliniową.
- Do innych technik analizy wielozmiennowej należą: analiza szeregów czasowych, służąca do badania procesów zachodzących w czasie, analiza ścieżkowa, pozwalająca na graficzne przedstawienie sieci powiązań przyczynowych między różnymi zmiennymi, oraz analiza czynnikowa, metoda, w której poszukuje się ogólniejszych czynników (wymiarów), odpowiadających wielu faktycznym zmiennym.
- Wnioskowanie statystyczne obejmuje techniki pozwalające uogólniać na całą populację wyniki uzyskane w badaniu wylosowanej z niej próby. Niektóre techniki wnioskowania statystycznego pozwalają szacować parametry rozkładu jednej zmiennej w populacji, inne – testy istotności statystycznej – służą uogólnianiu zaobserwowanych w próbie zależności między zmiennymi w populacji.
- Uogólnienia dotyczące cech populacji powinny być czynione poprzez podanie przedziału i poziomu ufności, które są obliczane zgodnie z zasadami rachunku prawdopodobieństwa, przy założeniu, że próba została wylosowana w standardowy sposób.
- Uogólniając na całą populację obserwacje, dotyczące związków między zmiennymi, stosujemy testy istotności statystycznej, za pomocą których wyznaczamy prawdopodobieństwo, że z populacji, w której taki związek nie zachodzi, zostanie wylosowana próba, w której zaobserwowany zostanie związek co najmniej tak silny jak ten, który testujemy. Również testy istotności statystycznej opierają się na rachunku prawdopodobieństwa i odwołują się do założenia, że próba została wylosowana w standardowy sposób.
- Jednym z najczęściej stosowanych w naukach społecznych testów istotności statystycznej jest test χ^2 (chi-kwadrat).

- Poziom istotności zależności stwierdzonej w próbie podawany jest jako prawdopodobieństwo, że związek ten jest skutkiem jedynie błędów próby. Stwierdzenie, że zależność jest istotna na poziomie 0,05, jest tożsame stwierdzeniu, że zależności takiej lub silniejszej będącej skutkiem błędu próby można się spodziewać nie częściej niż w 5 na 100 prób.
- Wielu badaczy stosuje przy testowaniu zależności jedynie kilka poziomów istotności, jak 0,05, 0,01, 0,001; jest to jednak tylko konwencja.
- Istotności statystycznej zależności nie można utożsamiać z jej znaczeniem merytorycznym, a więc z tym, że jest ona silna, ważna, bądź też warta tego, żeby o niej napisać w liście do rodziców.
- Testy statystyczne odwołują się do wielu założeń dotyczących charakteru danych oraz sposobu losowania próby, które w zasadzie prawie nigdy nie są do końca spełnione w realnych warunkach badań społecznych. Mimo to testy statystyczne mogą być bardzo pomocne w analizie i interpretacji danych empirycznych.

Kluczowe pojęcia

statystyka opisowa
proporcjonalna redukcja błędu (PRE)
analiza regresji
regresja liniowa
regresja wielokrotna
regresja cząstkowa
regresja krzywoliniowa
analiza ścieżkowa
analiza szeregów czasowych
analiza czynnikowa
wnioskowanie statystyczne
błędy systematyczne
istotność statystyczna
testy istotności statystycznej
poziom istotności
hipoteza zerowa

Pytania i ćwiczenia

1. Przedstaw teoretyczne założenia proporcjonalnej redukcji błędu (PRE) i opartych na niej miar siły związku.
2. Jakie cele stawia się analizie regresji?
3. Czym różnią się miary siły związku od testów statystycznej istotności?
4. Znajdź opis badań, zawierający odwołania do poziomów istotności uzyskanych w nich wynikach; oceń klarowność sposobu, w jaki są one podane.
5. Znajdź opis badań, w których zastosowano analizę czynnikową, i przedstaw ich wyniki.

Dodatkowe lektury

Babbie Earl, Halley Fred, Zaino Jeanne, *Adventures in Social Research*, Pine Forge Press, Newbury Park, Cal. 2000. Książka zawiera wprowadzenie do analiz statystycznych w badaniach społecznych przy użyciu SPSS dla Windows, oraz omówienie (wraz z przykładami) szeregu podstawowych technik statystycznych.

Blalock Hubert M. Jr., *Statystyka dla socjologów*, PWN, Warszawa 1975. Podręcznik Blalocka przez kilkadziesiąt lat był podstawową pomocą dla studentów socjologii (a także dla ich nauczycieli). Śmierć Blalocka była dla nauk społecznych dotkliwą stratą.

Frankfort-Nachmias Chava, *Social Statistics for Diverse Society*, Pine Forge Press, Newbury Park, Cal. 1997. Uniwersalny podręcznik statystyki, charakteryzujący się wyjątkowo efektywnym wykorzystaniem grafiki do objaśniania logiki wielu powszechnie stosowanych w naukach społecznych technik statystycznych.

Healey Joseph F., *Statistics. A Tool for Social Research*, Wadsworth, Belmont, Cal. 1999. Dobre wprowadzenie do statystyki dla socjologów.

Mohr Lawrence B., *Understanding Significance Testing*, Sage, Newbury Park, Cal. 1990. Bardzo dobra książka, zajmująca się wyczerpująco zarówno techniczną stroną stosowania testów istotności, jak i ich właściwą interpretacją.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology **Resource** Center, Virtual Society, gdzie znajdziesz **dodat-**

kowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Część
5

Społeczny kontekst badań

18. Etyka i polityka w badaniach społecznych

19. Zastosowania badań społecznych

Jak mogłeś się dowiedzieć z poprzednich rozdziałów, praktyka badawcza wymaga znalezienia przemyślanego kompromisu pomiędzy określonymi celami, które badania naukowe mają realizować, i kryteriami, które powinny spełniać, a ograniczeniami o charakterze praktycznym i organizacyjnym. W rozdziale 18 zajmiemy się kwestiami etycznymi i politycznymi, które również wymagają kompromisów w stosunku do tego, co mogłoby się wydawać doskonałą „naukową" realizacją badań.

Celem rozdziału 19 jest zebranie w zintegrowaną całość wielu kwestii omawianych w tej książce, aby pokazać, do czego może się przydać znajomość badań społecznych w twojej przyszłej karierze. W rozdziale tym przypomnimy zależności pomiędzy teorią, procesem zbierania danych i analizą, oraz popatrzymy, w jaki sposób wiedzę wyniesioną z lektury tej książki można wykorzystać w pozabadawczych dziedzinach życia.

Wreszcie na końcu rozdziału 19 znajdziesz „poradnik konsumenta", mający stanowić pomoc w oddzieleniu ziarna od plew w przypadku licznych badań społecznych, o których będziesz czytał przez całe życie. Przy okazji, „poradnik konsumenta" może ci się przydać jako streszczenie całej książki w przypadku, gdybyś miał zdawać egzamin. Powodzenia. ■

Rozdział
18

Etyka i polityka w badaniach społecznych

Przegląd holograficzny

Badania społeczne nie są prowadzone w społecznej próżni. Badacz, planując i realizując badania, musi uwzględnić oprócz kwestii naukowych także liczne kwestie etyczne i polityczne. Jednak trudne problemy etyczne i polityczne rzadko kiedy mają jednoznaczne rozwiązania.

Wprowadzenie

Zagadnienia etyczne w badaniach społecznych
Dobrowolne uczestnictwo
Nie krzywdzić badanych
Anonimowość i poufność
Oszukiwanie badanych
Analiza i prezentacja wyników
Instytucjonalne organa kontrolne
Kodeksy etyki zawodowej

Dwie kontrowersje etyczne
Kłopoty w kafejce
Obserwacja posłuszeństwa

Polityczne aspekty badań społecznych
Obiektywność a ideologia
Polityka przez małe „p"
Polityka we właściwych proporcjach

Główne zagadnienia

Kluczowe pojęcia

Pytania i ćwiczenia

Dodatkowe lektury

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Pisząc tę książkę, stawiałem sobie za cel dać czytelnikom realistyczne i użyteczne wprowadzenie do badań społecznych. Aby było ono rzeczywiście realistyczne, musi uwzględniać cztery główne ograniczenia projektów badawczych: naukowe, organizacyjne, etyczne i polityczne.

W poprzednich siedemnastu rozdziałach zajmowaliśmy się ograniczeniami naukowymi i organizacyjnymi. Jak widzieliśmy, czasami przeprowadzenie procedur badawczych, najwłaściwszych ze względu na kryteria logiki badań naukowych, może okazać się niemożliwe ze względów organizacyjnych – bo badania byłyby zbyt drogie czy też trwałyby zbyt długo. W całej książce zajmowaliśmy się zatem praktycznymi kompromisami.

W tym rozdziale zbadamy dwa inne aspekty prowadzenia badań w realnym świecie: etykę i politykę. Tak jak pewne procedury badawcze mogą okazać się niepraktyczne ze względów organizacyjnych, tak samo mogą okazać się niedopuszczalne etycznie, bądź też problematyczne ze względów politycznych.

Aby pokazać, o co mi chodzi, przedstawię przykład z życia. Parę lat temu zostałem poproszony o pomoc w przygotowaniu badań dotyczących kształcenia kalifornijskich prawników. Badania miały przeprowadzić wspólnie uniwersyteckie centrum badawcze i kalifornijskie stowarzyszenie adwokatów; celem badań była poprawa jakości kształcenia prawników poprzez ustalenie, jakie aspekty studiów prawniczych mają największy wpływ na wyniki egzaminów zawodowych. Zasadniczo nasz plan zakładał, że wszyscy podchodzący do egzaminów adwokackich będą musieli wypełnić szczegółowy kwestionariusz, dotyczący przebiegu ich studiów. Analizując jak kandydaci o różnych doświadczeniach z czasu studiów wypadają na egzaminie, powinniśmy być w stanie wskazać, jakie czynniki związane z przebiegiem studiów miały wpływ na sukces lub porażkę. Wyniki badań miały być udostępnione uczelniom kształcącym prawników, co powinno wpłynąć pozytywnie na poziom kształcenia prawniczego.

Prowadzenie badań we współpracy ze stowarzyszeniem adwokatów było obiecujące, gdyż można było liczyć na uwolnienie się od wszystkich trudności, będących zazwyczaj zmorą badaczy, jak na przykład uzyskanie zgody na przeprowadzenie badań w czasie egzaminów czy ryzyko wysokiego poziomu odmów odpowiedzi ze strony respondentów.

Wychodząc ze spotkania, byłem pewien powodzenia projektu. Gdy przedstawiłem go koleżance, rozwodziłem się zwłaszcza nad całkowitym rozwiązaniem problemu braków odpowiedzi. Jej natychmiastowa reakcja spowodowała, że sytuacja zmieniła się o 180 stopni. „To nieetyczne. Nie mamy prawa żądać wypełnienia ankiety; udział w badaniu musi być całkowicie dobrowolny". Projekt nie został zrealizowany.

Gdy teraz opowiadam tę historię, jest dla mnie całkowicie oczywiste, że żądanie udziału w badaniach byłoby niewłaściwe. Możliwe, że doszedłeś do tego wniosku, zanim wspomniałem o reakcji mojej koleżanki. Nadal, wspominając o tej sprawie, czuję się trochę głupio, ale miałem istotny powód, żeby opowiedzieć o tych własnych doświadczeniach.

Wszyscy uważamy się za ludzi wrażliwych moralnie – może nie idealnych, ale w każdym razie nie mniej, a zapewne nawet bardziej uczciwych niż inni. Problem w tym, że w badaniach społecznych (tak jak w życiu w ogóle) nie zawsze dostrzegamy etyczne aspekty pewnych działań. W efekcie często robimy rzeczy, których etycznego wymiaru nie zauważamy, choć może być on widoczny dla innych i mógłby być nawet oczywisty dla nas, gdyby ktokolwiek zwrócił nam na to uwagę. Kiedy grupie planującej badania, o których opowiadałem, przekazałem wątpliwości co do etycznej dopuszczalności domagania się od egzaminowanych udziału w badaniach, nikt nie zaprzeczył, że jest to niewłaściwe; wszyscy czuli się tylko nieco speszeni, że sami nie zauważyli tego wcześniej.

Nikt nie miałby wątpliwości, że badanie wymagające torturowania małych dzieci byłoby nieetyczne. Wiem też, że natychmiast byś zaoponował, gdybym zaproponował przeprowadzenie wywiadów pogłębionych, dotyczących życia seksualnego, po czym opublikował wypo-

wiedzi respondentów w lokalnej gazecie. Ale mimo twojej wrażliwości moralnej pominąłbyś zupełnie kwestie etyczne w pewnych innych sytuacjach – wszystkim nam się to zdarza.

W pierwszej części bieżącego rozdziału będziemy się zajmowali etyką badań społecznych, m.in. pewnymi powszechnie przyjętymi normami, określającymi, co w badaniach społecznych jest, a co nie jest dopuszczalne. Jednak ważniejsze niż znajomość ogólnych zasad jest to, byś planując badania był wyczulony na ich komponent etyczny. Nawet gdy aspekty etyczne danej sytuacji są dyskusyjne, powinieneś wiedzieć, że jest o czym dyskutować.

Polityczne aspekty badań również mogą mieć złożony, niejednoznaczny i sporny charakter. Zauważ, że w przykładzie, dotyczącym kształcenia prawników, problem miał charakter nie tylko etyczny, ale również polityczny. Zasada, że udział w badaniach powinien być całkowicie dobrowolny, jest powszechnie przyjętą normą etyki badań społecznych, ale wynika ona w sposób oczywisty z przyjętych w USA politycznych zasad ochrony wolności obywatelskich. W niektórych państwach badania takie mogłyby być uznane za w pełni dopuszczalne.

W drugiej części rozdziału zajmiemy się przypadkami projektów badawczych, które załamały się – bądź też były tego bliskie – z powodów politycznych. Podobnie jak w przypadku problemów etycznych, w pewnych sytuacjach nie ma rozwiązań po prostu „właściwych". Nie mam zamiaru narzucać ci żadnego konkretnego stanowiska co do tego, co jest politycznie akceptowalne, a co nie: podobnie jak z etyką, najważniejsze jest, by planując badania zwracać uwagę także na ich aspekty polityczne.

Zagadnienia etyczne w badaniach społecznych

Słowo „etyka", zarówno w większości słowników, jak i w codziennym użyciu, wiąże się na ogół z pojęciem moralności; oba te terminy odnoszą się do kwestii dobra i zła. Ale co jest dobre, a co złe? Co jest źródłem tego rozróżnienia? Na to pytanie padają różne odpowiedzi. Dla jednych jest to religia, dla innych ideologia polityczna albo pragmatyczna obserwacja konsekwencji różnych działań.

New World Dictionary Webstera, podobnie jak inne słowniki, definiuje słowo *etyczny* jako „zgodny z grupowymi bądź zawodowymi standardami postępowania" (*Conforming to the standards of conduct of a given profession or group*). Choć taka definicja może się nie podobać tym, którzy poszukują absolutnych zasad moralnych, to, co uważamy za moralne i etyczne w codziennym życiu, jest kwestią umowy między członkami danej grupy. Trudno się wobec tego dziwić, że różne grupy przyjmują różne normy postępowania. Dla dobrego funkcjonowania w konkretnej społeczności konieczna jest wiedza, co w niej jest, a co nie jest moralnie dopuszczalne. To samo dotyczy także społeczności naukowej.

Każdy, kto zaangażowany jest w prowadzenie naukowych badań społecznych, powinien znać powszechnie przyjęte w gronie naukowców normy określające, co w działalności badawczej jest, a co nie jest dopuszczalne. Niniejszy podrozdział podsumowuje niektóre najważniejsze normy etyczne, rozpowszechnione w badaniach społecznych.

Dobrowolne uczestnictwo

Często, jeśli nie zawsze, prowadzenie badań społecznych wiąże się z ingerencją w życie badanych. Dzwonek do drzwi, przed którymi stoi ankieter, albo znaleziony w skrzynce na listy kwestionariusz sygnalizują początek działań, których respondent się nie domagał, a które czasami wymagają od niego sporego nakładu czasu i energii. Udział w badaniach społecznych zaburza normalny plan dnia badanego.

Co więcej, badacz często oczekuje od respondenta ujawnienia pewnych osobistych informacji – być może dotyczących rzeczy, o których nie wiedzą nawet jego przyjaciele i współpracownicy. Teraz te informacje miałyby być ujawnione zupełnie obcej osobie. Inni specjaliści – na przykład lekarze czy adwokaci – też oczekują ujawniania tego rodzaju informacji, ale w ich przypadku jest to usprawiedliwione, gdyż wiedza taka jest gromadzona w interesie

respondenta. Badacze społeczni nic takiego nie mogą obiecać – co najwyżej, podobnie jak badacze w dziedzinie medycyny, mogą deklarować, że uzyskane wyniki będą spożytkowane ku dobru całej ludzkości.

Podstawową zasadą etyczną w badaniach medycznych jest dobrowolność udziału w eksperymentach. Ta sama zasada odnosi się do badań społecznych: nikt nie może być zmuszany do brania w nich udziału. Problem w tym, że zasadę tę znacznie łatwiej zaakceptować w teorii, niż zastosować w praktyce.

Można tutaj przytoczyć kolejną analogię z badaniami medycznymi. Wiele nowych leków jest testowanych na więźniach. W przypadkach najściślejszego przestrzegania norm etycznych więźniowie są informowani o charakterze eksperymentu oraz o wszelkich potencjalnych niebezpieczeństwach, wiążących się z udziałem w nim. Mówi się im także, że udział w eksperymencie jest w pełni dobrowolny, a jego uczestnicy nie będą w żaden sposób nagradzani (na przykład łatwiejszym uzyskaniem zwolnienia warunkowego). Jednak nawet w takich okolicznościach jasne jest, że ochotnicy będą się kierowali nadzieją na uzyskanie za udział w eksperymencie jakichś korzyści.

Gdy wykładowca prosi studentów pierwszego roku socjologii o wypełnienie ankiety, której wyniki chciałby zanalizować i opublikować, powinien im powiedzieć, że ich udział w badaniu jest całkowicie dobrowolny. Jednak nawet wtedy, gdy to zrobi, większość studentów może się obawiać, że odmowa będzie miała jakiś wpływ na ich końcową ocenę. Wykładowca musi być zatem szczególnie wyczulony na tego rodzaju obawy i zrobić wszystko, by je rozwiać. Może na przykład wyjść z sali w czasie, gdy studenci wypełniają ankiety, albo poprosić, by wypełnione kwestionariusze zostały odesłane do niego pocztą lub wrzucone do pudełka przed kolejnymi zajęciami.

Problem z zasadą dobrowolności udziału leży w tym, że znajduje się ona w bezpośrednim konflikcie z interesem naukowym badacza. Najogólniej rzecz biorąc, realizacja naukowego celu, związanego z uogólnianiem wyników, jest zagrożona w sytuacji, gdy uczestnikami eksperymentu czy respondentami w badaniach sondażowych są osoby, w których naturze leży chętne branie udziału w takich rzeczach. Ponieważ takie nastawienie może wynikać z pewnych bardziej ogólnych cech osobowości, wyniki takich badań mogłyby nie nadawać się do uogólniania na wszystkich ludzi. Ściślej – w przypadku badań sondażowych o charakterze opisowym badacz nie może uogólniać wyników uzyskanych w próbie na całą populację, jeśli badaniu nie została poddana zdecydowana większość osób wylosowanych do próby – a więc zarówno ci, którzy w badaniach udział biorą chętnie, jak i mniej chętni.

W badaniach terenowych, o czym była mowa w rozdziale 10, pojawia się podobny dylemat etyczny. Bardzo często badacz w ogóle nie może ujawnić, że przeprowadza jakieś badania, gdyż mogłoby to zasadniczo wpłynąć na przebieg badanych procesów społecznych. Oczywiście w takich przypadkach badani nie mają możliwości, by dobrowolnie wziąć udział w badaniu lub odmówić go.

Zasada dobrowolności udziału jest niewątpliwie ważna – ale w wielu przypadkach nie daje się zastosować. Jeśli już uznasz, że okoliczności uprawniają cię do jej złamania, musisz w jeszcze większym stopniu dbać o zachowanie innych norm etyki badań społecznych, takich jak zasada, że udział w badaniach nie może wyrządzić badanym żadnej szkody.

Nie krzywdzić badanych

Osoba uczestnicząca w badaniach, nawet całkowicie dobrowolnie, nigdy nie może być na ich skutek poszkodowana. Chyba najdobitniejszy przykład praktycznego zastosowania tej normy to zakaz ujawniania informacji, których ujawnienie byłoby dla badanych kłopotliwe lub mogłoby ich narazić na jakiekolwiek problemy rodzinne, towarzyskie, zawodowe czy inne (do tych kwestii jeszcze wrócimy).

Ponieważ uczestnik może w wyniku badania ponieść szkody o charakterze psychicznym, badacz musi przeanalizować wszelkie związane z tym elementy ryzyka i starać się badanego przed nimi uchronić. W wielu badaniach oczekuje się od badanego, by ujawnił swoje dewiacyjne zachowania, niepopularne poglądy, czy

podawał inne krępujące dla siebie informacje, na przykład o niskich dochodach czy korzystaniu z pomocy społecznej. Odpowiadanie na pytania dotyczące tych kwestii z reguły będzie dla badanego co najmniej nieprzyjemne.

Udział w badaniach często wiąże się także z koniecznością skonfrontowania się z tymi aspektami własnej osobowości, których badany normalnie nie dostrzega. Nie musi się to wiązać bezpośrednio z przekazywaniem jakichś informacji badaczowi: na przykład przy retrospekcji badany może uznać jakieś swoje zachowanie z przeszłości za niewłaściwe lub niemoralne. W konsekwencji uczestnictwo w badaniu może stać się dla badanego źródłem długotrwałego cierpienia: na przykład uczestnicząc w badaniach dotyczących moralności, badany może zacząć podawać w wątpliwość swój własny system moralny i będzie w rozterce jeszcze długo po zakończeniu całego projektu badawczego. Podobnie, konieczność wielokrotnego odpowiadania na pytania dotyczące tej samej kwestii może zranić osoby o niskim poczuciu własnej wartości.

Jak widać, ryzyko skrzywdzenia badanego wiąże się właściwie ze wszystkimi badaniami społecznymi. Nie da się go całkowicie wyeliminować, ale w przypadku niektórych typów badań jest ono większe niż przy innych. Jeśli konkretna procedura badawcza uprawdopodabnia pojawienie się tego rodzaju nieprzyjemnych dla badanego efektów – na przykład gdy w kwestionariuszu wywiadu mają się znaleźć pytania dotyczące dewiacyjnych zachowań respondenta – badacz musi mieć bardzo poważne naukowe uzasadnienie dla takiej decyzji. Jeśli badania mają być rzeczywiście wartościowe, a mogą być przy tym istotnie przykre dla badanych, badacz staje przed bardzo nieprzyjemnym dylematem etycznym. I choć jego rozterki same w sobie nic nikomu nie dają, to przynajmniej wskazują na pewne uwrażliwienie na ten problem.

Z zasady dobrowolności udziału w badaniach i zasady nieszkodzenia badanym wyprowadza się zasadę **świadomej zgody**. Oznacza ona, że badani wyrażają zgodę na udział w badaniach z pełną świadomością możliwych zagrożeń. Przykładowo, w przypadku eksperymentu medycznego potencjalni uczestnicy badania powinni być szczegółowo poinformowani o jego charakterze i wszystkich możliwych konsekwencjach. Wyrażając zgodę na uczestnictwo w eksperymencie, muszą oni podpisać oświadczenie, że zdają sobie sprawę z zagrożeń i mimo to decydują się na udział. Procedura taka nie budzi wątpliwości w przypadku testowania leku wywołującego skutki fizyczne, jednak trudno byłoby ją zastosować do, powiedzmy, obserwacji uczestniczącej zamieszek ulicznych. Mimo że w takim przypadku badacz obserwator nadal związany jest zasadą nieszkodzenia obserwowanym, uzyskiwanie „świadomej zgody" nie jest właściwym środkiem osiągnięcia tego celu.

Inne – często niezauważane – potencjalne źródło zagrożenia dla badanych związane jest z analizą i publikowaniem wyników badań. Zawsze trzeba się liczyć z faktem, że badani czytają publikacje dotyczące badań, w których uczestniczyli. Przy odpowiedniej dozie dociekliwości będą w stanie odnaleźć się w różnych indeksach czy tabelach, i w konsekwencji stwierdzić, że, choć nie z nazwiska, ale zostali określeni jako bigoci, ateiści czy osoby pozbawione patriotyzmu. Takie opisy mogą ich co najmniej zmartwić i zachwiać ich postrzeganie siebie, mimo że całe badania mogły mieć na celu znalezienie obiektywnych źródeł uprzedzeń.

W pewnych badaniach, dotyczących kobiet chodzących do kościoła (Babbie 1967), proszono pastorów w wylosowanych kościołach o przekazanie kwestionariuszy próbie wiernych wylosowanej w ich parafii, a następnie zebranie ich i zwrot do agencji badawczej. Jeden z pastorów przed oddaniem kwestionariuszy przejrzał je, po czym w czasie nabożeństwa obrzucił swoich parafian ogniem piekielnym i siarką, mówiąc, że wielu z nich to ateiści, którzy pójdą do piekła. Mimo że nie mógł on wiedzieć, od których parafian pochodziły konkretne kwestionariusze, jest oczywiste, że w wyniku tego działania wielu respondentów doznało osobistej krzywdy.

Podobnie jak dobrowolność uczestnictwa, zasada unikania szkodzenia badanym jest łatwa w teorii, ale często trudna w praktyce. Jednak wrażliwość na tę zasadę i doświadczenie w jej stosowaniu powinny ułatwić badaczowi taktowne zachowanie w delikatnych obszarach badań.

W ostatnich latach przestrzeganie przez badaczy norm etycznych jest w coraz większym stopniu wspierane przez czynniki zewnętrzne. Federalne i inne instytucje finansujące z reguły wymagają zamieszczenia w ofertach badawczych niezależnej ewaluacji zasad postępowania wobec badanych; na większości amerykańskich uniwersytetów zajmują się tym specjalne komisje. Chociaż takie wymogi nie zawsze są stosowane prawidłowo i bywają kłopotliwe, generalnie nie tylko chronią one przed nieetycznymi badaniami, ale często ujawniają kwestie etyczne przeoczone przez nawet najbardziej skrupulatnych badaczy.

Anonimowość i poufność

Podstawową zasadą, związaną z zabezpieczeniem interesów i dóbr osób badanych, jest ochrona ich tożsamości. Dotyczy to przede wszystkim badań sondażowych i jest szczególnie ważne w sytuacjach, gdy ujawnienie odpowiedzi na pytania ankiety mogłoby przynieść respondentowi jakąkolwiek szkodę. Dla chronienia tożsamości badanych stosuje się dwie, często mylone ze sobą techniki: anonimowość i poufność.

Anonimowość

Projekt badawczy gwarantuje badanym **anonimowość**, gdy nie tylko osoby zapoznające się z wynikami badań, ale także osoby, które badania przeprowadzały, nie mają możliwości identyfikacji autorów poszczególnych wypowiedzi. Wynika z tego, że badania sondażowe z zastosowaniem wywiadów nigdy nie mogą być traktowane jako anonimowe, ponieważ przeprowadzający wywiad uzyskuje wypowiedzi od osoby, której tożsamość może określić (zakładam, że stosowane były standardowe techniki losowania próby). Przykładem badań anonimowych może być ankieta pocztowa, pod warunkiem że przed rozesłaniem kwestionariuszy nie były na nie nanoszone żadne numery identyfikacyjne.

Jak pamiętamy z rozdziału 9, dotyczącego badań sondażowych, zapewnienie anonimowości utrudnia monitorowanie, kto zwrócił kwestionariusz, a kto nie. W pewnych sytuacjach jednak korzyści wynikające z zapewnienia anonimowości przewyższają koszty. Na przykład w pewnych badaniach dotyczących brania narkotyków przez studentów szczególnie zależało mi na anonimowości, gdyż oczekiwałem, że uczciwe jej zagwarantowanie wpłynie na większą szczerość odpowiedzi i w ogóle skłonność do ich udzielania. Poza tym nie chciałem znaleźć się w sytuacji, w której władze zażądałyby ode mnie ujawnienia nazwisk osób łamiących prawo. W kilku przypadkach, w których respondenci z własnej inicjatywy podpisali ankiety, ich nazwiska od razu zamazaliśmy.

Poufność

Projekt badawczy gwarantuje **poufność**, gdy badacz może zidentyfikować autora wypowiedzi, ale jednoznacznie obiecuje informacji tej nie ujawniać. Na przykład w przypadku badań sondażowych, przeprowadzonych za pomocą wywiadów, badacz mógłby podać do publicznej wiadomości wysokość dochodów, jaką zadeklarował dany respondent, ale zobowiązuje się wobec badanych, że czegoś takiego nie zrobi.

W każdym przypadku, gdy projekt badawczy zakłada poufność, a nie anonimowość, obowiązkiem badacza jest jasno poinformować o tym badanych. W szczególności niedopuszczalne jest stosowanie terminu *anonimowość* w znaczeniu *poufność*.

Z paroma wyjątkami (na przykład sondażowe badania z udziałem osób publicznych, które zgodziły się na ujawnienie ich wypowiedzi), informacje uzyskane od badanych muszą być traktowane co najmniej jako poufne. Normę tę nie zawsze łatwo jest zachować: z punktu widzenia sądów informacje pochodzące z badań społecznych nie są traktowane jako „przekaz uprzywilejowany" – tak jak to jest w przypadku informacji posiadanych przez księży lub adwokatów.

Brak prawnej ochrony poufności danych sondażowych omal nie doprowadził do katastrofy w 1991 roku. Działo się to dwa lata po tym, jak w pobliżu portu Valdez na Alasce wpadł na skały supertankowiec Exxon *Valdez*, powodując wyciek do morza kilkudziesięciu tysięcy ton ropy naftowej. Ekonomiczne i eko-

logiczne konsekwencje tamtej katastrofy były szeroko opisywane. Mniej natomiast uwagi przywiązywano do szkód o charakterze społecznym i psychologicznym, jakie ponieśli mieszkańcy tego obszaru. Pojawiały się legendarne doniesienia o wzroście alkoholizmu, przemocy w rodzinach i innych pośrednich skutkach osłabienia więzi społecznych, będących następstwem katastrofy ekologicznej. W konsekwencji 22 gminy z rejonu Cieśniny Księcia Williama i Zatoki Alaskańskiej zaskarżyły do sądu firmę Exxon za poniesione przez ich mieszkańców szkody ekonomiczne, społeczne i psychologiczne.

Aby określić wielkość owych szkód, gminy zleciły firmie badawczej z San Diego przeprowadzenie sondażu na próbie gospodarstw domowych, w którym zadawano badanym wiele bardzo osobistych pytań dotyczących ich zwiększonych problemów rodzinnych. Respondenci, którym zagwarantowano absolutną poufność wypowiedzi, proszeni byli o podanie wielu bolesnych i wstydliwych informacji. Wyniki badań potwierdziły ostatecznie, że po katastrofie tankowca istotnie wzrosła częstość występowania różnych niekorzystnych zjawisk dotyczących życia osobistego lub rodzinnego badanych.

Gdy Exxon dowiedział się o zamiarze użycia wyników sondażu jako dowodu dokumentującego cierpienia mieszkańców, jego reakcja była nieoczekiwana: zażądał załączenia do akt wszystkich kwestionariuszy! Sąd uznał zasadność tego żądania i nakazał badaczom przekazanie ankiet – wraz ze wszystkimi danymi identyfikującymi respondentów. Było jasne, że intencją Exxona było wezwanie respondentów w charakterze świadków i przesłuchanie ich w sądzie co do spraw, o których mówili ankieterom pod gwarancją poufności. Dodatkowym aspektem sprawy był fakt, że znaczną część badanych stanowili Indianie i Eskimosi, dla których, ze względów kulturowych, takie publiczne przesłuchanie byłoby szczególnie przykre.

Na całe szczęście sprawa zakończyła się ugodą, zanim sąd musiał podjąć decyzję o ewentualnych przesłuchaniach – ale niestety, ryzyko takiego obrotu wydarzeń nie zostało wyeliminowane.

Poważne problemy tego rodzaju mają nie tylko renomowane instytucje badawcze. Rik Scarce w czasie studiów na Uniwersytecie Stanu Waszyngton przeprowadził obserwację uczestniczącą w środowisku aktywistów ruchu obrony praw zwierząt. Wyniki swoich badań ogłosił w 1990 roku w książce *Ecowarriors: Understanding the Radical Environmental Movement*. W 1993 roku sąd zażądał od niego ujawnienia tożsamości obserwowanych przez niego aktywistów. Młody badacz, trzymając się zasady poufności, odmówił sądowi, co kosztowało go 159 dni spędzonych w więzieniu hrabstwa Spokane.

Istnieje kilka technik pozwalających uchronić się przed takim niebezpieczeństwem i skuteczniej gwarantować poufność. Po pierwsze, ankieterzy i inne osoby mające dostęp do danych personalnych respondentów powinni przejść szkolenie dotyczące ich etycznej odpowiedzialności. Niezależnie od szkolenia podstawową techniką jest niszczenie danych pozwalających zidentyfikować badanych, kiedy tylko przestają być potrzebne. Na przykład w przypadku badań sondażowych na kwestionariuszach nazwiska i adresy powinny zostać zastąpione numerem identyfikacyjnym. Należy stworzyć spis pozwalający powiązać numery z konkretnymi osobami, co umożliwia późniejsze poprawki w przypadku brakujących lub sprzecznych danych, ale nie powinien on być dostępny w innych sytuacjach.

Analogicznie, w badaniach sondażowych przeprowadzanych za pomocą wywiadów dane identyfikujące respondentów mogą nam być potrzebne do ponownego kontaktu w celu potwierdzenia, że wywiad był rzeczywiście przeprowadzony, i ewentualnego uzupełnienia brakujących danych. Jednak natychmiast po tym, gdy wywiad zostanie zweryfikowany i będziesz pewien, że nie będziesz potrzebował od respondentów żadnych dodatkowych informacji, możesz spokojnie usunąć wszystkie dane identyfikacyjne. Często kwestionariusze drukowane są w taki sposób, że wszystkie dane identyfikacyjne znajdują się na pierwszej stronie, którą można wyrwać, gdy przestaną być już potrzebne. Z kolei J. Steven Picou (1996a, 1996b) zwraca uwagę, że nawet usunięcie na-

zwisk i adresów respondentów ze zbiorów danych nie musi jeszcze skutecznie chronić ich poufności, o czym przekonał się podczas blisko rocznej batalii przed sądem federalnym. Dokładna analiza odpowiedzi danego respondenta w pewnych przypadkach pozwala wydedukować jego tożsamość: wyobraźmy sobie na przykład, że ktoś podał, że pracował kiedyś w pewnej firmie – gdy znana jest także płeć, wiek, grupa etniczna itd., firma mogłaby tę osobę zidentyfikować.

Nawet jeżeli planujesz usunąć dane identyfikacyjne, załóżmy, że jeszcze tego nie uczyniłeś. Co zrobisz, gdy policja lub sąd zażądają od ciebie ujawnienia odpowiedzi udzielonych przez respondentów?

Mamy tutaj do czynienia z rzeczywistym problemem badaczy społecznych, chociaż opinie co do tego, jak najlepiej chronić interesy badanych, są różne. Na przykład Harry O'Neill, wiceprzewodniczący *Roper Organization*[1] wyraził opinię, że najlepszym rozwiązaniem jest zupełna rezygnacja z możliwości powiązania udzielonych odpowiedzi z nazwiskami respondentów.

> „Jak to można osiągnąć? Po prostu trzeba nie mieć żadnych danych, umożliwiających identyfikację respondentów, możliwych do udostępnienia sądowi na żądanie. Kontaktując się po raz pierwszy z klientami adwokatów, od razu jednoznacznie deklaruję, że po zebraniu i zweryfikowaniu danych wszystkie dane identyfikujące badanych zostaną natychmiast usunięte zniszczone. Wszystkie inne materiały – wypełnione kwestionariusze, taśmy z danymi, opis metodologii, nazwiska ankieterów i koordynatorów – zostaną, oczywiście, zachowane".
>
> (1992, s. 4)

Przewodniczący Burns Roper (1992, s. 5) nie zgodził się z tą opinią, argumentując, że takie postępowanie może wywołać wątpliwości co do trafności metod badawczych. Zdaniem Ropera badacz musi być po prostu przygotowany na pójście w razie potrzeby do więzienia (jak dodał, jego zastępca O'Neill obiecał mu w takim przypadku odwiedziny).

Oszukiwanie badanych

Jak widać, etyka badań społecznych przywiązuje duże znaczenie do kwestii tożsamości osób badanych, ale istotne może być także to, co badacz robi z własną tożsamością. W pewnych sytuacjach ujawnienie przez badacza swojej roli może być dla realizacji badania korzystne lub wręcz nieodzowne. Musiałbyś być doświadczonym sztukmistrzem, żeby nakłonić ludzi do uczestniczenia w laboratoryjnym eksperymencie czy do wypełnienia długiej ankiety, nie mówiąc im, że przeprowadzasz badania.

Jednak nawet w sytuacjach, gdy badacz musi ukryć swoją tożsamość, powinien pamiętać o tym, że musi to mieć poważne naukowe lub organizacyjne uzasadnienie, które zresztą i tak zawsze może być kwestionowane: wprowadzanie ludzi w błąd samo w sobie jest nieetyczne.

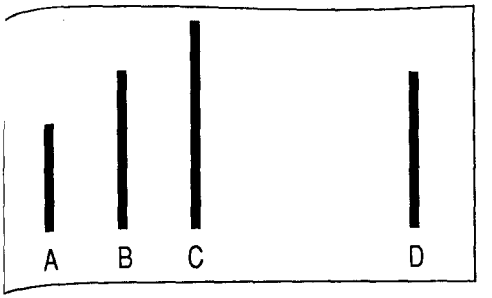
Czasami badacze przyznają, że prowadzą badania, ale kręcą co do ich charakteru lub zleceniodawcy. Wyobraźmy sobie, że na zlecenie instytucji zajmującej się opieką społeczną badamy poziom życia osób korzystających z pomocy społecznej. Nawet jeśli celem badań jest znalezienie sposobów na podniesie poziomu życia badanych, ci ostatni prawdopodobnie będą podejrzewali, że chodzi o polowanie na osoby, które pomoc wyłudzają, i w konsekwencji będą mieli skłonność do przedstawiania swojego życia w ciemniejszych niż rzeczywiste barwach. Ale jeżeli odpowiedzi badanych nie dadzą prawdziwego obrazu rzeczywistości, wyniki badań nie ułatwią skuteczniejszego udzielania im pomocy. Co zrobić w takiej sytuacji?

Jednym z możliwych rozwiązań byłoby powiedzenie respondentom, że badania prowadzone są przez uniwersytet – i ukrycie informacji o ich rzeczywistym zleceniodawcy. Postępowanie takie wpłynęłoby korzystnie na naukową jakość badania, ale budzi poważne wątpliwości etyczne.

Okłamywanie badanych co do celów badań jest bardzo częste w przypadku eksperymentów laboratoryjnych. Nawet wtedy, gdy trudno

[1] Roper Organization to jedna z najpoważniejszych amerykańskich instytucji badawczych (przyp. tłum.).

Rysunek 18.1
Eksperyment Ascha: linie różnej długości



ukryć fakt przeprowadzania badań, nie jest rzeczą trudną ukrycie ich celu, a niekiedy jest to wskazane. Przykładowo, wiele eksperymentów w psychologii społecznej przeprowadzono w celu zbadania, do jakiego stopnia ludzie są skłonni odrzucić świadectwo własnych zmysłów na rzecz opinii innych. Rysunek 18.1 przedstawia bodziec stosowany w klasycznym eksperymencie przeprowadzonym pierwotnie przez Ascha i często powtarzanym przez studentów na zajęciach z psychologii. W eksperymencie tym pokazywano badanym rysunek z trzema różniącymi się długością liniami (A, B i C), ci zaś mieli odpowiedzieć, która z nich jest tej samej długości co linia D.

Udzielenia poprawnej odpowiedzi – „B" – zapewne nie uznałbyś za szczególnie trudne. Napotkałbyś jednak pewną komplikację, ponieważ inni „badani", siedzący obok ciebie, zgodnie twierdziliby, że poprawna odpowiedź brzmi „A"! W rzeczywistości, oczywiście, byliby to współpracownicy eksperymentatora, poinstruowani, że mają udzielić błędnej odpowiedzi. Celem tego eksperymentu jest sprawdzenie, czy zrezygnujesz z własnej oceny na rzecz zgodnej opinii grupy. Myślę, że zgodzisz się, że konformizm jest zjawiskiem wartym badania i zrozumienia, oraz że nie dałoby się go badać eksperymentalnie, nie wprowadzając badanych w błąd co do istoty eksperymentu. Z podobną sytuacją spotkamy się jeszcze w tym rozdziale, omawiając słynny eksperyment Stanleya Milgrama. Jednak sam fakt, że wprowadzenie badanych w błąd jest konieczne dla powodzenia eksperymentu, nie rozwiązuje jeszcze wszystkich związanych z tym kwestii etycznych.

Jedną z rzeczy, które badacze uznają za właściwe, jest **debriefing** – sesja poeksperymentalna, w czasie której przeprowadza się wywiady z badanymi. Można wówczas zidentyfikować i skorygować wszystkie problemy, które wynikają dla nich z uczestnictwa w eksperymencie. Jeśli nawet były jakieś przyczyny, dla których nie można było ujawnić badanym prawdziwego celu eksperymentu przed badaniem, rzadko kiedy pozostają one aktualne po badaniu; ujawnienie prawdy może w jakimś stopniu zrekompensować fakt, że początkowo byli okłamywani. Należy jednak zrobić to w sposób dający pewność, że nie wywołało się przy tym u badanych wątpliwości co do własnej wartości, związanych z ich zachowaniem w sytuacji eksperymentalnej. Jeśli wydaje się to trudne, trzeba pamiętać, że jest to cena, jaką płacimy za traktowanie cudzego życia jako przedmiotu badań.

Jako badacz będziesz miał wiele zobowiązań moralnych wobec uczestników twoich badań. Ramka „Zagadnienia etyczne w badaniach ludzkiej seksualności" ilustruje pewne problemy etyczne, związane z tym specyficznym obszarem badań.

Analiza i prezentacja wyników

Oprócz moralnych zobowiązań wobec badanych badacze mają także zobowiązania moralne wobec swoich kolegów, związane z obszarem analizy danych i zasadami publikowania wyników.

Autor badań lepiej niż ktokolwiek inny zna wszystkie ich techniczne słabości i ograniczenia. Jego obowiązkiem jest poinformowanie o nich czytelników publikacji i raportów z badań, nawet jeżeli ma się czuć głupio, przedstawiając zastrzeżenia do własnej pracy i ujawniając swoje pomyłki.

Negatywne wyniki analiz również powinny być zawarte w publikacjach. Jest w nauce szkodliwym mitem (w który, niestety, wierzą również niektórzy redaktorzy periodyków naukowych), że jedynie wyniki o charakterze pozytywnym warte są publikacji. Jednak często wiedza, że jakieś dwie zmienne *nie* są od siebie zależne, jest dla nauki równie ważna, jak ta dotycząca ich ewentualnej zależności.

Zagadnienia etyczne w badaniach ludzkiej seksualności

Kathleen McKinney

Wydział Socjologii, Uniwersytet Stanowy Illinois

Gdy obszarem badań są ludzkie zachowania, kwestie etyczne nabierają szczególnego znaczenia. Ze względu na wyjątkowo osobisty, ważny i często krępujący charakter przedmiotu badań, w jeszcze większym stopniu niż w innych przypadkach dotyczy to badań nad ludzką seksualnością. Opinia publiczna i prawodawcy często wyrażają niepokój w związku z takimi badaniami. Trzy podstawowe zasady etyczne odnoszone są w szczególny sposób do badań nad ludzką seksualnością.

Świadoma zgoda. Zasada ta podkreśla znaczenie zarówno dokładnego poinformowania badanych czy respondentów o charakterze badań, jak i uzyskania ich ustnej lub pisemnej zgody na uczestniczenie w badaniach. Niedopuszczalna jest jakakolwiek forma przymusu, a badani mają możliwość wycofania się z udziału w każdej chwili. Norma ta bywa łamana na wiele sposobów. Może się zdarzyć, że badacz, obawiając się dużej liczby odmów lub fałszywych odpowiedzi, będzie przedstawiał temat badań, dotyczących spraw drażliwych lub osobistych, w sposób mylący lub fałszywy. Badania prowadzone w sposób niejawny (na przykład pewne typy obserwacji) także łamią zasadę świadomej zgody, gdyż badani w ogóle nie wiedzą, że są obserwowani. Kryterium świadomej zgody może stwarzać szczególne trudności dla badań prowadzonych w niektórych populacjach. Na przykład w przypadku badań seksualności dzieci, które mogą nie być – poznawczo i emocjonalnie – zdolne do wyrażenia świadomej zgody. Pomimo omówionych tu problemów większość badań przeprowadzana jest z zachowaniem pełnej dobrowolności, po wyrażeniu świadomej zgody przez uczestników.

Prawo do prywatności. Ze względu na bardzo osobisty charakter seksualności oraz ogromną troskę społeczeństwa o społeczną kontrolę seksualności, prawo do prywatności jest bardzo ważną normą etyczną w badaniach dotyczących tego obszaru. W przypadku ujawnienia pewnych aspektów swojego życia seksualnego badany może stracić pracę, mieć trudności rodzinne, bądź też może zostać poddany ostracyzmowi w swojej społeczności. Dotyczy to zwłaszcza osób, których zachowania seksualne postrzegane są jako dewiacje (na przykład transwestytów). Naruszenie prywatności ma miejsce, gdy badacz ujawnia tożsamość osób należących do określonej grupy objętej badaniami, ujawnia lub udostępnia dane lub odpowiedzi uzyskane od konkretnej osoby bądź w ukryty sposób obserwuje zachowania seksualne. Z reguły ochrona prywatności badanych nie powinna badaczowi sprawiać problemów. W badaniach sondażowych można zastosować ankiety wypełniane w sposób anonimowy samodzielnie przez badanych, w przypadku wywiadów zaś można zachować poufność. W przypadku obserwacji prawdziwa tożsamość badanych osób i grup może być w publikacjach zmieniona. W większości metod badawczych wyniki analizy danych powinny być podawane jedynie w postaci zagregowanej, na poziomie grupowym.

Ochrona badanych przed skrzywdzeniem. Szkoda może mieć charakter fizyczny, ale może to być także uraz psychiczny lub emocjonalny. Poszczególne metody badawcze różnią się poziomem ryzyka: jest ono większe w badaniach eksperymentalnych, gdzie badany poddawany jest jakiemuś działaniu czy manipulacji ze strony eksperymentatora, natomiast mniejsze w badaniach obserwacyjnych czy sondażowych. Urazy emocjonalne mogą być jednak skutkiem każdych badań dotyczących seksualności. Respondent może zostać postawiony wobec pytań wywołujących lęk, przywołujących przykre wspomnienia bądź też prowadzących do niekorzystnej samooceny. Badacz może ograniczyć ryzyko tego rodzaju komplikacji, stosując anonimowe, wypełniane samodzielnie przez badanych bądź za pośrednictwem specjalnie przeszkolonych ankieterów kwestionariusze, oraz przez szczególnie staranne formułowanie pytań dotyczących drażliwych kwestii.

Wszystkie trzy powyższe zasady są dość subiektywne. Ich łamanie bywa usprawiedliwiane korzyścią dla społeczeństwa, równoważącą ryzyko ponoszone przez badanych. Pozostaje oczywiście pytanie, kto ma podejmować tak ważne decyzje. Z reguły robią to badacze, często także komisje etyczne. Większość pomysłowych badaczy jest w stanie przestrzegać wszystkich tych trzech zasad etycznych i mimo to prowadzić znaczące badania.

Podobnie, nie należy ulegać pokusie przedstawiania wyników, uzyskanych przypadkiem, jako efektu starannie zaplanowanych strategii analitycznych. Wiele odkryć przychodzi nieoczekiwanie, mimo iż z perspektywy czasu wydają się oczywiste. A więc – jeśli na jakąś ciekawą zależność trafiliśmy przypadkiem – to co z tego? Dopisywanie do takiego wyniku fikcyjnych hipotez byłoby nieuczciwe, a w dodatku mogłoby mniej doświadczonych badaczy utwierdzić w przekonaniu, że wszystkie badania naukowe są precyzyjnie zaplanowane i zorganizowane.

Generalnie, postęp w nauce jest efektem uczciwości i otwartości; kłamstwa i próba obrony własnego „ja" jedynie go opóźniają. Najlepiej służy się społeczności naukowej – i nauce jako takiej – ujawniając prawdę o wszystkich pułapkach i problemach, które napotkało się na konkretnej drodze naukowej analizy. Być może pozwoli to uniknąć ich innym.

Instytucjonalne organa kontrolne

Wiele zasad etycznych dotyczących badań naukowych, których przedmiotem są ludzie, zostało zapisanych w USA jako normy prawa federalnego. Każda instytucja (szpital lub uniwersytet), ubiegająca się o federalne fundusze na sfinansowanie badań, musi powołać specjalną komisję etyczną, *Institutional Review Board* (IRB), w skład której wchodzą pracownicy naukowi (i inne osoby), mający za zadanie analizowanie wszystkich projektów badawczych przewidujących dokonywanie badań na ludziach, aby ocenić, czy w wystarczającym stopniu chronią one interesy i prawa badanych.

Głównym celem IRB jest upewnienie się, że ryzyko, na jakie narażeni będą uczestnicy badań, jest minimalne. W pewnych przypadkach IRB może zażądać poprawienia projektu; w innych może odmówić zatwierdzenia badań. Jeśli badania nieuchronnie niosą jakieś niewielkie ryzyko, badacz musi opracować informację dla potencjalnych uczestników w postaci formularza wyrażenia *„świadomej zgody"*, *w jasny sposób* opisującą wszystkie elementy ryzyka. W badaniach mogą wziąć udział wyłącznie osoby, które potwierdzą przez podpisanie tej informacji, że zapoznały się z nią i świadome ryzyka dobrowolnie godzą się na udział w badaniu.

Ustanowienie IRB było związane przede wszystkim z eksperymentami medycznymi na ludziach; znaczna część badań społecznych – takich jak na przykład anonimowe sondaże na dużych próbach – nie musi być poddawana ich ocenie. Zasadą, którą ma się kierować IRB, zgodnie z wykazem wyjątków, określonym prawem federalnym (45 CFR 46.101 [b]), jest to, czy badania „niosą ze sobą realne ryzyko narażenia badanego na odpowiedzialność karną lub cywilną, na szkody materialne lub ograniczenie zdolności do zarobkowania". Badania społeczne rzadko kiedy niosą dla badanych takie ryzyko – zwłaszcza jeśli przestrzega się zasady poufności.

Kodeksy etyki zawodowej

Problematyka etyczna badań społecznych jest tyleż ważna, co niejednoznaczna. Z tego względu większość organizacji zawodowych grupujących badaczy, tworzy i publikuje sformalizowane kodeksy etyczne, określające, jakie działania przystoją, a jakie nie przystoją profesjonalnym badaczom. Przykładem takiego kodeksu może być przedstawiony na rysunku 18.2 kodeks etyczny Amerykańskiego Stowarzyszenia Badania Opinii Publicznej (*American Association for Public Opinion Research* AAPOR), które jest interdyscyplinarnym stowarzyszeniem badaczy z obszaru nauk społecznych. Podobne kodeksy ma większość stowarzyszeń zawodowych – jak na przykład Amerykańskie Towarzystwo Socjologiczne (*American Sociological Association*), Amerykańskie Towarzystwo Psychologiczne (*American Psychological Association*), Amerykańskie Towarzystwo Nauk Politycznych (*American Political Science Association*) itd.; ich kodeksy etyczne możesz znaleźć w Internecie.

Dwie kontrowersje etyczne

Jak się zapewne domyślasz, z faktu, że opracowywane i publikowane są kodeksy etyczne, nie

Rysunek 18.2

Kodeks etyczny Amerykańskiego Stowarzyszenia *Badania Opinii Publicznej*

KODEKS ETYKI I PRAKTYKI ZAWODOWEJ

My, członkowie Amerykańskiego Stowarzyszenia Badania Opinii Publicznej, zobowiązujemy się przestrzegać zasad zapisanych w niniejszym kodeksie.

Naszym celem jest wspieranie zdrowych zasad postępowania w dziedzinie badania opinii publicznej. (Przez badania opinii publicznej rozumiemy takie badania, w których głównym źródłem informacji o przekonaniach, preferencjach i zachowaniach ludzi są ich własne deklaracje w tej sprawie).

Zobowiązujemy się do zachowywania wysokich standardów naukowej staranności w naszej pracy oraz w relacjach ze zleceniodawcami i ze społeczeństwem. Jednocześnie zobowiązujemy się nie przyjmować żadnych zleceń ani zobowiązań, które stałyby w sprzeczności z postanowieniami niniejszego kodeksu.

KODEKS

I. Zasady dotyczące prowadzenia prac badawczych

A. Będziemy zachowywać należytą staranność, zbierając i przetwarzając dane, podejmując wszelkie uzasadnione działania dla zapewnienia poprawności wyników.

B. Będziemy zachowywać należytą staranność, opracowując plany badań i analizując dane.

1. Będziemy stosować tylko takie narzędzia badawcze i metody analizy, które zgodnie z naszą kompetentną oceną są dobrze dobrane do danego problemu badawczego.
2. Nie będziemy dobierać narzędzi badawczych i metod analizy ze względu na większe prawdopodobieństwo uzyskania pożądanych przez nas wyników.
3. Nie będziemy świadomie formułowali interpretacji wyników badań, nie wynikających z zebranych danych, ani nie będziemy się milcząco godzili na takie interpretacje.
4. Nie będziemy świadomie przypisywali interpretacjom większego stopnia pewności, niż faktycznie wynikający z zebranych danych.

C. Uzyskane wyniki oraz zastosowane metody będziemy starannie i na odpowiednim poziomie szczegółowości prezentowali we wszystkich raportach z badań.

II. Zasady dotyczące odpowiedzialności zawodowej w naszych relacjach z ludźmi

A. Społeczeństwo

1. Będziemy współpracowali z uprawnionymi reprezentantami społeczeństwa poprzez podawanie metodologii przeprowadzanych badań.
2. Zachowujemy sobie prawo do wyrażania zgody na upublicznianie wyników naszych badań, niezależnie od tego czy będą one nam przypisane, czy też nie. Jeśli publikacje te zawierać będą nieuprawnione interpretacje, publicznie ujawnimy, co należy zrobić, by je skorygować, niezależnie od przyjętych wobec zleceniodawcy zobowiązań do zachowania poufności co do wszelkich innych aspektów badań.

B. Zleceniodawcy i sponsorzy

1. Będziemy zachowywać jako poufne wszelkie informacje, dotyczące działalności naszego klienta, a także wszelkie wyniki badań wykonanych na jego zlecenie, chyba że uzyskamy jednoznaczną zgodę na ich rozpowszechnianie.
2. Będziemy mieć na uwadze nasze ograniczenia techniczne i sprzętowe i będziemy przyjmowali tylko takie zlecenia, które można zrealizować przy tych ograniczeniach.

C. Społeczność zawodowa

1. Ponieważ Stowarzyszenie nie certyfikuje żadnych osób ani instytucji, nie będziemy podawali faktu członkostwa w Stowarzyszeniu dla potwierdzenia naszych kwalifikacji zawodowych.
2. Uznajemy naszą odpowiedzialność za przyczynianie się do rozwoju naukowego badań opinii publicznej oraz za upowszechnianie, w stopniu, w jakim jest to możliwe, idei i wyników, które wypracowaliśmy dzięki prowadzonym przez siebie badaniom.

D. Respondenci

1. Nie będziemy okłamywali naszych respondentów ani nie będziemy stosowali żadnych działań ani metod, prowadzących do wykorzystywania, poniżania lub stosowania przymusu wobec respondentów.
2. Będziemy chronili anonimowość wszystkich respondentów, z wyjątkiem przypadków, gdy respondent zwolni nas z tego obowiązku w określonym zakresie. Ponadto, wszystkie informacje mogące posłużyć do identyfikacji respondenta będziemy traktowali jako chronione i poufne.

Źródło: American Association for Public Opinion Research, *By-Laws*, maj 1977. Wykorzystano za pozwoleniem.

wynika jeszcze, że żadne kwestie etyczne w badaniach społecznych nie budzą już kontrowersji. Badacze w dalszym ciągu nie są w pełni zgodni co do niektórych ogólnych zasad, a tam, gdzie zgadzają się co do ogólnej zasady, często nadal dyskutują o szczegółach.

W tym podrozdziale przyjrzymy się dokładniej dwóm projektom badawczym, których etyczne aspekty były przedmiotem poważnych kontrowersji. Pierwszy z nich dotyczył zachowań homoseksualnych w toaletach publicznych, drugi to badania nad zjawiskiem posłuszeństwa, prowadzone w warunkach laboratoryjnych.

Kłopoty w kafejce

Jeszcze w czasie studiów Laud Humphreys zainteresował się badaniami zachowań homoseksualnych, a w szczególności badaniem sporadycznych i przelotnych kontaktów seksualnych z osobami tej samej płci, podejmowanych przez niektórych mężczyzn nie będących homoseksualistami. Jego badania skoncentrowały się na aktach homoseksualnych pomiędzy nieznajomymi sobie osobami, spotykającymi się w parkach w publicznych toaletach, nazwanych przez homoseksualistów „kafejkami" (*tearooms*). Efektem jego pracy była opublikowana w 1970 roku książka *Tearoom Trade*.

To, co szczególnie zainteresowało Humphreysa w działalności kafejek, to fakt, że mężczyźni zaangażowani w tego rodzaju akty poza tym sprawiali wrażenie osób prowadzących normalne życie rodzinne i akceptowanych w środowisku. Nie robili oni poza tym niczego innego, co mogłoby prowadzić do uznania ich za homoseksualistów, dlatego też było dla nich istotne, by ich wizyty w kafejkach miały charakter całkowicie anonimowy. W jaki sposób można by prowadzić nad nimi badania?

Humphreys postanowił wykorzystać społeczną strukturę sytuacji. Z reguły w kafejkowych spotkaniach brały udział trzy osoby: dwie, odbywające ze sobą stosunek, oraz trzecia, stojąca na czatach „czujka". Humphreys zaczął kręcić się w okolicach publicznych toalet, oferując swoje usługi jako ten trzeci zawsze, gdy wydawało się to potrzebne. Ponieważ wynagrodzeniem „czujki" była możliwość przyglądania się stosunkowi, Humphreys mógł prowadzić obserwacje terenowe analogiczne do tych, jakie mógłby prowadzić, gdyby badał zachowania na wiecach politycznych albo wzorce nieprawidłowego przechodzenia przez ulicę.

Aby móc lepiej zrozumieć mechanizmy wymiany, zachodzącej między uczestnikami spotkań w kafejkach, Humphreys potrzebował więcej informacji o swoich badanych. Ponieważ miałby bardzo niewielkie szanse na uzyskanie od nich wywiadu, zastosował inną strategię: gdy tylko mógł, odnotowywał numery rejestracyjne samochodów badanych, następnie korzystając z danych policji ustalał ich nazwiska i adresy, po czym, ucharakteryzowawszy się, by nie zostać rozpoznanym, odwiedzał ich w domach, deklarując, że przeprowadza badania opinii. W ten sposób Humphreysowi udało się uzyskać od badanych dotyczące ich informacje, których nie mógłby uzyskać w toaletach.

Jak można sobie wyobrazić, badania Humphreysa wywołały gwałtowny spór wśród naukowców, a także w szerszych kręgach opinii publicznej. Niektórzy krytycy zarzucali Humphreysowi przeprowadzenie, w imię nauki, ordynarnego zamachu na prywatność badanych. To, co robili mężczyźni w publicznych toaletach, było wyłącznie ich sprawą. Inni koncentrowali się na fakcie oszustwa – Humphreys okłamywał badanych, sugerując im, że jest jedynie podglądaczem. Nawet ci, którzy akceptowali obserwowanie uczestników schadzek w kafejkach – odbywających się ostatecznie w miejscach publicznych – protestowali przeciwko przeprowadzanej później ankiecie, gdyż, w ich opinii, nieetyczne było odszukiwanie badanych w domach i przeprowadzanie z nimi wywiadów pod fałszywym pozorem.

Byli jednak również tacy, którzy badania Humphreysa uznali za usprawiedliwione. Ich zdaniem temat wart był analizy. Nie mógł on być zbadany w żaden inny sposób, co do kłamstw zaś uznawali, że były w zasadzie nieszkodliwe dla badanych. Dodatkowo wskazywali na fakt, że Humphreys w pełni dbał o to, by nie zaszkodzić badanym poprzez ujawnienie ich kafejkowej działalności.

Kontrowersje wywołane przez Humphreysa nigdy nie zostały rozstrzygnięte. Są one wciąż

przedmiotem dyskusji – i zapewne zawsze będą, gdyż wywołują emocje i dotykają kwestii etycznych, co do których nie ma zgody. Co o tym wszystkim sądzisz? Czy postępowanie Humphreysa było etyczne? Czy może, twoim zdaniem, jakaś część jego badań była w porządku, a jakaś inna nie?

Obserwacja posłuszeństwa

Nasz drugi przykład różni się od pierwszego pod wieloma względami. Humphreys prowadził obserwację uczestniczącą, teraz zaś będziemy mówić o eksperymencie laboratoryjnym. Badania Humphreysa miały charakter socjologiczny, nasz drugi przykład to badania psychologiczne. Humphreys badał zachowania traktowane przez wielu jako dewiacyjne, natomiast badacz, z którym spotkamy się za chwilę, badał posłuszeństwo i konformizm.

Jednym z poruszających doświadczeń II wojny światowej było powszechne zjawisko tłumaczenia się niemieckich żołnierzy ze zbrodni za pomocą argumentu: „ja tylko wykonywałem rozkazy". Przyjmując tok rozumowania, prowadzący do formułowania takich usprawiedliwień, można by usprawiedliwić każde działanie – niezależnie od tego, jak nikczemne – jeśli tylko dałoby się za nie przerzucić odpowiedzialność na kogoś innego. Jeśli oficer nakazałby żołnierzowi zabicie dziecka – fakt działania na rozkaz zwalniałby jakoby żołnierza od jakiejkolwiek odpowiedzialności osobistej.

Chociaż sądy wojskowe, rozpatrujące sprawy o zbrodnie wojenne, nie uznawały takich usprawiedliwień, naukowcy zdali sobie sprawę, w jakim stopniu takie rozumowanie jest w życiu społecznym powszechne. Wydaje się, że często ludzie gotowi są robić rzeczy, o których wiedzą, że byłyby one uznane przez innych za niewłaściwe, jeśli mogą to uzasadnić poleceniem wyższej władzy. W ten sposób usprawiedliwiali się sprawcy tragedii w My Lai w Wietnamie, gdzie amerykańscy żołnierze zamordowali ponad 300 nieuzbrojonych cywilów – w tym także małe dzieci – tylko dlatego, że ich wioska uważana była za bazę Wietkongu. Tego rodzaju usprawiedliwienia spotyka się również w mniej dramatycznej formie w codziennym życiu w czasie pokoju. Chociaż w zasadzie mało kto podaje w wątpliwość samo zjawisko, eksperyment Stanleya Milgrama (1963, 1965) wywołał gwałtowną dyskusję.

Aby obserwować gotowość ludzi do krzywdzenia innych w następstwie słuchania rozkazów, Milgram zgromadził 40 dorosłych mężczyzn z wielu różnych środowisk w warunkach laboratoryjnych, zaprojektowanych tak, by odtwarzały badane zjawisko. Gdybyś był uczestnikiem tego eksperymentu, twoje doświadczenia byłyby mniej więcej takie.

Poinformowano by cię, że wraz z jeszcze jedną osobą macie wziąć udział w eksperymencie dotyczącym procesu uczenia się. W wyniku losowania ty masz być „nauczycielem", a twojemu towarzyszowi przypadła rola „ucznia". „Uczeń" jest prowadzony do drugiego pokoju, przypinany do krzesła, a do jego nadgarstków mocowane są elektrody. Ty, jako nauczyciel, zostajesz posadzony przed imponującym panelem pełnym przełączników, wskaźników i potencjometrów. Przy każdym wskaźniku widzisz tabliczkę, podającą w woltach inną wysokość napięcia prądu, od 15V do 315V. Dodatkowo przełączniki opisane są złowróżbnymi słowami, na przykład: „skrajne natężenie wstrząsu", „niebezpieczeństwo: silny wstrząs", „XXX".

Eksperyment ma następujący przebieg: odczytujesz „uczniowi" listy par słów, po czym sprawdzasz jego umiejętność łączenia ich ze sobą. Ponieważ nie masz z uczniem kontaktu wzrokowego, odpowiedź sygnalizowana jest przez zapalenie się odpowiedniej lampki na twoim panelu. Kiedy „uczeń" popełnia błąd, masz, zgodnie z instrukcją przekazaną przez eksperymentatora, nacisnąć jeden z przełączników, aplikując uczniowi wstrząs, począwszy od najłagodniejszego. Przez otwarte drzwi między pokojami słyszysz, jak uczeń reaguje na wstrząs elektryczny. Wtedy odczytujesz kolejną listę par słów i odpytujesz go dalej.

W miarę przebiegu eksperymentu aplikujesz coraz mocniejsze wstrząsy, aż w końcu uczeń zaczyna błagać o litość i o przerwanie eksperymentu. Otrzymujesz jednak polecenie zaaplikowania kolejnego wstrząsu. Po chwili uczeń, znowu krzycząc, zaczyna kopać w ścianę między pokojami. Nieporuszony tym eksperymentator

poleca ci zaaplikować jeszcze jeden wstrząs. W końcu, po przeczytaniu kolejnej listy, nie słyszysz żadnej odpowiedzi ucznia; z drugiego pokoju dobiega jedynie cisza. Eksperymentator mówi ci, że brak odpowiedzi należy traktować tak samo jak odpowiedź błędną, i poleca ci zaaplikować jeszcze silniejszy wstrząs. I tak dalej, do momentu gdy wstrząs „XXX" kończy serię.

Jak sądzisz, co byś naprawdę zrobił w momencie, gdy „uczeń" pierwszy raz zaczął krzyczeć? Kiedy zaczął kopać w ścianę? Albo kiedy kompletnie zamilkł i przestał dawać oznaki życia? Odmówiłbyś aplikowania kolejnych wstrząsów, prawda? Większość ludzi zapewne również by tak postąpiła.

Tak moglibyśmy sądzić – ale wynik uzyskany przez Milgrama był inny. Na pierwszych 40 badanych przez Milgrama dorosłych, żaden nie odmówił dalszego aplikowania wstrząsów aż do momentu, gdy „uczeń" zaczynał kopać w ścianę – wtedy uczyniło to pięć osób. Dwie trzecie badanych – 26 osób spośród 40 – wykonywało polecenia do końca całej serii, włącznie z zaaplikowaniem najsilniejszego wstrząsu.

Jak się zapewne domyślasz, wstrząsy były na niby, „uczeń" zaś był współpracownikiem eksperymentatora. Tylko „nauczyciel" był w tym eksperymencie prawdziwym badanym. Jako badany, w rzeczywistości nie robiłbyś nikomu krzywdy, ale zostałbyś postawiony w sytuacji, w której myślałbyś, że to robisz. Celem eksperymentu było sprawdzenie, czy byłbyś skłonny słuchać poleceń aż do momentu, gdy prawdopodobnie miałbyś kogoś zabić.

Wobec eksperymentu Milgrama stawiano zarzuty zarówno metodologiczne, jak i etyczne. Od strony etycznej koncentrowano się przede wszystkim na konsekwencjach, jakie dla badanych mogło mieć uczestnictwo w eksperymencie. Dla wielu z nich było to przeżycie niemal równie bolesne, jak elektrowstrząsy, które – jak sądzili – aplikowali komuś innemu. Błagali eksperymentatora, aby pozwolił im przestać. Byli skrajnie niezadowoleni i zdenerwowani. Niektórzy przestawali kontrolować swoje zachowanie.

Co sądzisz o tych badaniach? Czy ich przedmiot był, twoim zdaniem, na tyle ważny, że usprawiedliwiało to zastosowanie takich środków? Czy sesja z badanymi (*debriefing*) po zakończeniu eksperymentu wystarczyłaby, aby naprawić wszystkie szkody, jakie mogli ponieść? Czy umiałbyś wskazać inną metodę, którą badacz mógłby zastosować do badania posłuszeństwa?

Polityczne aspekty badań społecznych

Jak wspominałem już wcześniej, zarówno oceny etyczne, jak i polityczne zależą od przyjmowanych stanowisk ideologicznych. To, co dla jednych jest dopuszczalne, dla innych może dopuszczalne nie być. Choć jednak kwestie polityczne i etyczne są często ściśle ze sobą powiązane, można wskazać między nimi dwie istotne różnice.

Po pierwsze, etyka badań społecznych dotyczy przede wszystkim zastosowanych metod badawczych, natomiast kwestie polityczne koncentrują się na ogół na przedmiocie badań oraz ich zastosowaniach. Dla przykładu: część krytyków zarzucała Milgramowi, że jego eksperyment był nieetyczny, ponieważ zastosowane metody szkodziły badanym. Zarzuty o charakterze politycznym dotyczyłyby tego, że posłuszeństwo nie jest odpowiednim tematem badań, gdyż (1) nie należy podważać gotowości ludzi do okazywania posłuszeństwa władzy, albo też, z przeciwnych pozycji politycznych, (2) ponieważ wyniki takich badań mogłyby zostać wykorzystane do nakłaniania ludzi do *większego* posłuszeństwa.

Druga różnica między etycznymi a politycznymi aspektami badań społecznych wiąże się z faktem, że nie istnieją sformalizowane kodeksy określające, jakie działania są, a jakie nie są politycznie dopuszczalne. Mimo że niektóre normy etyczne mają powiązania z polityką – na przykład zasada, że badania nie mogą szkodzić badanym, jest bezpośrednio związana z ochroną wolności obywatelskich – nikt nie sformułował zbioru politycznych norm postępowania, który byłby powszechnie akceptowany przez badaczy społecznych.

Do pewnego stopnia wyjątkiem jest jedynie powszechnie przyjmowane przekonanie, że osobiste poglądy polityczne badacza nie po-

winny mieć wpływu na prowadzone przez niego badania naukowe. Niedopuszczalne jest także, by nieściśle lub fałszywie przedstawiać badania bądź przeprowadzać je w sposób niezgodny z zasadami sztuki, w celu forsowania własnych poglądów politycznych. Jednak, jak się zapewne domyślasz, zarzuty takie wobec różnych badań formułowane są dość często.

Obiektywność a ideologia

Jak pisałem w rozdziale 1, badania społeczne nigdy nie mogą być w pełni obiektywne: badacze są tylko ludźmi i jako tacy z konieczności są subiektywni. Odpowiadającą obiektywności cechą nauki, jako przedsięwzięcia zbiorowego, jest intersubiektywność. Oznacza ona, że różni uczeni, mający różne subiektywne poglądy, jeśli stosują uznane techniki badawcze, mogą i powinni dochodzić do takich samych wyników. W istocie rzeczy jest to możliwe, jeśli poszczególni naukowcy w toku badań potrafią abstrahować od swoich osobistych wartości i poglądów.

Klasyczne ujęcie obiektywności i neutralności w naukach społecznych zawarte zostało w wykładzie Maxa Webera *Nauka jako zawód i powołanie* (1998). Weber sformułował w nim postulat *socjologii wolnej od sądów wartościujących*, stwierdzając, że socjologia, jeśli jej uprawianie ma przynosić społeczeństwu jakąś korzyść, musi, tak jak inne nauki, być wolna od osobistych wartości wyznawanych przez różnych uczonych. Liberałowie i konserwatyści mogą uznawać te same „fakty" nauk społecznych, niezależnie od tego, jak się owe fakty mają do ich osobistych poglądów politycznych.

Powyższy abstrakcyjny ideał akceptuje większość badaczy społecznych, ale nie wszyscy. Uczeni marksistowscy i neomarksistowscy na przykład dowodzą, że nie da się i nie powinno się próbować rozdzielić nauk społecznych od działań społecznych. Wyjaśnianie źródeł społecznego *status quo*, jak twierdzą, może po cichu przeradzać się w obronę tego *status quo*. Zwykłe wyjaśnianie, powiedzmy, społecznych funkcji dyskryminacji jakichś grup, łatwo może się zamienić w usprawiedliwianie jej trwania. Na tej samej zasadzie samo badanie społeczeństwa i jego problemów, bez angażowania się w czynienie go bardziej ludzkim, nazywane jest działaniem nieodpowiedzialnym.

Niezależnie od abstrakcyjnych dyskusji o tym, czy nauki społeczne mogą, bądź też czy *powinny być* wolne od sądów wartościujących, wielokrotnie spierano się o to, czy konkretne przedsięwzięcia badawcze *są* wolne od wartościowania, czy też są obciążone wpływem wartości politycznych, wyznawanych przez ich autorów. Z reguły badacze zarzuty takie odrzucali, co jednak nie przekonywało tych, którzy je sformułowali. Przyjrzyjmy się dokładniej kilku przykładom kontrowersji, które rozgorzały wokół tej kwestii, i trwają nadal.

Badania społeczne a kwestie rasowe

Nigdzie indziej badania społeczne nie są w bardziej kontrowersyjny sposób splecione z polityką niż w obszarze stosunków rasowych. Problematyka ta jest od dawna przedmiotem badań społecznych, a wyniki badań często wykorzystywane są w praktyce politycznej – co pokażę na kilku przykładach.

Kiedy w 1896 roku Sąd Najwyższy Stanów Zjednoczonych ustanowił zasadę „odrębni ale równi", mającą pogodzić gwarantowaną przez Czternastą Poprawkę do Konstytucji równość Afroamerykanów z zasadami segregacji rasowej, to ani nie domagał się on żadnych badań społecznych, ani się na nie nie powoływał. Powszechnie uważa się jednak, że na orzeczenie sądu miały wpływ prace jednego z czołowych przedstawicieli nauk społecznych tamtych czasów, Williama Grahama Sumnera. Sumner wsławił się poglądem, że obyczaje i tradycje społeczne są w znacznym stopniu odporne na wpływ ze strony prawa i planowania społecznego. Jego opinia była często parafrazowana w postaci zdania *stateways do not make folkways*[2]. Tak więc sąd uznał, że nie może zaakceptować stanowiska, jakoby „uprzedzenia społeczne można było zwalczyć za pomocą prawa" i za-

[2] Dosłownie „obyczaje państwa nie czynią obyczajów [ludu]". *Stateways* – neologizm wywiedziony od *folkways* (obyczaje), w którym cząstkę *folk* – „lud" zamieniono na *state* – „państwo" (przyp. tłum.).

negował mądrość „praw stojących w sprzeczności z powszechnymi w społeczeństwie odczuciami" (*Civil Rights...* 1970, s. 308). Jak twierdzi wielu polityków, „nie można ustawą wprowadzać moralności".

Kiedy zasada „odrębni ale równi" została w 1954 roku odrzucona (sprawa Brown przeciwko Radzie Szkolnej w Topeka), nowe orzeczenie Sądu Najwyższego opierało się w części na stwierdzeniu, że segregacja jest szkodliwa dla afroamerykańskich dzieci. Opinię tę sąd uzasadniał cytując wiele raportów z badań socjologicznych i psychologicznych (*Civil Rights... 1970*).

Większość amerykańskich badaczy w XX wieku wspierała sprawę równouprawnienia Murzynów i często właśnie ich przekonania skłaniały ich do podejmowania badań. Co więcej, mieli oni nadzieję, że badania te doprowadzą do zmian społecznych. Dla przykładu – nie ma wątpliwości, że klasyczne, dwutomowe dzieło Gunnara Myrdala (1944) o stosunkach rasowych w USA miało istotny wpływ na przedmiot jego badań. Myrdal zgromadził ogromną ilość danych, aby pokazać, że sytuacja Afroamerykanów jest całkowicie sprzeczna z amerykańskimi wartościami równości społecznej i politycznej. Myrdal nawet nie próbował ukryć swojego osobistego stosunku do tych kwestii.

Wielu badaczy społecznych było bezpośrednio zaangażowanych w mniej lub bardziej radykalne ruchy obrony praw obywatelskich. Badania, których wyniki stanowiły wsparcie dla sprawy równości, w nikłym stopniu poddawane były krytyce lub w ogóle ich nie krytykowano. Aby zdać sobie sprawę, jak jednolite było w tych kwestiach ogólne stanowisko środowisk naukowych, wystarczy przyjrzeć się kilku projektom badawczym, których wyniki stały w sprzeczności z dominującym stanowiskiem ideologicznym.

Naukowcy zajmujący się badaniami społecznymi w większości – przynajmniej jawnie – wspierali zniesienie segregacji rasowej w szkołach. W 1966 roku gwałtowne dyskusje wywołało opublikowanie przez poważanego socjologa, Jamesa Colemana, wyników zakrojonych na dużą skalę ogólnokrajowych badań dotyczących kwestii rasowych w edukacji. W przeciwieństwie do innych badań, wyniki Colemana wskazywały na istnienie niewielkich różnic między osiągnięciami szkolnymi afroamerykańskich uczniów uczęszczających do szkół integracyjnych a osiągnięciami afroamerykańskich uczniów chodzących do szkół stosujących segregację. Wynikało z tego, że takie oczywiste czynniki jak biblioteki, wyposażenie pracowni, wysokość nakładów na jednego ucznia miały niewielkie znaczenie – tym, co według Colemana miało największy wpływ na osiągnięcia w nauce, były czynniki związane z sytuacją rodzinną i bezpośrednim otoczeniem ucznia.

Wyniki badań Colemana zostały źle przyjęte przez wielu badaczy społecznych, działających w ruchach obrony praw obywatelskich. Niektórzy uczeni stawiali pracy Colemana zarzuty o charakterze metodologicznym, ale wielu innych zgłaszało swój gorący sprzeciw, argumentując, że jej wyniki będą wykorzystywane jako polityczny argument na rzecz segregacji. Spór, jaki wybuchł wokół badań Colemana, był przypomnieniem sporu, jaki rok wcześniej został sprowokowany przez krytyczną analizę funkcjonowania afroamerykańskich rodzin w USA autorstwa Daniela Moynihana (1965).

Innym przykładem badań społecznych, związanych z przynależnością rasową, które wywołały kontrowersje polityczne, były badania dotyczące wyników testów inteligencji. W 1969 roku psycholog z Harvardu, Arthur Jensen został poproszony o przygotowanie dla „Harvard Educational Review" artykułu o rasowym zróżnicowaniu wyników testów IQ. W opublikowanym tekście (1969) Jensen doszedł do wniosku, że na niższe wyniki testów IQ Afroamerykanów wpływ mają różnice genetyczne między nimi a białymi. Jensen był tak utożsamiany z tym stanowiskiem, że przedstawiał je później w licznych dyskusjach odbywających się w campusach uniwersyteckich w całym kraju.

Badania Jensena były wielokrotnie atakowane z różnych pozycji metodologicznych. Zarzucano im, że dane, na których się opierają, były nieadekwatne i niechlujne – różne testy IQ mają różną wartość. Zarzucano też Jensenowi nieuwzględnienie w wystarczającym stopniu czynników związanych ze środowiskiem społecznym badanych. Formułowano także inne zastrzeżenia metodologiczne.

Jednak niezależnie od krytyki o charakterze naukowym, wielu potępiło Jensena jako rasistę. Był wygwizdywany, a jego publiczne wykłady były zagłuszane przez wrogo nastawiony tłum. Ironia losu sprawiła, że na wielu uniwersytetach Jensen był traktowany przez publiczność podobnie jak abolicjoniści sto lat wcześniej, gdy większość opinii publicznej była za utrzymaniem instytucji niewolnictwa.

Wielu badaczy społecznych ograniczało swoją krytykę badań Moynihana, Colemana czy Jensena jedynie do zarzutów naukowych, metodologicznych. Jednak wywołane przez te badania polityczne burze pokazują, w jaki sposób ideologia często może ingerować w badania społeczne. Choć teoretyczny model nauki zakłada jej oddzielenie od ideologii, praktyka bywa inna.

Projekt Camelot

Gdy w dyskusjach o związkach między badaniami a polityką, prowadzonych przez badaczy społecznych, pojawia się nazwa *Camelot*, wszyscy wiedzą, o co chodzi. Irving Louis Horowitz (1967, s. VI), człowiek, który od czasu do czasu krytykował instytucje rządowe, stwierdził, że projekt *Camelot* „był chyba największą wpadką w dziedzinie *public relations*, jaka przydarzyła się jakiejkolwiek rządowej instytucji Stanów Zjednoczonych". Skąd wzięło się takie zamieszanie?

Czwartego grudnia 1964 roku Biuro Badawcze do Zadań Specjalnych, działające przy American University, rozesłało do grupy badaczy społecznych ogłoszenie o przygotowywanym projekcie badawczym dotyczącym wojny domowej. W informacji tej znalazł się między innymi następujący opis:

> „Celem projektu *Camelot* jest określenie możliwości wypracowania ogólnego modelu systemów społecznych, który pozwalałby przewidywać i wpływać na politycznie istotne aspekty zmiany społecznej w krajach rozwijających się na całym świecie. W szczególności, celem projektu jest:
>
> *Po pierwsze*, stworzenie procedur oceny ryzyka wojny domowej w poszczególnych krajach.
>
> *Po drugie*, określenie, z możliwie dużym stopniem pewności, działań, które może podjąć rząd, aby złagodzić warunki oceniane jako powodujące ryzyko wojny domowej".
>
> (Horowitz 1967, s. 47)

Oczywiście, niewielu jest ludzi, którzy otwarcie sprzyjają wojnom; większość skłonna byłaby poprzeć badania, mające za cel zakończenie wojny lub zapobieganie jej. Jednakże w lecie 1965 roku, w czasie gdy narodowa debata dotycząca wojny w Wietnamie właśnie nabierała rozpędu, *Camelot* był w kręgach nauk społecznych gorąco krytykowany jako próba wciągnięcia badaczy przez Departament Obrony w wysiłki podejmowane dla zdławienia rebelii w Chile. Pojawiały się opinie, że Departament Obrony chciał zlecić badania, które mogłyby pomóc w stłumieniu nastrojów politycznego niezadowolenia, mogących doprowadzić do rewolucji w tym niestabilnym kraju Ameryki Łacińskiej. Niezależnie od osobistych motywacji, badacze obawiali się, że ich praca mogłaby zostać wykorzystana do umacniania istniejących reżimów i zwalczania reformatorów ludowych oraz ruchów rewolucyjnych w innych państwach.

Wielu badaczy, którzy w zasadzie zgodzili się na uczestnictwo w projekcie, szybko poczuło, że będą musieli przejść przez to, przez co kilkadziesiąt lat wcześniej przeszedł Robert Oppenheimer i niektórzy inni naukowcy, zaangażowani w projekt budowy bomby atomowej: odkrycia naukowe mogą być wykorzystywane do celów, których ich autorzy nie akceptują. W środowisku zawodowym zaczęły się pojawiać oskarżenia, na które odpowiadano kontroskarżeniami, padały wyzwiska, kwestionowano motywy postępowania, zrywano stare przyjaźnie. Departament Obrony był powszechnie krytykowany za próbę zdemoralizowania nauk społecznych. Stosunki międzynarodowe z krajami Ameryki Południowej uległy gwałtownemu ochłodzeniu i stały się napięte. Ostatecznie, w obliczu rosnącej krytyki, z realizacji projektu *Camelot* zrezygnowano.

Warto się zastanowić, jakie byłyby losy projektu *Camelot*, gdyby propozycja jego realizacji trafiła na społeczność badawczą, składającą się z zatwardziałych konserwatystów i antykomunistów. Nie mam wątpliwości, że w takich wa-

runkach projekt byłby zaakceptowany i zrealizowany bez specjalnych problemów czy wątpliwości. Z pewnością w całej tej sprawie nie chodziło o wojnę. Nikt nie krytykował Samuela Stouffera, który w czasie II wojny światowej zorganizował ośrodek badawczy armii, prowadzący badania mające zwiększyć skuteczność amerykańskich żołnierzy. Sama nauka jest w kwestiach wojny i pokoju neutralna – to naukowcy nie są.

Na zakończenie, zwróćmy jeszcze uwagę na pewną obserwację, dotyczącą etyki badań widzianej z perspektywy krajów, będących czasami przedmiotem badań amerykańskich uczonych. Hamnett ze współpracownikami zauważyli między innymi:

> „Rządy wielu krajów Trzeciego Świata stawiają badaczom prowadzącym badania międzynarodowe coraz większe wymagania, poczynając od ograniczeń nakładanych na badania, nie odnoszące się bezpośrednio do priorytetów rozwoju kraju, aż do wymogu współpracy z lokalnymi instytucjami i naukowcami. Żądania takie można uzasadnić poglądem, że dane społeczne czy kulturowe stanowią dobro narodowe, którego eksploatacja musi przynosić danemu krajowi jakieś korzyści".
>
> (Hamnett i in. 1984, s. 6)

Polityka przez małe „p"

Często okazuje się, że badania społeczne są powiązane z ideologiami politycznymi, jednak kwestie „polityczne" w badaniach społecznych sięgają znacznie głębiej. Badania odnoszące się do spornych kwestii społecznych po prostu nie mogą być w pełni obiektywne, zwłaszcza jeśli ich wyniki mogą służyć jako argumenty w sporach ideologicznych. To samo dotyczy także sytuacji, gdy badania są przytaczane jako argument w sytuacji sprzecznych interesów. Na przykład badacz powołany na biegłego w sądzie przekona się, że naukowy ideał „poszukiwania prawdy" w warunkach rozprawy sądowej jest beznadziejnie naiwny. Formalnie biegły nie reprezentuje w sądzie żadnej ze stron, jest jednak powołany na wniosek i za pieniądze jednej z nich, i zapewne jego opinia będzie dla tej strony korzystna. Nie musi to oznaczać, że taki świadek będzie kłamał na korzyść strony, która go wezwała, ale po prostu strony w postępowaniu sądowym mają większą skłonność do opłacania raczej ekspertyz, które są korzystne dla ich sprawy, niż takich, które ją kwestionują.

Tak więc zostałeś wezwany do sądu jako ekspert tylko z powodu przypadkowej zbieżności pomiędzy twoim uczciwym, naukowo motywowanym osądem a interesem strony, która płaci za twoją opinię. Kiedy stawisz się w sądzie i przysięgniesz mówić prawdę, całą prawdę i tylko prawdę, znajdziesz się nagle w świecie, któremu obcy jest ideał kontemplacji prawdy. Nagle ważne jest tylko to, co prowadzi do wygranej lub porażki. Oczywiście, jako biegły, możesz stracić co najwyżej prestiż (i ewentualnie szanse na powoływanie na biegłego w przyszłości oraz związane z tym wynagrodzenie), ale jest to wystarczająco wysoka stawka, by wywołać dyskomfort u większości badaczy społecznych.

Wspomnę tylko o jednej sprawie, w której przed federalnym sądem przedstawiałem swoją opinię. Zostałem powołany przez grupę urzędników służby cywilnej, którym obcięto dodatki do pensji na wyrównanie wzrostu kosztów utrzymania, uzasadniając to wynikami mocno wątpliwych, według mnie, badań. Moim zadaniem było przeprowadzenie bardziej „naukowych" badań, które wykazałyby, że urzędnicy zostali potraktowani niesprawiedliwie (Babbie 1982, s. 232–243).

Kiedy stanąłem przed sądem, uważałem się za poważanego profesora i autora podręczników, ale bardzo szybko przekonałem się, że przeniosłem się z uniwersytetu na boisko hokejowe. Testy istotności statystycznej czy błędy losowe okazały się nagle mniej istotne niż mocne uderzenia. Tylko chwilę zajęło adwokatowi z Waszyngtonu wyciągnięcie ode mnie, że słyszałem o pewnym piśmie naukowym, które niestety nigdy nie istniało. Byłem upokorzony, i nagle poczułem, że zmieniam się w kogoś innego: niepostrzeżenie dla samego siebie, w coraz mniejszym stopniu pozostawałem łagodnym i potulnym obywatelem, a coraz wyraźniej budził się we mnie profesor Ninja. Poczułem, że teraz z kolei ja muszę upokorzyć adwokata, i nawet mi się to udało.

Chociaż urzędnicy dostali swoje pieniądze z powrotem, muszę przyznać, że miało dla mnie istotne znaczenie, jak ja sam wypadłem przed sądem. Opowiedziałem tę historyjkę, żeby pokazać, iż na obiektywne, naukowe badania istotny wpływ ma także „polityka" relacji międzyludzkich. Musimy sobie zdawać sprawę, że badacze społeczni są tylko ludźmi i zachowują się jak inni ludzie – i że oceniając ich pracę musimy brać to pod uwagę. To oczywiście samo w sobie nie dyskwalifikuje wyników badań ani nie usprawiedliwia negowania ich tyko dlatego, że nam się nie podobają – ale jednak jest to czynnik, o którym trzeba pamiętać.

Polityka we właściwych proporcjach

Choć etyczny i polityczny wymiar badań są zasadniczo odrębne, to jednak są one ze sobą powiązane. Kiedy opinia publiczna lub politycy uznają, że badania społeczne naruszają standardy etyczne, nie będą zwlekali z podjęciem działań, mających uzdrowić sytuację. Może się przy tym zdarzyć, że pójdzie akurat o standardy, których społeczność naukowa nie uważa za własne. Może być i tak, że chociaż badacze popierają cele działań skierowanych przeciwko jakimś niewłaściwym praktykom badawczym, to środki przewidziane w regulacjach i przepisach mogą uniemożliwić realizację badań.

Przykład czegoś takiego można znaleźć w ramce „Ustawa Senatu nr 1126", zawierającej treść ustawy przyjętej w 1995 roku przez Senat stanu Wirginia. Jej celem było ograniczenie nadużyć, do jakich dochodziło przy prowadzeniu sondaży przedwyborczych. Ustawa dotyka rzeczywistego problemu: możliwego wpływu ośrodka finansującego badania na obiektywizm badaczy. Jeśli pamiętasz, co wcześniej mówiliśmy o „sondażowym przepychaniu" (*push-polls*, czyli pseudosondażach, których rzeczywistym celem jest wpłynięcie na opinię respondentów), to zdajesz sobie sprawę z możliwych zagrożeń. Ale szczegółowe regulacje, takie jak zawarte w omawianej ustawie, mogą przy okazji prób usuwania jednych problemów wywołać inne. Na przykład rutynowe podawanie respondentom, kto jest zleceniodawcą sondażu, również może mieć obciążający wpływ na ich odpowiedzi. (Respondent ma prawo wiedzieć, kto zlecił badania, po zakończeniu wywiadu; oczywiście może także w ogóle odmówić udziału w badaniach).

Szczególnym obszarem zainteresowań prawodawców są badania prowadzone na dzieciach. Chociaż zasady etyczne badań społecznych, przedstawione w tym rozdziale, nie dopuszczają prowadzenia badań, które mogłyby przynieść dzieciom jakiekolwiek szkody fizyczne lub emocjonalne, to od czasu do czasu w różnych miejscach pojawiają się regulacje prawne, które zbliżają się do działań w pewnym mieście na zachodzie USA (nazwy nie wymienię). W reakcji na sygnały, że pewien nauczyciel w publicznej szkole puszczał dzieciom w czasie lekcji muzykę New Age i namawiał je do medytacji, rada miejska uchwaliła przepis zakazujący nauczycielom robienia czegokolwiek, co „miałoby wpływ na umysły uczniów!".

Chciałbym, żebyś z powyższych rozważań wyciągnął trzy główne wnioski. Po pierwsze, nauka nigdy nie jest uprawiana w politycznej próżni. Wpływ polityki i politycznych ideologii nie ogranicza się do badań społecznych – podobne doświadczenia dotyczyły i dotyczą także nauk przyrodniczych – ale sytuacja nauk społecznych jest szczególna: są one częścią życia społecznego. Przedmiotem badań społecznych są problemy, które ludzie uważają za ważne, do których mają ugruntowany, osobisty stosunek, i które mogą mieć bezpośredni wpływ na ich sprawy. Jednocześnie ludźmi są także badacze – i ich postawy często przenikają do ich działalności zawodowej. Negowanie tego faktu byłoby naiwnością.

Po drugie, uprawianie nauki może zawieść w sam środek gwałtownego politycznego sporu. Ale naukowe poszukiwania nie ustają nawet wtedy, gdy napięcia między badaczami doprowadzają do wymiany obelg, bądź też gdy społeczność naukowa staje się obiektem ataków z zewnątrz. Nawet w takich warunkach prowadzone są badania, publikowane ich wyniki, dokonywane nowe odkrycia. Mówiąc w skrócie – spory ideologiczne nie powstrzymują nauki, wzmagają za to współzawodnictwo i czynią naukę jeszcze bardziej ekscytującą.

Ustawa Senatu nr 1126

24 lutego 1995

Ustawa wprowadzająca poprawkę do Kodeksu Wirginii, polegającą na dodaniu sekcji 24.2–1014.1, dotyczącej sondaży telefonicznych w okresach kampanii wyborczych; kary.

Sprawozdawca – Robb

z inicjatywy Gubernatora

skierowano do Komisji Immunitetów i Wyborów

Zgromadzenie stanu Wirginia uchwaliło, co następuje:

1. Uzupełnia się Kodeks Wirginii poprzez dodanie sekcji 24.2–1014.1 o następującej treści:

§ 24.2–1014.1 Określenie osób odpowiedzialnych za niektóre sondaże telefoniczne: kary

A. Gdziekolwiek w tej sekcji używa się terminu „sondaż telefoniczny", należy przez to rozumieć serię rozmów telefonicznych, przeprowadzonych (i) z co najmniej 25 numerami na terenie stanu, (ii) w okresie 180 dni przed wyborami powszechnymi lub specjalnymi bądź w okresie 90 dni przed prawyborami, (iii) w celu zebrania informacji, dotyczących opinii publicznej lub preferencji wobec kandydatów w wyborach lub prawyborach.

B. Przeprowadzanie przez kogokolwiek sondaży telefonicznych będzie uznane za nielegalne, jeśli przed zakończeniem rozmowy telefonicznej respondent nie zostanie poinformowany o tym, kto zlecił lub sfinansował sondaż. Osoba, prowadząca wywiad telefoniczny, ujawni następujące informacje:

1. Nazwisko kandydata, jeśli badania były zlecone przez kandydata lub jego komitet wyborczy.
2. Nazwę komitetu wyborczego partii politycznej, jeśli sondaż był zlecony przez komitet wyborczy partii politycznej.
3. Nazwę komitetu, grupy lub dane osoby zlecającej badania, jeśli badania nie były zlecone przez kandydata, jego komitet wyborczy ani komitet wyborczy partii politycznej. Jeśli respondent domaga się dodatkowych informacji o tym komitecie, grupie lub osobie, osoba prowadząca wywiad powinna podać numer rejestracyjny, nadawany przez władze stanu komitetom lub osobom, które zarejestrowały się jako organizacje zgodnie z przepisem 24.2–9098, a jeżeli komitet lub osoba nie rejestrowała się, pełne nazwisko i adres zamieszkania osoby, która zleciła lub sfinansowała sondaż.

C. Zaniechanie przez osobę nadzorującą realizację sondażu telefonicznego dostarczenia osobom przeprowadzającym wywiady informacji wymaganych w niniejszej sekcji, będzie uznawane za nielegalne.

D. Podawanie komukolwiek fałszywych lub fikcyjnych danych identyfikacyjnych, wymaganych w podsekcji B, będzie uznawane za nielegalne.

E. Ktokolwiek złamie jakikolwiek przepis niniejszej sekcji, podlega karze grzywny do pięćdziesięciu dolarów, a gdyby było to działanie celowe, winny będzie wykroczenia I stopnia. Procedura nałożenia grzywny za naruszenia przepisów tej sekcji jest taka, jak określono w przepisie 24.2–929. Naruszenie przepisów tej sekcji nie stanowi podstawy do unieważnienia jakichkolwiek wyborów.

I wreszcie, świadomość istnienia aspektów ideologicznych wzbogaca teorię i praktykę metodologii nauk społecznych. Wiele technik badawczych wymyślono po to, by zrównoważyć lub kontrolować nasze ludzkie ograniczenia – zwłaszcza te, z których nie zdajemy sobie sprawy. Gdyby nie one, moglibyśmy, obserwując świat, dostrzegać jedynie odbicie naszych własnych iluzji i przekonań.

Główne zagadnienia

- Planując badania społeczne, należy wziąć pod uwagę nie tylko ich aspekty naukowe i techniczne, ale również organizacyjne, etyczne i polityczne.
- Co w badaniach jest etycznie dopuszczalne, a co niedopuszczalne, jest w ostatecz-

nym rachunku kwestią tego, co w danej społeczności uważane jest za dobro i zło.

- Badacze zgodni są co do tego, że zasadniczo udział w badaniach powinien być dobrowolny. Jednak norma ta w pewnych sytuacjach pozostaje w sprzeczności z naukową potrzebą uogólniania wyników.
- Badacze są zgodni co do tego, że badania nie mogą szkodzić badanym, chyba że w pełni poinformowani badani dobrowolnie zaakceptują ryzyko, wiążące się z udziałem w badaniach, wyrażając na nie swoją świadomą zgodę.
- Anonimowość odnosi się do sytuacji, w której nawet osoby prowadzące badania nie mogą powiązać uzyskanych danych z konkretnymi uczestnikami badań. Natomiast poufność odnosi się do sytuacji, w której badacz obiecuje uczestnikom, że nie ujawni informacji ich dotyczących. Najprostszą metodą zapewnienia poufności jest zniszczenie danych, pozwalających na identyfikację badanych od razu, kiedy przestaną być potrzebne.
- Wiele projektów badawczych zakłada, w szerszym lub węższym zakresie, wprowadzanie badanych w błąd. Ponieważ oszukiwanie ludzi jest sprzeczne z ogólnymi zasadami etyki, działanie takie dopuszczalne jest tylko z bardzo poważnych powodów; zawsze jednak należy się liczyć z tym, że jego dopuszczalność etyczna będzie podawana w wątpliwość.
- Badacze mają moralne zobowiązania nie tylko wobec badanych, ale i wobec społeczności badawczej. Należy do nich m.in. obowiązek pełnego i ścisłego przedstawiania wyników badań, a także ujawnianie ich błędów, ograniczeń i wszelkich innych niedoskonałości.
- Stowarzyszenia zawodowe w różnych dziedzinach ogłaszają kodeksy etyczne zawierające wskazówki dla badaczy. Choć kodeksy takie są potrzebne i pomocne, nie rozwiązują one wszystkich problemów etycznych.
- Badania Lauda Humphreysa dotyczące spotkań w „kafejkach" oraz badania nad posłuszeństwem Stanleya Milgrama wywołały trwające do dziś dyskusje nad związanymi z nimi problemami etycznymi.
- Badania społeczne nieuchronnie mają także wymiar polityczny i ideologiczny. Choć nauka jest neutralna politycznie, to nie dotyczy to samych naukowców. Co więcej, wiele badań społecznych nieuchronnie wiąże się z przekonaniami politycznymi ludzi, nie należących do społeczności naukowej.
- Chociaż większość badaczy zgodna jest co do tego, że przekonania polityczne nie powinny wpływać na uprawianie nauki, w praktyce oddzielenie polityki i ideologii od działalności badawczej może być bardzo trudne. Niektórzy badacze uważają, że badania mogą i powinny być instrumentem działalności na rzecz zmian społecznych. Bardziej subtelne efekty dotyczą tego, w jaki sposób wspólnie wyznawana ideologia może wpływać na odbiór badań przez innych badaczy.
- Chociaż zasady prowadzenia działalności naukowej nie mogą zmusić badaczy do porzucenia osobistych systemów wartości, intersubiektwny charakter nauki gwarantuje, że nie zostaną uznane za „naukowe" wyniki będące jedynie projekcją indywidualnych obciążeń.

Kluczowe pojęcia

świadoma zgoda
anonimowość
poufność
sesja poeksperymentalna (*debriefing*)

Pytania i ćwiczenia

1. Przyjrzyj się następującym rzeczywistym i hipotetycznym sytuacjom badawczym. Jakie są etyczne aspekty każdego z przykładów? Co o nich sądzisz? Czy twoim zdaniem opisane działania są ostatecznie akceptowalne czy nieakceptowalne? Być może łatwiej będzie odpowiedzieć na te pytania po przedyskutowaniu ich w gronie kolegów.
   a. Osoba, prowadząca wstępne zajęcia z psychologii, prosi swoich studentów o wypełnienie kwestionariuszy, których

analizę wykorzysta, pisząc artykuł do czasopisma naukowego.

b. Prokurator żąda od badacza, który prowadził badania dotyczące zachowania ludzi w czasie zamieszek ulicznych, ujawnienia danych osobowych tych badanych, którzy dopuszczali się plądrowania sklepów. Badacz, nie chcąc ryzykować aresztowania za współudział, ujawnia te dane.

c. Po skończeniu pierwszej wersji książki prezentującej wyniki badań badacz odkrywa, że 25 spośród 2000 wywiadów zostało sfałszowanych przez ankieterów. Aby uniknąć kwestionowania wyników całych badań, pomija tę informację i publikuje książkę.

d. Badacze uzyskują listę prawicowych ekstremistów, których chcą badać. Kontaktują się z nimi, wyjaśniając, że zostali dobrani „losowo" z całej populacji w celu zbadania próby „opinii publicznej".

e. Wykładowca w college'u, chcąc zweryfikować hipotezę dotyczącą konsekwencji niesprawiedliwego traktowania studentów, przeprowadza godzinny test dla dwóch grup studentów uczęszczających na jego zajęcia. Studenci w obu grupach uzyskują średnio takie same wyniki. Jednak oceny w jednej z grup zostają sztucznie zaniżone, a wykładowca robi awanturę studentom z tej grupy, zarzucając im, że nie przyłożyli się do nauki. Na zakończenie kursu wykładowca przeprowadza taki sam egzamin końcowy dla obu grup i odkrywa, że wyniki studentów z grupy potraktowanej niesprawiedliwie są gorsze. Hipoteza została potwierdzona, a wyniki badania opublikowane.

f. Aby przełamać opór badanych przed podawaniem informacji o zachowaniach, które mogliby traktować jako przynoszące wstyd, osoba prowadząca badania nad zachowaniami seksualnymi zadaje pytanie w następującej formie: „Obecnie wszyscy się onanizują. Jak często ty się onanizujesz?".

g. Badania dotyczące życia w akademikach na terenie campusu uniwersyteckiego wykazały, że 60% ich mieszkańców regularnie łamie obowiązujące ograniczenia dotyczące spożywania alkoholu. Ponieważ opublikowanie tych danych spowodowałoby prawdopodobnie oburzenie wśród społeczności campusu, a nie planowano żadnych głębszych analiz dotyczących spożywania alkoholu, badacz decyduje się nie ujawniać tego wyniku.

h. Aby zbadać, w jakim stopniu ludzie są skłonni dla zachowania twarzy wyrażać swoje opinie w kwestiach, o których nie mieli żadnych informacji, badacz prosi badanych o wyrażenie opinii dotyczącej zdarzenia, które w ogóle nie miało miejsca.

i. Aby zapewnić wysoki odsetek realizacji ankiet, w pewnych badaniach kwestionariusz dostarcza się studentom wraz z dokumentami dotyczącymi ich rejestracji, w nadziei, że choć nikt im nie powiedział, że muszą go wypełnić, będą myśleli, że są do tego zobligowani.

j. Osoba prowadząca obserwację uczestniczącą w pewnym radykalnym ugrupowaniu politycznym zostaje włączona do wąskiego kręgu osób planujących szczegółowe działania grupy. Co powinna zrobić, jeśli grupa planuje:
   (1) zorganizowanie nielegalnej, ale pokojowej manifestacji;
   (2) zdetonowanie bomby w budynku publicznym, ale w czasie, gdy na pewno nie będzie tam żadnych ludzi;
   (3) przeprowadzenie zamachu na urzędnika państwowego.

2. Zapoznaj się z omówieniem eksperymentu Milgrama, dotyczącego posłuszeństwa. Jak zaplanowałbyś badania, mające takie same cele, ale którym nie można by postawić zarzutów etycznych stawianych Milgramowi? Czy twoje badania przyniosłyby równie wartościowe rezultaty? Czy miałyby te same konsekwencje?

3. Wyobraź sobie, że badacz, będący zwolennikiem małych rodzin jako rozwiązania

problemu przeludnienia, chce przeprowadzić badania sondażowe mające określić, dlaczego niektórzy ludzie chcą mieć wiele dzieci, a inni nie chcą. Jakie problemy dla tych badań mogą wyniknąć z osobistego zaangażowania badacza? Jak można ich uniknąć?

4. Jakie aspekty etyczne powinien uwzględnić badacz, planując badania sondażowe omówione w ćwiczeniu 3?
5. Znajdź w Internecie jakiś oficjalny dokument, określający procedury uzyskiwania od badanych świadomej zgody na udział w projekcie badawczym. Podaj adres internetowy tego dokumentu oraz zacytuj lub streść jedną z opisanych w nim procedur. (Podpowiedź: w którejś z wyszukiwarek użyj wziętej w cudzysłów frazy „informed consent").

Dodatkowe lektury

Hamnett Michael P., Douglas Porter J., Singh Amarjit, Krishna Kumar, *Ethics, Politics and International Social Science Research,* University of Hawaii Press, Honolulu1984. Omówienia etyki badawczej koncentrują się z reguły na interesach osób uczestniczących w badaniach; ta książka poszerza zakres dyskusji, włączając do niej prawa całych społeczeństw.

Homan Roger, *The Ethics of Social Research,* Longman, London 1991. Głęboka analiza zagadnień etycznych w badaniach społecznych, przeprowadzona przez czynnego brytyjskiego badacza społecznego.

Lee Raymond, *Doing Research on Sensitive Topics*, Sage, Newbury Park, Cal. 1993. Książka analizuje sytuacje konfliktu pomiędzy potrzebami badań naukowych a prawami osób uczestniczących w badaniach – wraz z zaleceniami co do postępowania w takich sytuacjach.

Sweet Stephen, *Using a Mock Institutional Review Board to Teach Ethics in Sociological Research,* „Teaching Sociology" 27 stycznia 1999, s. 55–59. Chociaż artykuł ten jest adresowany do nauczycieli, zawiera on wiele przykładów badań, na których można poćwiczyć swoją wrażliwość etyczną.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Rozdział
19

Zastosowania badań społecznych

Przegląd holograficzny

Nadszedł już czas, abyśmy popatrzyli wstecz na całą drogę, którą do tej pory przeszliśmy, i jeszcze raz przyjrzeli się wzajemnym relacjom między poszczególnymi komponentami badań społecznych. Nawet jeśli nie planujesz w przyszłości aktywnie zajmować się prowadzeniem działalności badawczej, to dzięki temu, czego do tej pory się nauczyłeś, możesz stać się kompetentnym odbiorcą badań, na których wyniki natykamy się na co dzień.

Wprowadzenie

Teoria – zbieranie danych – analiza

Wykorzystanie umiejętności badawczych na co dzień

Poradnik konsumenta badań społecznych
Plan badań
Pomiar
Dobór próby
Eksperyment
Badania sondażowe
Badania terenowe
Analiza istniejących danych statystycznych
Badania ewaluacyjne
Analiza danych
Raportowanie

Zakończenie

Socjologiczna strona w sieci

Infotrac College Edition

Wprowadzenie

Tak więc, jak widzisz, zbliżamy się już do końca naszej wspólnej wyprawy w dziedzinę badań społecznych. Ostatni rozdział poświęcimy na próbę powiązania rzeczy, którymi zajmowaliśmy się wcześniej, w jedną całość, łatwiejszą do przyswojenia i późniejszego wykorzystania.

Zaczniemy od przypomnienia sobie znaczenia teorii, zbierania danych oraz analizy – trzech głównych aspektów badań społecznych. Następnie przyjrzymy się, w jaki sposób można wykorzystać znajomość technik badań społecznych w codziennym życiu.

Na zakończenie tego rozdziału przedstawię ci zbiór wskazówek dla odbiorców badań społecznych. Nawet jeżeli nigdy nie zdarzy ci się prowadzić złożonego projektu badawczego, stale będziesz się stykał z badaniami prowadzonymi przez innych, a ważne jest, żebyś umiał właściwie oceniać ich wartość. Ta część rozdziału posłuży także jako podsumowanie całej książki – myślę, że sam będziesz pod wrażeniem tego, ile już umiesz.

Teoria – zbieranie danych – analiza

Badania społeczne stanowią próbę zrozumienia życia społecznego. Oczywiście nie jest to jedyny sposób: każdy człowiek codziennie stara się dostrzec jakieś prawidłowości w tym, co się wokół niego dzieje. Chcemy zrozumieć świat i chcemy być choć trochę zdolni do przewidywania przyszłych zdarzeń. W końcu chcielibyśmy mieć jakieś pojęcie o konsekwencjach naszych różnych możliwych decyzji. Chcemy robić to, co doprowadzi do rezultatów, które chcemy uzyskać.

W codziennym życiu nasze próby zrozumienia i przewidywania podejmujemy na podstawie różnych strategii. Po części po prostu wchodzimy w nawyk, robiąc to, co w przeszłości przynosiło pożądany efekt – w nadziei, że będzie tak zawsze. Społeczeństwo jako całość ustanawia w tym celu tradycje i zwyczaje. Nie musimy rozumieć, *dlaczego* przybijanie podkowy do drzwi domu przynosi szczęście, ale robimy tak, bo to działa.

Wiele tradycji dietetycznych w różnych kulturach powstało prawdopodobnie na skutek obserwacji, że ludzi umierali po spożyciu skażonej żywności. Podobnie obowiązujący hinduistów zakaz spożywania wołowiny mógł powstać jako mechanizm gwarantujący, że głodny rolnik nie zje swoich wołów – co spowodowałoby, że nie miałby potem czym zaorać pola.

Jeśli będziesz badać obyczaje obowiązujące w różnych społecznościach, często będziesz napotykał wyjaśnienia, które próbują uzasadnić poszczególne nakazy i zakazy. Wiele z nich będzie miało charakter religijny i odwoływać się będzie do sił nadprzyrodzonych, inne mogą ci się wydać głupimi zabobonami.

W przypadku wszystkich tych tradycji i ich wyjaśnień spotkasz ludzi, którzy będą gotowi opowiedzieć wiele legend potwierdzających ich zasadność. Na przykład gdybyś mieszkał na zboczu czynnego wulkanu, mógłbyś nie widzieć żadnego sensu w praktyce wrzucania do krateru butelki z ginem, ale w tradycji hawajskiej zjednywanie sobie Pani Pele jest jedynym rozsądnym działaniem, i będziesz mógł usłyszeć mnóstwo historii o nieszczęściach, których udało się w ten sposób uniknąć, i drugie tyle o nieszczęściach, które spadły na tych, co zirytowali boginię wulkanów.

Gdy tradycyjne wzorce już raz zakorzenią się w praktyce i uzyskają potwierdzenie w religii lub w innym systemie wierzeń, bardzo trudno jest nimi zachwiać. Nawet jeśli zmienią się okoliczności, które doprowadziły do powstania tradycji, wzorce mogą pozostać. Na przykład posiadanie dużej liczby dzieci w warunkach pewnego typu funkcjonującej w przeszłości gospodarki rolniczej było zrozumiałe – dzieci zapewniały siłę roboczą do prac polowych. Jednak wzorzec ten, raz przyjęty i uzasadniony, trudno jest zmienić, nawet w warunkach przeludnienia, głodu i wyczerpania zasobów.

Chociaż z reguły naukę przeciwstawia się tradycji, warto zwrócić uwagę na pewne podobieństwa. Ludzie chcą uważać swoje indywidualne i społeczne zachowania za: (1) racjonalne oraz (2) powiązane z rzeczywistością. To samo dotyczy także nauki.

Rola systemów wierzeń i uzasadniających tradycje kulturowe legendarnych przykładów

odpowiada mniej więcej roli, jaką w badaniach odgrywają teoria i analiza danych. Badacze społeczni chcą z obserwacji ułożyć sensowną całość – i to jest rola teorii. Natomiast w dziedzinie konkretnych obserwacji oczekują, że ich teorie będą znajdowały potwierdzenie w faktach. Tak więc jeśli teoria głosi, że kobiety są bardziej religijne od mężczyzn, wyniki analiz statystycznych, według których 40% mężczyzn i 70% kobiet regularnie chodzi do kościoła, będą wspierały teorię. Gdyby jednak wyniki pokazywały, że odsetek ten jest w obu grupach taki sam, poważnie podważałoby to teorię.

W naszych codziennych próbach zrozumienia otaczającej nas rzeczywistości możemy zbłądzić na trzy sposoby. Po pierwsze, możemy być tak przywiązani do naszych wyjaśnień, że będziemy ignorowali wyniki naszych obserwacji. Odnosi się to zwłaszcza do wyjaśnień o charakterze religijnym, ale podobny efekt dotyczy także głęboko zakorzenionych ideologii politycznych albo lojalności wobec rodziny. Po drugie, nasze analizy prawidłowości życia codziennego są nierzadko dosyć przypadkowe – częściej opieramy się na anegdotach niż na ścisłej analizie danych. Tak więc chętniej przypominamy sobie kilka przemawiających do wyobraźni przykładów potwierdzających nasze wyjaśnienia, niż liczymy, w jakim naprawdę procencie przypadków pojawił się określony wynik. Po trzecie, w życiu codziennym nie ma bezpośredniego powiązania pomiędzy naszymi wyjaśnieniami a obserwacjami. Często odwołujemy się do „dowodów" nie mających logicznego związku z systemem przekonań, który chcemy uzasadnić, bądź też jako jego uzasadnienie podajemy liczne, ale nie związane ze sobą doświadczenia.

Inaczej jest w badaniach społecznych, gdzie teorie poddawane stałemu testowaniu w każdej chwili mogą zostać sfalsyfikowane. Żadna teoria nie jest święta (i nie zmienia tego fakt, że niektórzy naukowcy czasem o tym zapominają, przynajmniej przejściowo). Od teorii wymagamy, by potwierdzane były nie przez anegdoty, ale poprzez jednoznaczne empiryczne dowody. Naukowcy, różniący się osobistymi doświadczeniami i wyznawanymi wierzeniami, zgodzą się ze sobą co do tego, że 40% to jest 40%, i że jest to mniej niż 70%.

Plan badawczy, procedura doboru próby, pomiar, techniki zbierania danych umożliwiają zebranie obserwacji, które nadają się do wyjaśnień teoretycznych – albo potwierdzających nasze pierwotne hipotezy, albo przeczących im. Te same metody badawcze pozwalają nam odkryć prawidłowości w danych obserwacyjnych, co czasami prowadzi do sformułowania nowej, ogólnej teorii. Techniki analityczne pozwalają na podsumowanie naszych obserwacji i wyciąganie z nich wniosków.

Zapoznanie się z metodami i ideami, którym poświęcona była ta książka, może spowodować, że twoje codzienne obserwacje i rozumowanie będą ściślejsze, że częściej będziesz zauważał sytuacje, w których ludzie myślą nielogicznie i niewłaściwie prowadzą obserwacje, a i sam także częściej będziesz się łapać na robieniu takich błędów. Mam nadzieję, że tak będzie – bo będzie to znaczyło, że stałeś się wrażliwszy na logikę badań i interpretacji – że nauczyłeś się bardziej krytycznie myśleć.

Jednocześnie w zaskakująco wielu sytuacjach stwierdzisz przydatność konkretnych umiejętności badawczych – kilka przykładów pokażę w następnym podrozdziale.

Wykorzystanie umiejętności badawczych na co dzień

W tym podrozdziale będziemy mówić o badaniach społecznych, które może ci przyjść prowadzić, nawet jeśli zasadniczo nie będziesz się tym zajmować zawodowo. Oczywiście szanse na to będą większe, jeśli ludzie w twoim otoczeniu będą cię uważali za osobę kompetentną w tej dziedzinie.

Może się na przykład zdarzyć, że będziesz kiedyś uczestniczył w zaplanowaniu i przeprowadzeniu badań sondażowych. Żyjąc w państwie demokratycznym, jesteśmy w tak znacznym stopniu przywiązani do podejmowania decyzji poprzez głosowanie, że odwołujemy się do niego automatycznie w sytuacji braku zgody. Jeśli na przykład na spotkaniu klubu działkowiczów nie ma zgody co do miejsca urządzenia dorocznego pikniku, ktoś w końcu mówi –

„głosujmy!" – i nikt nie sprzeciwi się przyjęciu takiej metody rozwiązania sporu. Bardzo często takie „głosowanie" przybiera formę sondażu wśród członków danej organizacji. Twoja umiejętność sformułowania odpowiednich pytań, skutecznego rozprowadzenia i odbioru kwestionariuszy oraz opracowywania wyników pozwoli przeprowadzić głosowanie prawidłowo.

Może być też tak, że będziesz zaangażowany w jakąś lokalną sprawę, na przykład ktoś będzie chciał usytuować składowisko odpadów toksycznych na terenach mieszkalnych. Możesz wykorzystać swoją umiejętność prowadzenia badań sondażowych do przeprowadzenia badania opinii mieszkańców w tej sprawie. Możliwe, że środki masowego przekazu nagłośnią wyniki twojego sondażu; mógłbyś wykorzystać je także jako argument w rozmowach z władzami.

Ten ścisły związek między badaniami sondażowymi a demokracją sprawia, że badania opinii są powszechnym zjawiskiem we współczesnym społeczeństwie, ale stanowią one tylko jeden ze sposobów praktycznego wykorzystania znajomości metod badań społecznych – teraz powiemy o czymś zupełnie innym.

Znajomość problemów pomiaru, zwłaszcza układania kwestionariusza, może okazać się przydatna wszędzie tam, gdzie będziesz miał do czynienia z formularzami. Wyobraźmy sobie, że zostajesz dyrektorem personalnym i chciałbyś uzyskać od kandydatów do pracy różnorodne informacje, pozwalające ocenić ich przydatność na poszczególnych stanowiskach. Z pewnością musiałeś już wielokrotnie wypełniać różne formularze, które wydawały ci się dziwne lub nawet głupie. I mogły takie rzeczywiście być – nie zawsze jest łatwo opracować formularz, który we właściwy sposób uwzględni potrzebne informacje. Wiedza, którą uzyskałeś dzięki lekturze tej książki, powinna zwiększyć twoje umiejętności w zakresie przygotowywania jasnych i użytecznych kwestionariuszy.

Znajomość zasad prowadzenia obserwacji uczestniczącej może przydać się najbardziej w codziennym życiu. Prowadzeniem obserwacji uczestniczącej zajmowałeś się od dziecka – teraz możesz to robić trochę bardziej profesjonalnie. Możesz prowadzić swoje obserwacje dokładniej, pamiętając przy tym, że patrząc i analizując świat subiektywnie, widzimy go w nieco innych niż rzeczywiste kolorach. Przyswajając sobie zasady prowadzenia badań społecznych, stałeś się zapewne bardziej wiarygodnym świadkiem. Z pewnością wzrosła też twoja wartość jako rozmówcy, bo umiesz słuchać innych i jesteś zainteresowany tym, co mają do powiedzenia.

Mam nadzieję, że twoja znajomość metod badań jakościowych pozwoli ci dostrzec wiele ukrytych aspektów życia społecznego. To, co ludzie mówią, nie zawsze jest tym samym, co mają na myśli, a to, co robią, może stać w sprzeczności z tym, w co wierzą i co sądzą. Umiejętność dostrzegania niuansów ludzkich interakcji powinna pozwolić ci rozumieć innych ludzi lepiej niż przed lekturą tej książki.

Metody analizy treści mogą okazać się przydatne w bardzo wielu okolicznościach. Jeśli będziesz uprawiać krytykę literacką, nie tylko będziesz „po prostu wiedział", co ci się podoba, ale będziesz mógł się także głębiej zastanowić, dlaczego ci się to podoba. Te same umiejętności mogą się przydać prawnikom analizującym historię regulacji prawnych dotyczących jakiegoś problemu. Specjaliści od reklamy mogą wykorzystać techniki analizy treści do ustalenia, dlaczego pewne reklamy działają lepiej od innych.

Umiejętność analizowania istniejących danych statystycznych ma bardzo wiele zastosowań. Załóżmy, że masz wybrać jedną z wielu posad albo uczelni w różnych regionach kraju. Warto w takiej sytuacji poświęcić trochę czasu na określenie, jakie zmienne, charakteryzujące region, są dla ciebie najważniejsze – klimat, koszty utrzymania, przekrój demograficzny, instytucje kulturalne itp. Gdy już wiesz, czego najbardziej chcesz, możesz poszukać danych, pozwalających odpowiednio ocenić wszystkie opcje.

Znajomość logiki eksperymentów i badań ewaluacyjnych może się przydać zarówno w życiu osobistym, jak i zawodowym. Mniej lub bardziej otwarcie, zawsze w swoich działaniach będziesz dążył do różnych celów. Może będziesz dążył do tego, by bardziej efektywnie pracować, by mieć więcej wolnego czasu, więcej zarabiać,

mniej ważyć, mniej się użerać z rodziną, a może to będą zupełnie inne rzeczy. Ludzie często próbują realizować swoje cele po omacku, ale ty po przeczytaniu tej książki możesz jasno określić: (1) czego chcesz, (2) co robisz, aby to osiągnąć i (3) czy to, co robisz, prowadzi cię do tego, czego chcesz. Nie bój się prowadzić notatek dotyczących tego, jak ci się różne rzeczy udają – choć nie musisz mówić wszystkim, że je prowadzisz.

Generalnie – jeśli pominiemy dowcipy o zwariowanych naukowcach i stukniętych profesorach, umiejętność naukowego spojrzenia na większość aspektów życia przynosi pewne rzeczywiste korzyści. W takim stopniu, w jakim polega to na trzymaniu się logiki, dokładnej obserwacji i ścisłości w interpretowaniu tego, co się zaobserwowało, na naukowym podejściu do życia możesz zyskać i ty, i twoje otoczenie.

Poradnik konsumenta badań społecznych

Być może uważasz, że pewne aspekty tej książki nie są odpowiednie dla ciebie. Poza prostymi badaniami, o których mówiliśmy przed chwilą, możesz nie przypuszczać, byś kiedykolwiek angażował się w prowadzenie badań społecznych. Możliwe – badacze społeczni stanowią niewielki odsetek społeczeństwa.

Nie jest zbyt prawdopodobne ($p<0{,}001$), że przyjdzie ci kiedykolwiek planować i zarządzać badaniem sondażowym na wielostopniowej, ogólnokrajowej reprezentatywnej próbie siły roboczej, mającym określić, czy istnieje statystyczna zależność między płcią a zarobkami w powiązanych podzbiorach pracowników etatowych. Zauważ jednak, że mniej więcej wiesz, o co chodzi, i możesz coś powiedzieć o różnych etapach tych badań: o budowie kwestionariusza, doborze próby, procedurach zbierania danych, obliczaniu średnich zarobków w różnych grupach pracowniczych (zatrudnieni na pełny etat, zatrudnieni w niepełnym czasie pracy itp.). Możliwe, że będziesz musiał zajrzeć przy tym do tej lub innej książki, żeby sobie przypomnieć jakieś szczegóły – ale fachowcy zawsze to robią.

Z grubsza szacując, wiesz teraz o badaniach społecznych więcej niż jakieś 99,98% ludności świata (jeśli zrobiło to na tobie wrażenie, namów kolegów, żeby kupili tę książkę). Należysz do grupy określanej niekiedy jako „świadomi konsumenci" – nawet jeżeli nigdy nie będziesz prowadził na własną rękę żadnych większych badań, będziesz umiał ocenić wartość badań przeprowadzonych przez innych. Jest to ważna umiejętność – bo na wyniki badań społecznych będziesz natykał się przez całe życie: będziesz czytał raporty z badań w gazetach, słuchał o nich w radiu i telewizji, a politycy i sprzedawcy będą ci mówili, co „pokazują badania". Nawet jeśli sam nie uprawiasz zboża, umiejętność oddzielania ziarna od plew zawsze może się przydać.

Poniżej znajdziesz pewne wskazówki dla odbiorców raportów z badań społecznych. Częściowo jest to streszczenie materiału, zawartego w tej książce, ponieważ jednak teraz skoncentrujemy się raczej na ocenie badań, niż na ich planowaniu i przeprowadzaniu, pominiemy tutaj te kwestie, które nie odnoszą się do umiejętności odbioru.

A więc jeśli ktoś ci mówi, „co pokazują badania", możesz zadać mu któreś z poniższych pytań. Nie wszystkie te pytania odnoszą się do wszystkich typów badań – ale zawsze znajdziesz tu coś do wyboru.

Plan badań

- Co było celem badań: eksploracja, opis, wyjaśnianie czy jakaś ich kombinacja?
- Kto przeprowadził badania? Kto je sfinansował? Dlaczego te badania zostały przeprowadzone? Jeśli tak się składa, że wyniki badań są korzystne dla zleceniodawcy lub badaczy, to powinniśmy traktować je ze szczególną ostrożnością – choć nie jest to powód, żeby je od razu odrzucać.
- Co było jednostką analizy? Czy została ona właściwie dobrana ze względu na cele badań? Czy konkluzje wyprowadzone przez badaczy odpowiadają przyjętej jednostce analizy (na przykład jeśli jednostką analizy były miasta, to czy badacze przypadkiem nie sformułowali wniosków, odnoszących się od pojedynczych ludzi)?

- Czy badania miały charakter przekrojowy, czy dynamiczny? Trzeba szczególnie uważać na twierdzenia odnoszące się do zmian w czasie, sformułowane na podstawie wyników obserwacji przekrojowych.
- Czy jeśli były to badania dynamiczne, to za każdym razem stosowano porównywalne pomiary? Jeśli były to badania sondażowe, to czy za każdym razem zadawano takie same pytania? Jeśli w badaniach porównywane są, powiedzmy, wskaźniki przestępczości czy poziom ubóstwa, to czy za każdym razem były one tak samo definiowane (na przykład definicje ubóstwa zmieniają się dość często)?
- W przypadku badań panelowych – ile osób wypadło z próby pomiędzy poszczególnymi etapami tych badań?

Pomiar

- Za pomocą jakichś pojęć określono przedmiot badań?
- Czy badacze nakreślili różne wymiary badanej zmiennej? Czy trzymali się tego rozróżnienia, prowadząc analizy i przygotowując raport?
- Jakie wskaźniki jakościowe lub ilościowe zostały wybrane jako mierniki tych wymiarów i pojęć? Czy każdy z nich rzeczywiście mierzy to, co ma mierzyć? Co jeszcze może mierzyć dany wskaźnik? Czy jest on rzetelny? Czy ta rzetelność była testowana?
- Jaki jest poziom pomiaru każdej zmiennej: nominalny, porządkowy, interwałowy (przedziałowy) czy ilorazowy? Czy jest to właściwy poziom pomiaru?
- Czy były stosowane jakieś miary złożone (indeksy, skale lub typologie)? Jeśli tak, czy były one odpowiednie do celów badań? Czy zostały poprawnie skonstruowane?

Dobór próby

- Czy ograniczenie badań do próby było uzasadnione, czy też należało przeprowadzić badania wyczerpujące? Pamiętaj, że dobranie próby losowej nie zawsze jest możliwe.
- Jeśli wymagana była próba, to czy należało ją wylosować, czy też właściwsza byłaby próba celowa, kwotowa, lub też dobierana metodą „kuli śnieżnej"? Czy zastosowano właściwy typ doboru próby?
- Na temat jakiej populacji badacz chciał wyciągać wnioski?
- Co było celem badań? Jeśli opis statystyczny, to próba powinna mieć ściśle losowy charakter.
- Jeśli zastosowano próbę losową, co było operatem losowania? Czy operat ten odpowiednio reprezentował populację, będącą przedmiotem zainteresowania badacza? Jakie elementy populacji nie wchodziły w skład operatu losowania, jakie zawierał on elementy dodatkowe?
- Jaki konkretny schemat losowania zastosowano: losowanie proste, systematyczne, grupowe? Czy przed losowaniem przeprowadzono warstwowanie? Czy dokonano go na podstawie właściwych cech, tzn. cech istotnych dla zmiennych analizowanych w badaniu?
- Jak duża była próba? Jak duży był procent jej realizacji? Czy można się spodziewać różnic pomiędzy częścią próby, która została zbadana, a tą, która nie została?
- Zakładając, że biorący udział w badaniu są reprezentatywni dla całej dobranej próby – jak duży jest błąd z próby przy próbie tej wielkości?
- W jaki sposób badacz sprawdzał reprezentatywność próby – na przykład czy porównywał rozkłady płci, wieku, narodowości, wykształcenia albo dochodu w całej populacji i w próbie?
- Generalnie – czy badane osoby (lub inne jednostki analizy) są reprezentatywne dla szerszej populacji, z której zostały wybrane? A więc czy wnioski z próby można odnieść do jakiejś znaczącej populacji albo do życia w ogóle?
- Jeśli, ze względu na charakter badań, nie miała zastosowania statystycznie reprezentatywna próba losowa, to czy badani i obserwacje zostali dobrani w sposób zapewniający szeroki ogląd analizowanego zjawiska? Czy zwrócono szczególną uwagę na przypadki nietypowe lub nie potwierdzające sformułowanych tez?

Eksperyment

Jaka jest w danym eksperymencie podstawowa zmienna zależna? Na przykład – jaki efekt eksperymentator próbował uzyskać? Jakiego bodźca (bodźców) użyto w eksperymencie?

Jakie były inne zmienne istotne dla eksperymentu? Czy były one mierzone?

Jak każda ze zmiennych była definiowana i mierzona? Jakie problemy z trafnością i rzetelnością mogły wyniknąć z tych definicji i pomiarów?

Czy posłużono się właściwą grupą kontrolną? Czy badani zostali przydzieleni do grup eksperymentalnej i kontrolnej losowo, czy przez dopasowanie? Czy procedury te zostały przeprowadzone prawidłowo? Czy eksperymentator wykazał, że grupy eksperymentalna i kontrolna były ze sobą początkowo w pełni porównywalne?

- Czy dokonywano pomiarów zmiennej zależnej przed zaaplikowaniem bodźca i po jego zaaplikowaniu, a więc czy zastosowano zarówno pretest, jak i posttest?
- Jakie jest prawdopodobieństwo wystąpienia w tym eksperymencie „efektu placebo" („efektu Hawthorne")? Czy kwestia ta była przez eksperymentatora uwzględniona (na przykład przez zastosowanie schematu podwójnie ślepego)?
- Czy są jakieś wątpliwości co do trafności wewnętrznej, związane z historią, dojrzewaniem, pomiarem, narzędziami, regresją statystyczną, obciążeniami doboru, „wymieraniem" grup, przyczynowym porządkiem czasowym, rozpowszechnianiem lub naśladowaniem zachowań, rekompensatami, rywalizacją o rekompensatę, zniechęceniem?
- Czy są jakieś problemy z trafnością zewnętrzną? W jaki sposób eksperymentator upewnił się, że wyniki uzyskane w laboratorium można odnieść także do świata realnego?

Badania sondażowe

- Sprawdź wszystkie pytania odnoszące się do doboru próby.
- Jakie pytania zadawano respondentom? Jak dokładnie były sformułowane? Uważaj na raporty, w których badacze jedynie streszczają pytania.
- Jeśli zadawano pytania zamknięte, to czy kategorie odpowiedzi były właściwie dobrane, wyczerpujące i wzajemnie rozłączne?
- Jeśli zadawano pytania otwarte, w jaki sposób kategoryzowano odpowiedzi? Czy badacz ograniczył w jakiś sposób ryzyko, że jego własne obciążenia zakradną się do procesu kodowania odpowiedzi na pytania otwarte?
- Czy wszystkie pytania są jasne i jednoznaczne? Czy nie zachodzi ryzyko, że były przez respondentów niewłaściwie rozumiane? Jeśli tak, to czy udzielone odpowiedzi mogły mieć inny sens, niż ten przyjęty przez badacza?
- Czy respondenci byli w stanie odpowiadać na stawiane im pytania? Jeśli nie, mogli mimo to udzielać odpowiedzi, aczkolwiek mogą one nie mieć żadnego sensu.
- Czy pojawiły się pytania, które dotyczą więcej niż jednej kwestii? Przyjrzyj się pytaniom, w których występują spójniki „i" albo „lub". Czy zdarzało się, że respondenci mieli się odnieść jednocześnie do dwóch poglądów w sytuacji, gdy mogliby jeden z nich podzielać, a drugiego nie?
- Czy pojawiły się pytania zawierające przeczenia? Pytania takie mogły zostać źle zrozumiane, co prowadzi od udzielania niewłaściwych odpowiedzi.
- Czy na odpowiedzi na niektóre pytania mogła wpłynąć chęć dostosowania się do norm społecznych? Jeśli na jakieś pytania są „dobre" lub „złe" odpowiedzi, respondenci, odpowiadając na nie, mogli kierować się tym, „co ludzie o nich pomyślą".
- Sprawdź wszystkie pytania kwestionariusza, zastanawiając się, jak byś na nie odpowiadał. Jeśli miałbyś z którymś z nich trudności, mogliby je mieć również inni. Następnie zastanów się, jak dane pytanie mogłoby być odbierane przez osoby o różnych poglądach (na przykład liberała i konserwatystę, osobę religijną i niereligijną).

- Czy przeprowadzano wtórną analizę danych zebranych dawniej? Jeśli tak, oceń wartość tamtych oryginalnych badań. Poza tym sprawdź, czy analizowane dane odpowiadały aktualnym potrzebom? Czy oryginalnie zadawane pytania wystarczająco dobrze oddają aktualnie analizowane zmienne?
- Krajowa Rada do Spraw Badania Opinii Publicznej (*National Council on Public Polls*) opracowała listę 20 pytań dotyczących sondaży – znajdziesz je w załączniku J.

Badania terenowe

- Według jakiego teoretycznego paradygmatu badacz prowadzi te badania?
- Czy badania podjęto w celu testowania hipotez, czy zbudowania teorii na bazie obserwacji? A może też badacz w ogóle nie odnosi się do kwestii teoretycznych?
- Jakie są najważniejsze zmienne w tych badaniach? Jak są definiowane i mierzone? Czy dostrzegasz jakieś problemy z trafnością?
- Co z rzetelnością badań? Czy inny badacz, obserwując te same wydarzenia, sklasyfikowałby je tak samo?
- Czy można być pewnym, że na dokonaną klasyfikację obserwacji nie miały wpływu konsekwencje, jakie klasyfikacja mogłaby mieć dla hipotez i/lub wniosków formułowanych przez badacza?
- Jeśli wyciągane były wnioski o charakterze opisowym, na przykład „grupa zasadniczo przyjmowała dość konserwatywne standardy", to co dokładnie oznaczają użyte w opisie terminy?
- W jakim stopniu uzyskane wyniki można uogólniać na jakąś szerszą część społeczeństwa? Co w tej kwestii deklaruje autor badań? Na czym opiera swoje deklaracje?
- Jeśli przeprowadzano wywiady, to jak dobierano respondentów? Czy reprezentowali oni wszystkie istotne typy badanych?
- W jakim stopniu badacz uczestniczył w wydarzeniach, które analizował? Na ile fakt jego uczestnictwa mógł wpłynąć na ich przebieg?
- Czy badacz ujawnił badanym fakt, że prowadzi badania? Jeśli tak, to jaki mogło to mieć wpływ na ich zachowania?
- Czy można dostrzec u badacza jakieś osobiste uczucia – pozytywne lub negatywne – wobec osób przez niego obserwowanych? Jeśli tak, to jaki wpływ mogły mieć te odczucia na poczynione obserwacje i wnioski z nich wyciągnięte?
- W jaki sposób tożsamość kulturowa bądź osobiste doświadczenia badacza mogły wpłynąć na interpretację jego obserwacji?

Analiza istniejących danych statystycznych

- Kto pierwotnie zebrał analizowane teraz powtórnie dane? Czy proces zbierania danych przebiegał całkowicie prawidłowo?
- W jakim pierwotnie celu były zbierane te dane? Jaki mogło to mieć na nie wpływ?
- Co było jednostką analizy dla tych danych? Czy pasuje ona do aktualnych pytań badawczych i formułowanych wniosków? Czy można wykluczyć ryzyko błędu ekologicznego?
- Kiedy dane były zebrane? Czy nadal są one odpowiednie dla analizowanych obecnie problemów?
- Jakie zmienne były analizowane w aktualnych badaniach? Czy ich definicje, przyjęte przez autorów oryginalnych badań, odpowiadają współczesnym potrzebom?

Badania ewaluacyjne

- Jaka interwencja społeczna była przedmiotem analizy? Jakie zastosowano metody pomiaru? Czy pojawiły się jakieś problemy z trafnością i rzetelnością?
- Czy do obserwacji wybrano właściwe osoby (jednostki analizy)?
- Jak zdefiniowano „sukces"? Kogo ów sukces dotyczył – jednostek, organizacji, poziomu przestępczości? Czy był on właściwie mierzony?
- Czy badacz ocenił ewaluowaną interwencję jako sukces lub porażkę? Czy ta ocena jest dobrze uzasadniona?
- Kto sfinansował badania, a kto je faktycznie przeprowadził? Czy można mieć zaufanie do obiektywizmu badacza? Czy zleceniodawca wpływał jakoś na przebieg badań?

Analiza danych

- Czy cel i plan badań wymagały analizy jakościowej, czy ilościowej?
- Jak zostały zakodowane dane niestandaryzowane? Pytanie to dotyczy zarówno analizy jakościowej, jak i ilościowej. W jakim zakresie kategorie kodowania: (1) opierały się na wcześniejszej teorii, a w jakim (2) zostały stworzone na podstawie danych?
- Czy badacz przeprowadził wszystkie istotne dla badań analizy? Czy wszystkie istotne zmienne zostały określone i zbadane? Czy korelacja, zaobserwowana pomiędzy dwiema zmiennymi, mogła być spowodowana przez trzecią, wcześniejszą zmienną, a zaobserwowana zależność była jedynie pozorna?
- Czy konkretne wnioski z analizy są rzeczywiście istotne? Czy na przykład zaobserwowana pomiędzy dwiema podgrupami różnica jest rzeczywiście duża lub znacząca? Czy wymaga podjęcia jakichś działań?
- Czy badacz, formułując wnioski, nie wykroczył poza rzeczywiste wyniki?
- Czy analiza i interpretacja danych prowadzona była poprawnie od strony logicznej?
- Czy w obserwacjach empirycznych można dostrzec nowe wzorce zależności, pozwalające sformułować nowe teorie ugruntowane? Czy badacz analizował te przypadki, które nowych teorii nie potwierdzają?
- Czy użyte techniki analizy statystycznej były odpowiednie dla poziomu pomiaru analizowanych zmiennych?
- Jeśli stosowano testy istotności statystycznej, to czy ich wyniki były prawidłowo interpretowane? Czy istotności statystycznej nie mylono ze znaczeniem merytorycznym?

Raportowanie

- Czy badacz przedstawił swoje badania w kontekście poprzednich badań dotyczących tego samego tematu? Czy te badania uzupełniają, zmieniają, powtarzają, czy też kwestionują wyniki wcześniejszych badań?
- Generalnie – czy szczegóły planu badań oraz ich przebiegu zostały opisane wyczerpująco? Czy jakieś części raportu z badań wydają ci się mętne i niedokładne w opisie szczegółów?
- Czy badacz pisał o jakichś usterkach lub ograniczeniach planu lub realizacji badań? Czy proponował, w jaki sposób można by ulepszyć badania tego typu w przyszłości?

Zakończenie

Niezależnie od tego, czy związesz swoją karierę z prowadzeniem badań społecznych, czy po prostu wykorzystasz nabyte umiejętności jako kompetentny konsument, mam nadzieję, że ta wyprawa w dziedzinę systematycznych badań nad problemami społecznymi rozbudziła twoje zainteresowanie tą dziedziną. Emocje, których dostarcza uprawianie nauk społecznych, trudno jest porównać z czymkolwiek innym w nauce. Po długiej nocy spędzonej na analizie danych, znajdowaniu prawidłowości i w końcu na dostrzeganiu logiki relacji międzyludzkich – poranny spacer po mieście, budzącym się właśnie do kolejnego pełnego społecznych działań i relacji dnia, daje szczególne odczucie poznania czegoś, co właśnie udało ci się trochę lepiej zrozumieć.

Podobną satysfakcję odczuwają także konsumenci badań społecznych. Jak mówiłem, wszyscy chcemy zrozumieć otaczający nas świat, a twoje nowo nabyte umiejętności w dziedzinie badań społecznych mogą sprawić, że staniesz się aktywnym i krytycznym uczestnikiem czegoś, co czasem jest nazywane „społeczeństwem informacyjnym". Masz szansę uwolnić się od jakiejś części frustracji, wywoływanej przez niekończące się różnice zdań „ekspertów". Będziesz w stanie choć trochę zrozumieć, o co chodzi prezentującym odmienne opinie specjalistom, a nawet ocenić, które wyniki są najbardziej odpowiednie w jakich sytuacjach i dla jakich twoich celów.

I jeszcze coś. Ufam, że to, czego się tu nauczyłeś, doda ci siły do działania na rzecz poprawy twojej społeczności i świata. W miarę, jak życie społeczne staje się coraz bardziej złożone, ludziom coraz częściej wydaje się, że nie mają wpływu na kształtowanie struktur relacji międzyludzkich. Co za tym idzie – godzimy się na niedoskonałość, stoimy z boku i pozwalamy so-

bie na marnowanie szans. Namawiam cię, byś odłożył na bok swój opór przed zaangażowaniem się w sprawy, które cię dotyczą. Weź je w swoje ręce – teraz, kiedy wyostrzyłeś umiejętność krytycznego myślenia i wnikliwej obserwacji, uczyń z nich wkład na rzecz świata, w którym żyjesz i tych, którzy żyją w nim wraz z tobą.

Socjologiczna strona w sieci

Odwiedź stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz dodatkowe linki, internetowe ćwiczenia i quizy do rozdziału oraz inne materiały:

http://www.sociology.wadsworth.com

Infotrac College Edition

Podsumowanie – wyszukiwanie słów

Wejdź na stronę Wadsworth Sociology Resource Center, Virtual Society, gdzie znajdziesz listę szukanych słów do każdego rozdziału. Używając wyszukiwarki idź do Infotrac College Edition – biblioteki internetowej z zasobami ponad 900 czasopism. Możesz tam przeprowadzić badanie online i znaleźć lektury związane z twoimi badaniami. Pomocą w poszukiwaniach i przydatnymi wskazówkami służy The Student Guide to Infotrac College Edition na stronie The Virtual Society:

http://www.sociology.wadsworth.com

Załączniki

A. Jak korzystać z biblioteki

B. Badania społeczne w cyberprzestrzeni

C. Raport z badań

D. Kwestionariusz struktury gospodarstwa domowego z badania GSS

E. Liczby losowe

F. Rozkład chi-kwadrat

G. Powierzchnie pod krzywą normalną

H. Szacunkowy błąd z próby

I. Podręcznik SPSS

J. Dwadzieścia pytań, które dziennikarz powinien zadać, zapoznając się z wynikami sondażu opinii

A

Jak korzystać z biblioteki

Wprowadzenie

Żyjemy w świecie pełnym raportów z badań. Gazety, czasopisma, pisma naukowe, informatory dla maturzystów, biuletyny różnych organizacji społecznych – właściwie wszystko, co zdarza nam się czytać, zawiera raporty na różne tematy. Formalne rozpracowywanie problemu badawczego wciąż jednak najlepiej zacząć od dobrej biblioteki uniwersyteckiej. Obecnie istnieją dwie metody poszukiwania materiałów bibliograficznych: tradycyjna i elektroniczna. Zacznijmy od systemu tradycyjnego, następnie zaś przyjrzyjmy się wariantowi elektronicznemu.

Jak uzyskać informacje

Jeśli chcesz znaleźć coś w bibliotece, twoim najlepszym przyjacielem jest dyżurny bibliotekarz, który jest specjalnie przygotowany w zakresie prowadzenia poszukiwań w księgozbiorze. W niektórych bibliotekach jest wielu dyżurnych bibliotekarzy, specjalizujących się w naukach społecznych, humanistyce, dokumentach rządowych itd. Odszukaj specjalistę od twojej dziedziny i umów się na spotkanie, po czym wyjaśnij, co cię szczególnie interesuje. Bibliotekarz zaproponuje ci prawdopodobnie kilka spośród wielu dostępnych źródeł bibliograficznych.

Źródła bibliograficzne

Zapewne znane ci jest określenie „eksplozja informacyjna". Twoja biblioteka to jedno z głównych pól walki. Na szczęście wiele wydawnictw oferuje przewodniki po dostępnej informacji.

Książki w druku

To wydawnictwo zawiera listę wszystkich książek, które są aktualnie drukowane w Stanach Zjednoczonych – uporządkowane według tytułu i autora. Książki już wydrukowane znajdziesz w starszych wydaniach.

Przewodnik po czasopismach dla czytelników

Ten aktualizowany raz na miesiąc rocznik zawiera listę artykułów publikowanych w większości gazet i czasopism. Przewodnik jest doskonałym źródłem informacji o lekturach dotyczących twoich zainteresowań, ponieważ hasła są uporządkowane tematycznie. Rysunek A.1 zawiera próbkę tekstu z *Przewodnika*.

Oprócz tomów ogólnych znajdziesz też wiele bibliografii specjalistycznych, jak na przykład:

- Abstrakty socjologiczne
- Abstrakty psychologiczne
- Indeks nauk społecznych
- Indeks cytowań w naukach społecznych

Rysunek A.1

Strona z *Przewodnika czytelnika po czasopismach*

Przewodnik czytelnika po czasopismach

MUZYKA – cd.

Nauka i nauczanie

Patrz też
Gitara – nauka i nauczanie

Tematy, motywy etc.

Patrz też
Samochody w muzyce

Teoria

Patrz też
Atonalność

Japonia

Muzyka japońska i zachodnia. L. Futoransky *The Courier (Unesco)* **40:38 + D '87**

MUZYKA AMERYKAŃSKA

Patrz też
Muzyka jazzowa

MUZYKA ELEKTRONICZNA

Patrz też
Komputery – zastosowanie w muzyce

MUZYKA I PAŃSTWO

Punkt widzenia (rządowe dotacje dla oper) J. L. Poole, „Opera News", 52:4 F 13 '88

Związek Radziecki

Gorbatchov sets the beat for rock, „U.S. News & World Report" 104: 8–9 F 8 '88

MUZYKA I NIEWIDOMI

Call him Doc (D. Watson) F. L. Schultz, „Country Journal" 15:44 53 F '88

MUZYKA I NIEPEŁNOSPRAWNI

Patrz też
Gitarzyści, Niepełnosprawni

MUZYCZNE STOWARZYSZENIE AMERYKAŃSKIE *Patrz* MCA INC.

MUZYCZNA KRYTYKA I KRYTYCY

Patrz też
Recenzje operowe

MUZYCZNE FESTIWALE

Austria

Bregenz. H. Koegler, „Opera News" 52:38 F 13 ' 88

Niemcy (Zachodnie)

Bayreuth. J. F. Sutcliff „Opera News" 52:36 Ja 30 ' 88

Wielka Brytania

Buxton. E. Forbes, „Opera News" 52:40–1 F 13 ' 88

Włochy

Torre del Lago (Festiwal Pucciniego) M. Hamlet-Mets, „Opera News" 52:38–40 F 13 '88

Pensylwania

Filadelfia (Amerykański Festiwal Teatrów Muzycznych) R. Baxter, „Opera News" 52:34 Ja 30 ' 88

MUZYCZNE KOMEDIE, REWIE ETC. *Patrz* musicale, rewie etc.

MUZYCZNE INSTRUMENTY, ELEKTRONIKA

It's alive with the sound of – well, just about everything **(Synclavier Syntetyzer)** L. Helm, „Business Week" p75 F 8 ' 88

MUZYCZNE INSTRUMENTY – PRZEMYSŁ

Patrz też
Stowarzyszenie Przemysłu Elektronicznego Nowej Anglii

MUSICALE, REWIE ETC.

Choreografia

Patrz Choreografia

Recenzje

prace pojedyncze

Anything goes
„Dance Magazine" 62:52–7 Ja '88 J. Gruen

Do lasu
„Dance Magazine" 62:64 Ja '88 K. Grubb

Kabaret
„Dance Magazine" 62:73–4 Ja '88 H. M. Simpson

Oil City Symphony
„The Nation" 246:175–6 F 6 '88 T. M. Ditch

Upiór w operze
„Life" 11:88–92 F '88 M. Stasio
„McLean's" 101:51 F 8 '88 J. Simon
„New York" 21:89–90 F 8 '88 M. Kramer
„The New Yorker" 63:97–8 F 8 '88 J. Kroll
„Newsweek" 111:68–70 + F 8 '88 D. Handelman
„Time" 131:83–4 F 8 '88 WA. Henry

Wybrany
„The Nation" 246:176 F 6 '88 T. M. Ditch

Dekoracje i sceneria

High-tech magic: follow that gondola [*Upiór w operze*] J. Kroll „Newsweek" 111:70 F 8 '88

Scenariusze

Changing the face of Broadway [A. Lloyd Webber] M. Stasio „Life" 11:88–92 F '88

MUZYCY

Patrz też
Narkotyki i muzycy
Muzycy rockowi

MUSKE, KAROL 1945–

Skid [wiersz] „The New Yorker" 63:38 F 8 '88

MUZUŁMANIE

Patrz też
Islam

Afganistan

Beyond Afghan Stalemate, L. Komisar, „The New Leader" 71: 5–6 Ja 11–25 '88

Bliski Wschód

The Islamic resurgence: a new phase? R. Wright, bibl f „Current History" 87:53–6 +F '88

MUTACJA

Patrz też
Transpozycje

MUTUALIZM (BIOLOGIA) *Patrz* Symbioza

MUZIEKTHEATER (AMSTERDAM, HOLANDIA)

Patrz Opery

MYASTHENIA GRAVIS

Suzanne Rogers: *I looked at my face and* **thought: „Who would hire a freak?"** A. W. Petrucelli, „Redbook 170:104 +F '88

MYKOBAKTERYJNE CHOROBY

Patrz też
Gruźlica

MYKSOTYNY *Patrz* Toksyny i antytoksyny

N

N.W. AYER & SON, INC.

Ayer to the throne (Burger King ad campaign) B. Kanner, **„New York" 21:24** + F 29 '88

NADIS, STEPHEN J.

Robot observatories „Omni" („New York") 10:24 + Ja '88

NSPE *Patrz* Narodowe Stowarzyszenie Postępu Edukacji

NAKAGAMI, KENJI, 1946–

o

Two contemporary writers D. Palmé „The Courier" (Unesco) 40:44 **D '87**

NANDINA

Nandina does the unexpected „Southern Living" 23:50 Ja '88

NEAPOL (WŁOCHY)

Muzyka

Patrz też
Opera – Włochy

NAROKTYKI – UZALEŻNIENIE Patrz Nadużycie narkotyków

NARKOTYKI – REGULACJA PRAWNA

Patrz też
Statki w przepisach antynarkotykowych
Roboty w przepisach antynarkotykowych

Austria

A five year penalty call [Chech hockey legend J. Bubla serving prison sentence for smuggling heroin] J. Holland „Maclean's" 101:6 F 8 '88

Kolumbia

Battling the drugs lords [Attorney general C. Hoyod murdered] E. Tolnie „Maclean's" 101:26 F 8 '88

Day of the assasins [Attorney general C. Hoyod murdered] M. S. Serrill, „Time" 131:42 F 8'88

How cocaine rules the law in Columbia [Attorney general C. Hoyod murdered] C. A. Robins „U.S. News & World Report" 104:28–9 F 8 '88

Murderers of Medellin [Assasination of Columbia's Attorney general C. Hoyod] F. Willey „Newsweek" 111:33 F 8 '88

NARKOTYKI – HANDEL

Patrz też
Narkotyki – regulacja prawna
Statki w przepisach antynarkotykowych
Roboty w przepisach antynarkotykowych
Teen drug dealers: uncovering the real story. W. White, K. Dickerson, „Teen" 32:36–9 + F '88

Panama

The dictator in the dock [M.A.Noriega] N. Cooper „Newsweek" 111:33 F 22'88

Drugs, money and death [cover story, special section] „Newsweek" 111:37 F 8 '88

Noriega's money-maschine [aides testify before Senate subcommitee] M. S. Serrill, „Time" 131:39–40 F 22 '88

- Popularny przewodnik po publikacjach rządowych
- Indeks „New York Timesa"
- Fakty
- Badanie raportów wydawniczych
- Indeks czasopism ekonomicznych
- Miesięcznik wydawnictw oficjalnych
- Biuletyn informacyjny służby publicznej
- Indeks edukacyjny
- Indeks nauk stosowanych i technologii
- Przewodnik po czasopismach geograficznych
- Ogólny indeks nauk
- Indeks biologiczny i rolniczy
- Indeks opieki i ochrony zdrowia
- Indeks studiów nad ochroną zdrowia
- Indeks wydawnictw niskonakładowych
- Popularny indeks czasopism
- Indeks biograficzny
- Kwartalnik tygodniowych raportów Kongresu
- Literatura biblioteczna
- Indeks bibliograficzny

Jak znaleźć książkę

Poważne badania wymagają na ogół przeszukiwania regałów, na których zwykle przechowywane są książki. Poniższy fragment zawiera informacje na temat poszukiwania książek na półkach.

Katalog

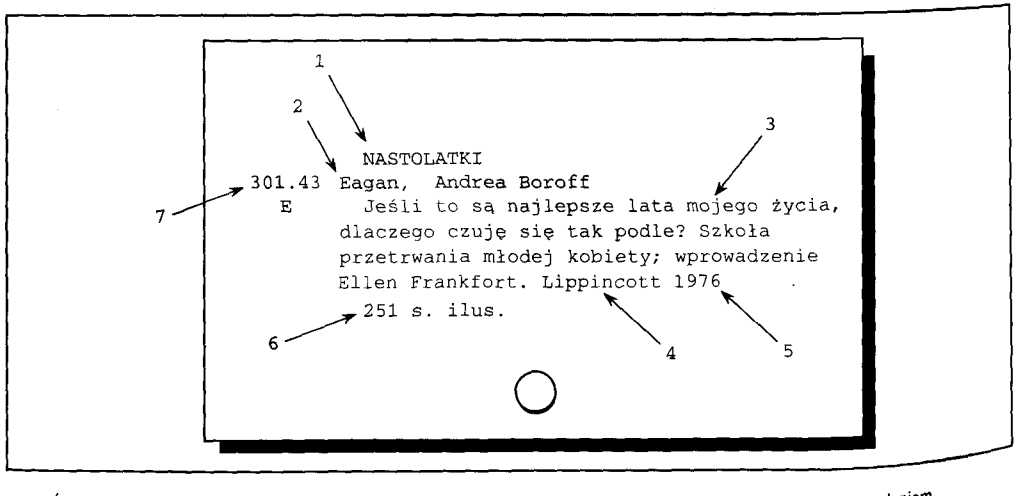
W systemie tradycyjnym katalog, zwierający karty katalogowe, jest na ogół głównym źródłem informacji o położeniu książki. Każda książka jest opisana na oddzielnej karcie o wymiarach 3 na 5 cali. Karty są następnie porządkowane alfabetycznie w trzech zestawach – według nazwiska autora, tytułu i tematu.

Jeśli szukasz konkretnej książki, możesz sprawdzić jej położenie pod nazwiskiem autora albo tytułem. Jeśli natomiast interesuje cię pewne zagadnienie ogólne, powinieneś przejrzeć katalog przedmiotowy. Rysunek A.2 przedstawia przykładową kartę katalogową. Zawiera ona następujące elementy:

1. Hasło przedmiotowe (zawsze wielkimi literami)
2. Nazwisko autora (nazwisko, imię)
3. Tytuł książki
4. Wydawca
5. Data wydania
6. Liczba stron i inne informacje dodatkowe (na przykład czy książka jest ilustrowana)

Rysunek A.2

Przykład karty z katalogu rzeczowego



Źródło: Lilian L. Shapiro, *Teaching Yourself in Libraries*, H. W. Wilson, New York 1978, s. 3–4. Wykorzystano za pozwoleniem.

7. Kod służący do odszukania książki naukowej na półce bibliotecznej; beletrystyka jest zazwyczaj ułożona w porządku alfabetycznym według nazwisk autorów

Klasyfikacja Biblioteki Kongresu

Oto praktyczna strategia przy badaniu określonego zagadnienia: skoro znasz już kod interesującej cię książki, podejdź do półki, znajdź tę książkę i obejrzyj przy okazji inne książki na sąsiednich półkach. Znajdziesz w ten sposób inne prace o podobnej tematyce, ponieważ książki ułożone są według haseł przedmiotowych.

Możesz też od razu udać się do magazynu i przejrzeć książki na interesujący cię temat. W większości bibliotek książki są ułożone i ponumerowane według klasyfikacji przedmiotowej, wypracowanej przez Bibliotekę Kongresu (acz niektóre stosują także system dziesiętny Deweya). Oto lista wybranych kategorii Biblioteki Kongresu.

Klasyfikacja Biblioteki Kongresu (fragment)

A DZIEŁA OGÓLNE
B FILOZOFIA, PSYCHOLOGIA, RELIGIA
- B-BD Filozofia
- BF Psychologia
- BL-BX Religia

C NAUKI POMOCNICZE HISTORII
D HISTORIA (poza Ameryką)
- DA-DR Europa
- DS Azja
- DT Afryka

E-F HISTORIA (Ameryka)
- E Stany Zjednoczone
- E51-E99 Indianie północnoamerykańscy
- E185 Murzyni w Stanach Zjednoczonych
- F101-1140 Kanada
- F1201-3799 Ameryka Łacińska

G GEOGRAFIA-ANTROPOLOGIA
- G-GF Geografia
- GC Oceanologia i oceanografia
- GN Antropologia
- GV Sporty, rozrywka, gry

H NAUKI SPOŁECZNE
- H62.B2 *Badania społeczne w praktyce*
- HB-HJ Ekonomia i biznes
- HM-HX Socjologia

J NAUKI POLITYCZNE
- JK Stany Zjednoczone
- JN Europa
- JQ Azja, Afryka
- JX Stosunki międzynarodowe

K PRAWO
L EDUKACJA
M MUZYKA
N SZTUKI PIĘKNE
- NA Architektura
- NB Rzeźba
- NC Sztuki graficzne
- ND Malarstwo
- NE Rytownictwo
- NK Ceramika, tekstylia

P JĘZYK I LITERATURA
- RE Literatura angielska
- PG Języki słowiańskie
- PJ-PM Języki wschodnie
- PN Dramat, oratorium, dziennikarstwo
- PQ Literatura romańska
- PR Literatura amerykańska
- PT Literatura germańska

Q NAUKA
- QA Matematyka
- QB Astronomia
- QC Fizyka
- QD Chemia
- QE Geologia
- QH-QR Biologia

R MEDYCYNA
- RK Stomatologia
- RT Położnictwo

S ROLNICTWO – PRODUKCJA ROŚLINNA I ZWIERZĘCA
T TECHNOLOGIA
- TA-TL Inżynieria
- TR Fotografia

U NAUKI WOJSKOWE
V NAUKA O MARYNARCE WOJENNEJ
Z BIBLIOGRAFIA I BIBLIOTEKOZNAWSTWO

Katalogi skomputeryzowane

Coraz częściej zasoby bibliotek są katalogowane elektronicznie. Systemy komputerowe używane przez biblioteki różnią się oczywiście między sobą, poniżej przedstawiam jednak typowy przykład ich funkcjonowania.

Siedząc przed monitorem – w bibliotece, w pracowni komputerowej lub w domu – możesz wpisać tytuł książki i chwilę później obejrzeć jej kartę katalogową. Jeśli chcesz się bliżej zapoznać z daną książką, możesz wpisać odpowiednią komendę i zobaczyć streszczenie.

Możesz też wpisać hasło tematyczne i otrzymać listę książek i artykułów na ten temat. Listę tę możesz przejrzeć, zaznaczając pozycje, które chciałbyś zobaczyć.

Wiele bibliotek zapewnia obecnie dostęp do czasopism i książek poprzez Internet. Skomputeryzowany system biblioteczny powinien zawierać informacje o materiałach dostępnych w sieci. Niekiedy można pobrać przez Internet nawet całe artykuły czy książki. Największa biblioteka w twojej okolicy zapewnia z pewnością dostawy materiałów dla swoich czytelników. Załącznik B zajmuje się także innymi aspektami badań społecznych w cyberprzestrzeni.

Wiele bibliotek uniwersyteckich ma obecnie dostęp do Centrum Informacyjnego Zasobów Edukacyjnych (*Educational Resources Information Center* – ERIC). Ten system komputerowy pozwala na przeszukiwanie setek wiodących czasopism edukacyjnych i znalezienie interesujących cię publikacji w dziedzinie edukacji. W zasadzie każda biblioteczna strona www powinna zawierać listę baz danych, które możesz przejrzeć, co ułatwi ci ograniczenie liczby tytułów związanych z danym hasłem. Upewnij się, że zawęziłeś swoje pole poszukiwań przez określenie, powiedzmy, daty publikacji i języka. Kiedy wybierzesz już artykuły, którymi jesteś zainteresowany, komputer wydrukuje ich streszczenia.

Szczególnie wartościowe dla socjologów wydawnictwa „Abstrakty Socjologiczne" (*Sociological Abstracts*) i „Abstrakty Psychologiczne" (*Psychological Abstracts*) zawierają – przygotowane często przez samych autorów – opisy treści książek i artykułów, możesz więc szybko i bez trudu zlokalizować wiele pozycji bibliograficznych. Skoro znajdziesz adekwatne pozycje, możesz przejść do dzieła oryginalnego i przejrzeć je w całości. Streszczenia dostępne są zarówno na papierze, jak i w formie elektronicznej.

Rysunek A.3 zawiera streszczenie artykułu, otrzymane w czasie komputerowego przeszukiwania „Abstraktów Socjologicznych". Zacząłem od pytania o listę artykułów dotyczących podręczników do socjologii. Po przejrzeniu listy postanowiłem obejrzeć streszczenia każdej z wybranych pozycji. Oto, co otrzymałem chwilę później: artykuł socjologa Grahama C. Kinlocha, opublikowany w „International Review of Modern Sociology".

Rysunek A.3

Streszczenie raportu z badań w *Abstraktach Socjologicznych*

AU – Kinloch-Graham-C.
T – *The Changing Definition and Content of Sociology in Introductory Textbooks*, 1894–1981
P – *International Review of Modern Sociology*, 1984, 14, 1, wiosna, s. 89–103
HP – Socjologia–edukacja (D810300). Podręcznik (D863400)
S – Analiza 105 podręczników ze wstępu do socjologii, opublikowanych w latach 1894–1981. Wskazuje na zmiany historyczne w definicji dyscypliny i jej głównych zagadnień w zależności od wpływu czynników zawodowych i zmian społecznych. Omawiane są dominujące w każdej dekadzie koncepcje socjologii, przy czym przeważająca opinia głosi, że socjologia to „naukowe badanie struktury społecznej w celu zmniejszenia liczby konfliktów i dewiacji oraz wzmocnienia kontroli społecznej". Trwałość tego poglądu, wraz ze stosunkowo niewielką wrażliwością podręczników na problemy społeczne, analizowana jest w kontekście względnie homogenicznej struktury wieku i wykształcenia ich autorów. 1 tabela, 23 przypisy.
Zmienione HA.

W razie gdyby znaczenie skrótów na rysunku A.3 nie było dla ciebie oczywiste, wyjaśniam: AU to autor, T – tytuł, P – źródło lub położenie oryginalnej publikacji, HP – hasło przedmiotowe, pod którym dostępne jest streszczenie, S – streszczenie. Komputerowa dostępność zasobów, takich jak „Abstrakty Socjologiczne", stanowi potężne narzędzie badawcze dla współczesnych socjologów. Będziesz mógł ściągnąć lub wydrukować, ze streszczeniem lub bez niego, każdy tytuł, który znajdziesz dzięki bibliotecznym przeglądarkom.

Jeśli dany dokument nie jest dostępny w samej bibliotece lub przez Internet, zawsze możesz skorzystać z tzw. rewersu międzybibliotecznego, którego użycie jest na ogół darmowe. Biblioteki nie mają wszystkich książek czy materiałów multimedialnych (CD-ROM-y, kasety wideo, płyty kompaktowe, filmy), mogą jednak zawierać między sobą porozumienia o wzajemnych wypożyczeniach, z których możesz skorzystać. Musisz jednak być świadom czasu, który musi upłynąć między złożeniem wniosku a otrzymaniem książki czy artykułu. Jeśli chodzi o książkę, która jest w innej bibliotece, być może rozsądniej będzie udać się tam osobiście. Kluczem do dobrego poszukiwania bibliografii jest informacja. Współpracuj więc z bibliotekarzami, profesorami i kolegami!

Dodatkowe lektury

A *Guide to Writing Sociology Papers*, Judith Richlin-Klonsky, Ellen Strenski (red.), St. Martin's Press, New York 1998. Świetna, nieduża książka pełna dobrych rad dla badaczy. Szczególnie polecana młodym socjologom, którzy muszą się nauczyć najbardziej podstawowych aspektów badań.

Bart Pauline, Frankel Linda, *The Student Sociologist's Handbook*, Random House, New York 1986. Kamizelka ratunkowa dla badacza społecznego. Wprowadza krok po kroku w tajniki tworzenia dokumentacji badawczych, zawiera rozdziały na temat czasopism, indeksów i abstraktów, pomoc bibliograficzną i inne wtórne źródła, poza tym zaś kompletny przewodnik po rządowych i pozarządowych źródłach danych. Oddzielny rozdział na temat ról płciowych i studiów kobiecych.

Li Tze-chung, *Social Science Reference Sources: A Practical Guide*, Greenwood Press, Westport, Con. 1990. Wymienia i opisuje wszystkie rodzaje źródeł bibliograficznych, w tym bazy danych, archiwa i źródła drukowane. Książka podzielona jest na dwie części: nauki społeczne w ogóle i wedle dyscyplin.

Załącznik
B

Badania społeczne w cyberprzestrzeni

Nowe źródło danych staje się dla socjologów niezbędnym narzędziem: Internet. Jako że Internet, www i inne elementy tej „informatycznej autostrady" zmieniają się z miesiąca na miesiąc, poniższy aneks obejmuje tylko podstawy, które mogą się przydać czytelnikom nieobeznanym z cyberprzestrzenią. Omówione zostaną cztery tematy: e-mail, listy mailingowe, usługi *gopher* i strony www. Jeśli właśnie zaczynasz odkrywać to niezwykle użyteczne narzędzie, przedstawione tu informacje mogą pozwolić ci się zorientować, gdzie skutecznie poszukiwać pomocy.

E-mail

Internet jest obecnie najczęściej używany jako substytut usług pocztowych i telefonicznych. Podobnie jak w przypadku listu, piszesz wiadomość, w tym przypadku na ekranie komputera. Następnie, zamiast ją drukować jako list, wysyłasz ją przez sieć komputerową. Adresat odczytuje twoją wiadomość na własnym komputerze.

Aby używać poczty elektronicznej lub któregokolwiek innego systemu, opisanego w tym załączniku, musisz mieć konto komputerowe na uczelni albo u jakiegoś operatora. Jeśli twoja uczelnia nie oferuje kont pocztowych, możesz zastanowić się nad skorzystaniem z usług komercyjnych operatorów internetowych. Zapewne słyszałeś o największym z nich – America Online – są jednak i inne możliwości. Lista ze stycznia 2000 roku informuje, że istniało wówczas już 8300 komercyjnych operatorów poczty. Meta-lista Operatorów Dostępu (http:/www.herbison.com/herbison/iap_meta_list.html) oferuje podobne zestawienie.

Użytkownicy sieci lokalnych mogą się z łatwością komunikować między sobą, podobnie jak użytkownicy tej samej usługi komercyjnej. Postaram się pokazać, jak można komunikować się z innymi częściami Internetu.

Skoro już podłączyłeś się do sieci, wysyłasz wiadomości zawierające następujące elementy: wiadomość (jak nota, przypomnienie czy list), tytuł (krótki nagłówek określający wiadomość, na przykład „Plany urlopowe") oraz adres, pod który chcesz wysłać wiadomość. Twój adres będzie dołączany do każdej wiadomości automatycznie.

Adres skrzynki pocztowej zawiera zazwyczaj imię, @ (*at*) i lokalizację. Każdy adres ma trzy elementy:

<nazwa> @ <serwer>.<typ>

Nazwa jest zazwyczaj stosunkowo prosta. Nadaje się ją przy zakładaniu konta – moja na przykład brzmi „babbie" i nic w tym dziwnego.

Typ konta to także stosunkowo prosta sprawa. Seria skrótów informuje o typie instalacji używanej dla uzyskania dostępu do Internetu. Oto niektóre skróty używane powszechnie w Stanach Zjednoczonych:

edu – instytucja edukacyjna
com – operator komercyjny, jak Compuserve albo przedsiębiorstwo
org – organizacja non-profit
gov – urząd lub władza publiczna
net – operator regionalny
bit – Bitnet

Obok tego typu skrótów można także spotkać międzynarodowe oznaczenia państw, jak na przykład:

ar – Argentyna
au – Australia
ca – Kanada
de – Niemcy
eg – Egipt
fr – Francja
gh – Ghana
jp – Japonia
mx – Meksyk
sg – Singapur
uk – Wielka Brytania
za – Afryka Południowa

Środkowy fragment adresu (nazwa serwera, z którego korzysta komputer) może przybierać różną formę. Dla serwera America Online brzmi on po prostu „aol". Przykładem adresu na America Online może być:

JDoe@aol.com

Operator Compuserve używa dwuczęściowego numeru zamiast nazwy i słowa „Compuserve" jako serwera. Oto przykład:

76424.156@Compuserve.com

Nazwy niektórych serwerów edukacyjnych są proste. Mój adres na przykład brzmi: babbie@chapman.edu. Niektóre lokalizacje mogą dysponować więcej niż jednym serwerem, wówczas adres może zawierać więcej elementów, oddzielonych kropką.

Kiedy ludzie podają ci swój adres, znak „@" nazywają „małpą". Tak więc podając ci mój adres, powiedziałbym: „Babbie, małpa, chapman, kropka, edu".

Oprócz wysyłania własnych wiadomości, w większości systemów pocztowych możesz też odpowiadać na wiadomości otrzymane. Kiedy to robisz, komputer automatycznie adresuje twoją odpowiedź do pierwotnego nadawcy, załączając twój adres. Może też automatycznie przedrukować otrzymaną wiadomość jako część twojej odpowiedzi, opatrując niekiedy kolejne wiersze znakiem „>".

Załączenie otrzymanej wiadomości (lub jej części) można wykorzystać, by przypomnieć jej nadawcy, na co odpowiadasz. Możesz przetworzyć jego wiadomość, umieszczając w różnych miejscach swoje odpowiedzi. Oto krótki przykład:

Wiadomość otrzymana:
Patrycja:
Chodźmy w sobotę do kina.
Wolałbym jechać twoim samochodem, jeśli to możliwe.
Odezwij się do czwartku.
Ściskam, Janek

Element odpowiedzi automatycznej:
Patrycja:
>Chodźmy w sobotę do kina.
>Wolałbym jechać twoim samochodem, jeśli to możliwe.
>Odezwij się do czwartku.
>Ściskam, Janek

Odpowiedź przetworzona:
Janek:
>Chodźmy w sobotę do kina.
Świetnie, co chciałbyś obejrzeć?
>Wolałbym jechać twoim samochodem, jeśli to możliwe.
Mój samochód jest w warsztacie. Co z twoim?
Na razie, Patrycja

Jako że możliwe jest wysłanie odpowiedzi na odpowiedź, może się okazać, że wdałeś się w rozmowę bez końca.

Odpowiadając, nie należy więc odsyłać całej otrzymanej wiadomości. Zostaw tylko te jej fragmenty, które przydadzą ci się w sformułowaniu twojej odpowiedzi. Jeśli wyślesz całość, zmuszasz odbiorcę do przedzierania się przez długi dokument w poszukiwaniu twojego komentarza. Nadmiar słów zużywa też zasoby sieci, nie mówiąc już o twardym dysku odbiorcy.

Zgodnie z powstającą etykietą mailowania, DUŻE LITERY wykorzystywane są jako równoważnik okrzyku lub dla dodatkowego zaakcentowania słów i powinny być używane oszczędnie. Możesz napisać: „MUSZĘ mieć nowy komputer albo zwariuję", ale pisanie całego tekstu wielkimi literami wyraża głośny krzyk.

Od czasu do czasu natkniesz się na wielkie litery, które nie mają sensu, jak BTW czy IMHO. Oto niektóre skróty używane w kulturze Internetu – to tylko niewielka część tych, które możesz spotkać:

BTW	– *By the way* (Przy okazji)
FYI	– *For your information* (Dla twojej informacji)
IMHO	– *In my humble opinion* (Moim skromnym zdaniem)
IMNSHO	– *In my not so humble opinion* (Moim nie tak znowu skromnym zdaniem)
FAQ	– *Frequently asked questions* (Często zadawane pytania)
LOL	– *Laughing out loud* (Ale ubaw)
RTFM	– *Read the f——g manual* (Przeczytaj p ... ą instrukcję)

Ponieważ nie możesz w pisemnych wiadomościach przekazać wyrazu twarzy czy tonu głosu, wymyślono kilka konwencji, by temu zaradzić. Możesz zamieścić znaki, wyrażające emocje, na przykład: „Po tym jak mi przypomniałeś o egzaminie w poniedziałek, cały dzień zrobił się do kitu". Możesz też użyć symbolu, jak „K", na oznaczenie faktu, że dzień był do kitu.

Bardziej twórczym rozwiązaniem jest tworzony ciągle przez użytkowników Internetu zbiór „emotikonów", służących wyrażaniu emocji przez e-mail. Oto kilka przykładów. Żeby zrozumieć, o co chodzi, musisz przekręcić głowę o 90 stopni w lewo.

:-)	szczęśliwy
8-)	szczęśliwy w okularach
8-)=	szczęśliwy w okularach i z brodą
:-(	nieszczęśliwy
:-O	zdziwiony
:-p	pokazujący język

Prawdopodobnie jesteś w stanie wymyślić nowe znaki tego typu i zaprezentować je społeczności Internetu.

```
  ---
 / øø \
 o
 \_/
```

Listy mailingowe

Poza wysyłaniem wiadomości między pojedynczymi osobami, może ci się przydać także jedna z tysięcy elektronicznych konwersacji. Najczęściej ich nazwa zawiera słowa *listserv*. Dwa inne duże systemy to *listproc* i *majordomo*. Jest kilkanaście takich list, dotyczących metod prowadzenia badań społecznych.

METODY (*METHODS*) na przykład to ogólnodostępna lista mailingowa, głównie dla ludzi nauczających metod badań społecznych. Każdy z użytkowników może wysłać wiadomość na listę i wiadomość ta pojawi się w skrzynkach wszystkich pozostałych użytkowników. Jeśli jeden z nich zechce odpowiedzieć, odpowiedź także pojawi się u wszystkich użytkowników listy. W efekcie nauczyciele metod badawczych mogą dyskutować o często występujących problemach i wspólnie poszukiwać rozwiązań.

Aby zapisać się na listę mailingową, wysyłasz wiadomość do komputera zarządzającego daną listą. W przypadku listy METODY (*METHODS*) odbywa się to następująco:

1. Wyślij wiadomość na adres *listserv@mail.unm.edu*.
2. Pozostaw pole „tytuł" puste.
3. Wyślij wiadomość o treści: „zapisz Jane Doe na listę METODY" (*subscribe Methods Jane Doe*) – zamiast Jane Doe wstaw swoje nazwisko.

Po zapisaniu się możesz wysyłać wiadomości na adres methods@mail.unm.edu.

To standardowa procedura, aczkolwiek istnieją pewne odmiany i wyjątki. Aby zapisać się na którąkolwiek z poniżej wymienionych list, wyślij wiadomość, zastępując w adresie nazwę listy słowem *listserv* i użyj nazwy listy w treści wiadomości, jak to przed chwilą zademonstrowałem na przykładzie METOD. Powinieneś otrzymać wiadomość z potwierdzeniem i z adresem, na który możesz wysyłać wiadomości do

innych użytkowników. (Zauważ, że wnioski administracyjne, na przykład o zapisanie na listę lub wykreślenie z niej, wysyłasz na inny adres niż ten, na który trafiają wiadomości przeznaczone do umieszczenia na liście).

Listy mailingowe są bezpłatne. Nie musisz też brać udziału w rozmowach – wystarczy tylko słuchać, jeśli chcesz. Możesz też wypisać się w każdej chwili. Oto kilka innych list, które mogą cię zainteresować. Zauważ, że w adresach nie ma przerw ani odstępów, nawet jeśli są dłuższe niż jeden wiersz.

por@gibbs.oit.unc.edu
badania opinii publicznej

social-theory@mailbase.ac.uk
teoria społeczna

qualrs-l@uga.cc.uga.edu
badania jakościowe

qual-software@mailbase.ac.uk
oprogramowanie do badań jakościowych

cjust-l@cunyvm.cuny.edu
dyskusja o sądownictwie karnym

demographic-list@coombs.anu.edu.au
australijska lista demograficzna
Uwaga: jeśli chcesz się zapisać, wyślij wiadomość do
majordomo@coombs.anu.edu.au
zamiast na „listserv"

familysci@ukcc.uky.edu
dyskusja z zakresu nauki o rodzinie

ipe@csf.colorado.edu
dyskusja o międzynarodowej ekonomii politycznej

ncs-l@umdd.umd.edu
dyskusja o Krajowym Badaniu Przestępczości

qualnet@chimera.sph.umn.edu
dyskusja o badaniach jakościowych

socgrad@ucsd.edu
dyskusja studentów socjologii

uncjin-l@albnyvm1
sieć informacyjna Trybunału Karnego Narodów Zjednoczonych

sos-data@unc.edu
dyskusja o danych z zakresu nauk społecznych

To tylko kilka list interesujących dla badacza społecznego. W dodatku ich liczba wzrasta codziennie. Zgodnie z efektem domina odkryjesz jednak niebawem, że zapisanie się na jedną listę doprowadzi cię do kolejnych, te zaś do następnych. Z całą pewnością informacji, które mogłyby ci się przydać, jest więcej, niż jesteś w stanie zebrać i przeczytać.

Gopher

Wyobraź sobie, że twoja przyjaciółka napisała program komputerowy albo tekst, który chciałaby, żebyś obejrzał. Prawdopodobnie skopiowałaby po prostu plik na dyskietkę i dała ci, żebyś mógł go przegrać na swój komputer. Wyobraź sobie teraz, że masz dziesiątki tysięcy przyjaciół, z których każdy postępuje w ten właśnie sposób. Oczywiście nie mogą wszyscy dać ci plików na dyskietkach. Od tego są usługi *gopher*, wykonujące *protokół transferu plików (file transfer protocol – ftp)*.

Programy takie jak Turbogopher czy Fetch wymyślono, abyś mógł się łączyć z komputerami na całym świecie. Poza programem do ftp potrzebujesz tylko adresu („nazwy hosta") dostępnego komputera. Być może do komputera twojej uczelni można dotrzeć przez *gopher*.

Jako że jest to coś więcej niż tylko zapisywanie się na listę mailingową czy wysyłanie wiadomości, powinieneś postarać się o jakąś pomoc na miejscu, jeśli chcesz podłączyć się do usługi typu *gopher*. Kiedy połączysz się z odległym komputerem, będziesz mógł przeszukać katalogi i pliki tam dostępne. Musisz znaleźć plik w tamtym komputerze i przekopiować go do własnego. Przekonasz się, że na początku popełnia się przy tym wiele błędów. Program do ftp jest jednak zaprojektowany tak, by pomóc ci tworzyć kopie (na przykład poprzez kliknięcie przycisku z napisem „kopiuj").

Wśród dostępnych na komputerach całego świata plików znajdziesz teksty, gry, programy dźwiękowe, programy do analizy danych itp. Nigdy jednak nie dowiesz się, jak wiele różnorodnych materiałów jest do dyspozycji, dopóki sam się o tym nie przekonasz. Możesz zacząć od zajrzenia do Books A to Z Gopher Servers pod adresem http://booksatoz.com/Gopher.htm.

Strony www

WWW (*World Wide Web*) to być może najbardziej ekscytujący aspekt Internetu w naszych czasach. Działa to trochę jak sieć stron ftp – źródła informacji rozrzucone po całym świecie – dostęp do informacji jest jednak nieporównanie łatwiejszy, a format prezentacji bardziej atrakcyjny.

Aby uzyskać dostęp do stron www, potrzebujesz przeglądarki internetowej, w rodzaju Microsoft Internet Explorer albo konkurencyjnej przeglądarki wyprodukowanej przez NetScape. Komercyjni usługodawcy internetowi, jak America Online, oferują własne przeglądarki. Następnie wchodzisz do Web URL – *uniform resource locator*, czyli systemu lokalizacji zasobów i ruszasz w podróż. Przekonasz się, że niemal każda strona www zawiera połączenia (tzw. linki) w formie wyróżnionych słów lub przycisków, których naciśnięcie przenosi cię pod inny adres.

Adresy stron www na ogół zaczynają się od „http://", ale współczesne wyszukiwarki często uzupełniają tę część adresu za ciebie. Możesz na przykład odwiedzić Uniwersytet Chapmana pod adresem

http://www.chapman.edu/

– ale zapewne wystarczy, gdy wpiszesz część adresu, zaczynając od liter „www". Skoro znajdziesz się na miejscu, odkryjesz wiele informacji, łącznie z listą zajęć prowadzonych na tej uczelni. Jeśli klikniesz „Wilkinson College", będziesz miał m.in. okazję zapoznać się z jednostkami administracyjnymi wchodzącymi w jego skład. Kliknij „Nauki społeczne" (*Social Science*), potem zaś „Socjologia" (*Sociology*). Zobaczysz wtedy m.in. listę kadry tego wydziału (*Faculty*). Kliknij ją, a następnie „Babbie" – aby otrzymać zdjęcie autora niniejszego podręcznika.

Gdybyś kliknął „Wydział Nauk Politycznych" (*Political Science Department*), znalazłbyś interesujące linki, odsyłające cię do Białego Domu czy Izby Reprezentantów. Pokazuje to, jak silne są wzajemne powiązania między różnymi stronami www.

Korzystanie z www wymaga często odnalezienia adresów nieznanych ci jeszcze stron. Możesz na przykład powziąć zamiar wyszukania wszystkiego, co się da, na konkretny temat. W tym celu możesz użyć jednej z setek dostępnych w sieci *wyszukiwarek*.

Wyszukiwarki zawierają wykazy ogromnej ilości stron internetowych stworzonych na świecie. Niektóre z nich, jak Yahoo!, są uporządkowane tematycznie, podobnie jak katalogi biblioteczne. Możesz przechodzić od ogólnego hasła w rodzaju „Historia" do coraz bardziej szczegółowych tematów, póki nie dotrzesz do stron, które cię interesują. Inne wyszukiwarki zorganizowane są inaczej. W każdej z nich musisz po prostu wpisać hasło oznaczające interesujący cię temat, na przykład „Przestępczość w USA" czy „Marilyn Monroe". W toku pracy wyszukiwarki może się okazać, że powinieneś zawęzić pole wyników, na przykład uzupełniając lub zmieniając hasło. Wiele wyszukiwarek odnajduje w sieci oprócz stron www także sieciowe grupy dyskusyjne czy nawet usługi *gopher*.

Ponieważ wyszukiwarki pracują różnie, dobrze jest: (1) sprawdzić więcej niż jedną wyszukiwarkę, jeśli chce się przeprowadzić naprawdę dogłębne poszukiwania i (2) przeczytać wskazówki do przeszukiwania (albo przeszukiwania zaawansowanego), które można zwykle znaleźć na stronach wyszukiwarek.

Oto adresy kilku popularnych wyszukiwarek. Jeśli chciałbyś znać ich więcej, jest na to prosty sposób: możesz użyć jednej z wyszukiwarek do... wyszukiwania wyszukiwarek!

http://www.altavista.com
Alta Vista
http://www.askjeeves.com
Ask Jeeves
http://www.excite.com
Excite!
http://www.hotbot.com
Hotbot

Wymienione wyszukiwarki mają profil ogólny. Poniżej przedstawiam natomiast niektóre strony www, odnoszące się do badań społecznych. Pamiętaj jednak, że adresy w sieci zmieniają się często, zwłaszcza te bardziej złożone, oznaczające zasoby dokumentów na danej stronie. Jeśli nie możesz znaleźć strony o danym adresie złożonym, możesz spróbować usuwać pojedynczo od końca kolejne człony

adresu (oddzielone „slashami"). Wypróbuj taki obcięty adres; jeśli nie działa, odetnij kolejny człon. Kiedy dotrzesz do strony podstawowej, możesz na ogół przejść pod żądany adres za pomocą wskazanych tam połączeń.

Miłego surfowania!

Książki

gopher://ftp.std.com/11/obi/book/
http://www.bookwire.com/
http://ask.elibrary.com/

Fundacje

FEDIX-Federal Informational Exchange,Inc.
http://www.id.ucsb.edu/detche/library/fedix.html/
National Science Foundation
http://www.nsf.gov/
Lista fundacji na Yahoo!
http://www.yahoo.com/Business_and_Economy/Organizations/

Gazety

„Asahi Shimbun" (Japonia)
http://www.asahi.com/english/english.html
„Chicago Sun Times"
http://www.suntimes.com/
„Christian Science Monitor"
http://www.csmonitor.com/
„The Standard"
http://www.thestandard.com.hk/
„Houston Chronicle"
http://www.chron.com/
„India World"
http://www.indiaworld.com/
„Irish Times"
http://www.ireland.com/
„Jerusalem Post"
http://www.jpost.com/
„Joong-ang Daily" (Korea)
http://joongangdaily.joins.com/
„New York Times"
http://www.nytimes.com/
„The New Zealand Herald"
http://nzherald.co.nz/
„St. Petersburg Times" (Rosja)
http://www.sptimes.ru/
„San Jose Mercury News"
http://www.sjmercury.com/
„Singapore Business Times"
http://www.asia1.com.sg/biztimes/
„The Star" (Malezja)
http://www.thestar.com.my/
„Sydney Morning Herald"
http://www.smh.com.au/
„The Times of London"
http://www.timesonline.co.uk/
„USA Today"
http://www.usatoday.com/
Yahoo (Reuters News Service)
http://news.yahoo.com/

Dane demograficzne

Princeton
http://opr.princeton.edu/archive/archive. html/
Uniwersytet Michigan
http://www.umich.edu/
Uniwersytet Stanowy Floryda
http://www.fsu.edu/
Uniwersytet Stanowy Pennsylwania
http://www.pop.psu.edu/

Dane na temat opinii publicznej

Organizacja Gallupa (Gallup Organization)
http://www.gallup.com/
Generalny Sondaż Społeczny (*General Social Survey at ICPSR*)
http://www.icpsr.umich.edu/gss
Wydział Socjologii Queen's College
http://www.soc.qc.edu/
Roper Center przy Uniwersytecie Connecticut
http://www.ropercenter.uconn.edu/

Problemy społeczne

Aborcja
http://dir.yahoo.com/Health/Reproductive_Health/Abortion/
Dzieci
http://www.yahoo.com/Society_and_Culture/Cultures_and_Groups/Children/
Feminizm
http://www.igc.apc.orghome/womensnet/index.html/
Głód
http://www.thp.org/index.html/
Ludzie starsi
http://www.yahoo.com/Society_and_Culture/Cultures_and_Groups/Seniors/

Niepełnosprawni
http://www.yahoo.com/Society_and_Culture/Disabilities/
Prawa człowieka
http://www.yahoo.com/Society_and_Culture/Issues_and_Causes/Human_Rights/
Prawa obywatelskie
http://www.yahoo.com/Society_and_Culture/Issues_and_Causes/Civil_Rights/
Przestępczość
http://www.yahoo.com/Society_and_Culture/Crime/
Seksualność
http://www.yahoo.com/Society_and_Culture/Sexuality/
Zagadnienia związane z płcią
http://www.yahoo.com/Society_and_Culture/Gender/

Zasoby związane z badaniami społecznymi
Australian National University
http://www.anu.edu.au/
California State University, Hayward
http://www.csubak.edu/
Clearing House for Social Sciences & Social Issues
http://www.lib.umich.edu/house/tree/social.html/
Corporate Information Database
http://www.sec.gov/edgarhp.htm/
Economic and Social Research Council Data Archive
http://www.essex.ac.uk/
Institute for Research in Social Science
http://www.irss.unc.edu/
Institute for Research in the Social Science
http://www.ciesin.org/datasets/irss/irss.html/
Inter-university Consortium for Politcal and Social Research
http://www.icpsr.umich.edu/
National Election Survey
http://www.umich.edu/~nes/
National Survey of Families & Households
http://www.ssc.wisc.edu/nsfh/home.htm/
Public Use Microdata Samples
http://www.ipums.umn.edu/
Research for Social Science
http://www.ssrn.com/
Social Sciences Information Gateway
http://sosig.esrc.bris.ac.uk/
Social Sciences Data Centre
http://www.lib.virginia.edu/subjects/sociology/
University of California, San Diego
http://ssdc.ucsd.edu/
The Urban Information Center
http://oseda.missouri.edu:80/uic/
www Wirtual Library of Sociology
http://www.w3.org/hypertext/DataSources/bySubject/Sociology/Overview.html/
Yahoo-Social Science
http://www. yahoo.com/Social_Science/
Yahoo-Society and Culture
http://www.yahoo.com/Society_and_Culture/

Rządy stanowe i lokalne
http://www.statelocalgov.net/index.cfm/

Rząd USA
Biuro Spisowe
http://www.census.gov/
Biuro Statystyk Pracy
http://stats.bls.gov/
Biuro Statystyk Transportu
http://www.bts.gov/
Centra Kontroli Medycznej
http://www.cdc.gov/
Centralna Agencja Wywiadowcza (CIA)
http://www.cia.gov/
Departament Oświaty
http://www.ed.gov/
Agencja Ochrony Środowiska
http://www.epa.gov/
Federalne Biuro Śledcze
http://www.fbi.gov/
Izba Reprezentantów
http://www.house.gov/
Krajowe Centrum Statystyk Oświatowych
http://www.ed.gov/NCES/
Krajowy Instytut Zdrowia
http://www.nih.gov/
Biały Dom
http://www.whitehouse.gov/

Historia świata
http://www.hyperhistory.com/online_n2/History_n2/a.html/

Powyższa lista nie jest nawet czubkiem *czubka* góry lodowej. W jednym z ostatnio przeprowadzonych badań oceniono, że jest po-

nad 800 milionów stron www. W tej chwili jest ich znacznie więcej. Gdybym mógł przedstawić ci kompletny wykaz stron www dotyczących badań społecznych, wykaz ten byłby nieaktualny jeszcze zanim mój rękopis dotarłby do wydawcy, nie mówiąc już o chwili, w której dotarłby do ciebie. Aby zrozumieć, jak szybko rozwija się sieć, pomyśl o chwilach, w których pracowałeś nad czymś tak intensywnie, że w końcu zabierało ci to całe dnie. Nie myślałeś wówczas właściwie o niczym innym. Wiedz zatem, że w tej chwili dziesiątki tysięcy ludzi pracują z takim samym zapałem nad stronami www. Niezliczone mnóstwo innych pracuje nieco wolniej. A wszystko to tylko dla ciebie.

Uzupełnienie do wydania polskiego

(opracowała Agnieszka Kloskowska-Dudzińska)

Wyszukiwarki

Polski czytelnik, poza wymienionymi przez autora, może korzystać także z wyszukiwarek polskojęzycznych. Szybką wyszukiwarką, dostępną po polsku, jest Google (www.google.com). Inne i popularne wyszukiwarki dostępne są na portalach internetowych, takich jak np. Wirtualna Polska (www.wp.pl), Onet (www.onet.pl) Hoga (www.hoga.pl) czy Interia (www.interia.pl).

Prasa

Wiele polskich gazet i czasopism prowadzi wersje on-line.

Dotyczy to m.in. obu największych dzienników: „Rzeczpospolitej" (www.rp.pl), na stronie której dostępny jest m.in. serwis ekonomiczny i jego „Notowania", czyli wskaźniki ekonomiczne, a także archiwum, oraz „Gazety Wyborczej" (www1.gazeta.pl/wyborcza), która prowadzi także portal internetowy (www.gazeta.pl), zawierający wiele przydatnych linków. Wersję on-line ma też „Trybuna" (www.trybuna.com.pl).

Z głównych polskich czasopism społeczno-politycznych wersje on-line mają m.in. „Wprost" (www.wprost.pl), „Polityka" (polityka.onet.pl), „Gazeta Polska" (www.gazetapolska.pl), „Newsweek" (http://newsweek.redakcja.pl/ uruchamia się także przez www.newsweek.pl), „Najwyższy Czas" (www.nczas.com).

Z czasopism z zakresu nauk społecznych dostępny jest np. kwartalnik „Kultura i Społeczeństwo" (www.isppan.wa/kis/index.htm).

Instytucje administracji publicznej

- Wszystkie instytucje rządowe w Polsce mają własne strony internetowe. Ich adresy najłatwiej odnaleźć poprzez stronę internetową Kancelarii Prezesa Rady Ministrów (www.kprm.gov.pl), na której pod zakładką „Administracja" zamieszczony jest wykaz ministerstw i urzędów centralnych wraz z linkami. Strona KPRM zawiera także liczne inne użyteczne zakładki, m.in. „Polska" wraz z dalszą zakładką „Polska w liczbach", gdzie możesz znaleźć podstawowe informacje o państwie.
- Links do urzędów centralnych i ministerstw oraz parlamentów na świecie znajdziesz też na stronie internetowej Sejmu RP (www.sejm.gov.pl) pod ikonką „WWW". Na tej samej stronie pod zakładką „Prawo" zamieszczone są: konstytucja, ordynacje wyborcze oraz Internetowy System Informacji Prawnej (opisy aktów prawnych oraz ujednolicone teksty ustaw), a pod zakładką „Polska – UE" akty prawne związane z akcesją oraz linki do polskich stron dotyczących integracji europejskiej.
- Własne strony internetowe mają także urzędy wojewódzkie i marszałkowskie oraz wiele gmin. Ich adresy z łatwością znajdziesz za pomocą którejś z wyszukiwarek. Wiele informacji o Polsce oraz liczne linki (m.in. do urzędów administracji lokalnej – w serwisie Samorządy) zawiera portal www.polska.pl, prowadzony przez NASK.

Ogólnodostępne dane statystyczne

Instytucje publiczne, gromadzące dane statystyczne, są zobowiązane do ich udostępniania.

- Kopalnię danych statystycznych stanowi strona internetowa Głównego Urzędu Statystycznego (www.stat.gov.pl). Zawiera ona m.in. bazy danych regionalnych i de-

mograficznych, a także dane społeczno-gospodarcze, m.in. raport z wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2002 i raport z wyników Powszechnego Spisu Rolnego 2002, jak również raporty z wyników Spisów Powszechnych w województwach. Na stronie GUS zamieszczone są linki do urzędów statystycznych na świecie i różnych światowych serwisów statystycznych.

- Wyniki wyborów i referendów (na najniższym poziomie agregacji danych, a więc na poziomie obwodów głosowania) dostępne są na stronie internetowej Państwowej Komisji Wyborczej (www.pkw.gov.pl).

Dane z badań społecznych

Dane z tego typu badań rzadko są ogólnodostępne. Jest jednak kilka wyjątków.

- Dane z Polskiego Generalnego Sondażu Społecznego (będącego polskim odpowiednikiem wielokrotnie przytaczanych przez autora badań amerykańskich *GSS – General Social Survey*) to najobszerniejszy zbiór danych z badań społecznych dostępny w Polsce. Sondaże przeprowadzane są obecnie co dwa lata (początkowo co rok), w jednolitej metodologii, umożliwiającej zarówno porównania w czasie, jak i porównania międzynarodowe na podstawie wyników podobnych sondaży w innych krajach (np. amerykańskiego GSS). Strona internetowa PGSS znajduje się pod adresem: http://pgss.iss.uw.edu.pl/osrodki/obs/pgss/pl/index.php/
- Polski Generalny Sondaż Wyborczy, przeprowadzany przez Instytut Studiów Politycznych PAN, to przykład innych badań, z których dane zostały udostępnione online. Badania te są polskim odpowiednikiem przytaczanych przez autora badań *NES – National Election Survey*. Dane dostępne są pod adresem: www.isppan.waw.pl/pgsw/index.html/
- Obszerny zbiór omówień wyników badań o charakterze niekomercyjnym można znaleźć na stronach internetowych Centrum Badania Opinii Publicznej w postaci komunikatów (www.cbos.com.pl). Zaletą tego zbioru jest powtarzalność niektórych tematów.
- Pracownia Badań Społecznych w Sopocie na swoich stronach internetowych (www.pbssopot.com.pl) udostępnia najnowsze wyniki badań, dotyczących preferencji politycznych oraz oceny instytucji i urzędów w Polsce. Cennym uzupełnieniem aktualnych wyników jest możliwość dostępu do wcześniejszych edycji obu badań (poprzez Archiwum) i przeprowadzenia analizy trendów.
- Liczne raporty z badań społecznych od 1989 roku udostępnia przez Internet także TNS OBOP (www.tns-global.pl). Jednak raporty z ostatnich dwóch lat udostępniane są odpłatnie.
- Omówienie wyników niektórych badań społecznych i politycznych znajdziesz także na stronach internetowych polskiego oddziału Ipsosu (www.demoskop.com.pl lub www.ipsos.pl).

Dodatkowe lektury

Cassell James W., *Internet Data Archives for Sociologists*, praca zaprezentowana Amerykańskiemu Stowarzyszeniu Socjologicznemu, Nowy Jork, 19 sierpnia 1996.

Ferrante Joan, Vaughn Angela, *Let's Go Sociology: Travels on the Internet*, Wadsworth, Belmont, Cal. 1999. Przeznaczony dla studentów przewodnik po ponad 500 stronach www, z krótkim wprowadzeniem do problematyki sieci.

Kardas Edward P., Milford Tommy M., *Using the Internet for Social Research and Practice*, Wadsworth, Belmont, Cal. 1996. Jak sam tytuł wskazuje, to wprowadzenie do Internetu koncentruje się na zasobach związanych z badaniami społecznymi. Książka jest pełna adresów stron www i innych informacji, a załączona dyskietka ułatwia korzystanie z nich.

Kurland Daniel J., *The Net, the Web and You*, Wadsworth, Belmont, Cal. 1996. Dobre wprowadzenie do tematów omawianych w tym załączniku i do zagadnień komunikacji w ogóle, począwszy od słowa mówionego, skończywszy na druku.

Załącznik C

Raport z badań

Wprowadzenie

Ta książka dotyczy wielu czynności składających się na *prowadzenie* badań społecznych. W tym załączniku zajmiemy się problemem często zaniedbywanym, a mianowicie relacjonowaniem wyników badań innym ludziom. Jeśli badania nie są należycie zakomunikowane, wszystkie starania w zakresie poprzednio omówionych procedur są bezużyteczne. Oznacza to, że dobry raport z badań wymaga przede wszystkim dobrej polszczyzny (albo hiszpańszczyzny czy innego języka, którym się posługujesz). Kiedy każemy liczbom „mówić samym za siebie", liczby najczęściej pozostają nieme. Kiedy używamy skomplikowanej terminologii i konstrukcji pojęciowych, zakres komunikacji jest ograniczony. Każdy badacz powinien przeczytać i co najmniej raz na trzy miesiące odświeżać znajomość świetnej książeczki Williama Strunka Juniora i E. B. White'a, zatytułowanej *The Elements of Style**. Jeśli zrobisz to uczciwie, to nawet gdyby umknęło ci 90% treści, pozostała część wystarczy, by cię zrozumiano i doceniono twoje wyniki.

Raport naukowy pełni kilka funkcji. Po pierwsze, komunikuje konkretne dane i idee. Raport powinien przedstawiać je jasno i na tyle szczegółowo, by umożliwić innym kompetentną ocenę. Po drugie, raport naukowy powinno się traktować jako wkład w całokształt wiedzy naukowej. Zachowując należytą skromność, powinieneś jednak być świadom, że twój raport dodaje coś do naszej wiedzy o zachowaniach społecznych. Raport powinien wreszcie stymulować i ukierunkowywać dalsze badania.

Kilka podstawowych kwestii

Abstrahując od powyższych ogólnych wskazówek, każdy raport służy różnym celom. Raport stosowny dla jednego celu może się okazać zupełnie nieodpowiedni dla innego. Ta część książki przedstawia kilka podstawowych kwestii z tym związanych.

Odbiorcy

Zanim naszkicujesz swój raport, musisz się zastanowić, kto – zgodnie z twoimi oczekiwaniami – będzie go czytał. Powinieneś na ogół rozróżnić uczonych i zwykłych czytelników. Jeśli piszesz raport dla tych pierwszych, możesz zrobić pewne założenia co do ich wiedzy i w związku z tym przedstawić pewne punkty ogólnie, zamiast wdawać się w szczegóły. Możesz też używać języka bardziej technicznego niż w stosunku do niespecjalistów.

* Wyd. 4, Macmillan, New York 1999. Inna dobra książka dotycząca pisania: R. W. Birchfield, *The New Fowler's Modern English Usage*, wyd. 3 Oxford University Press, New York 1998.

Pamiętaj jednak również, że każda nauka ma swoje szkoły. Terminy, założenia i poszczególne techniki bliskie twojemu otoczeniu, mogą wprawiać w zakłopotanie innych uczonych. Socjolog religii, pisząc dla socjologów innych specjalności, powinien na przykład objaśnić swoje spostrzeżenia bardziej szczegółowo, niż gdyby zwracał się wyłącznie do socjologów religii.

Forma i długość raportu

Moje uwagi w tej części odnoszą się zarówno do raportów pisanych, jak i ustnych prezentacji. Każda z tych form wpływa jednak na kształt raportu.

Dobrze jest zastanowić się nad wielością raportów, które mogą opisywać jeden projekt badawczy. Na początek być może zechcesz przygotować krótką *notę badawczą*, w celu publikacji w czasopiśmie naukowym. Tego typu raporty mają na ogół od jednej do pięciu stron znormalizowanego maszynopisu, powinny być zwięzłe i precyzyjne. W tekście o tak niewielkiej objętości nie można prezentować szczegółowo stanu badań w danej dyscyplinie, musisz więc skrócić uwagi metodologiczne. Powinieneś w zasadzie wyjaśnić czytelnikowi, dlaczego uważasz, że krótka nota jest uzasadniona na gruncie twoich wyników, potem zaś przedstawić same wyniki.

Badacze muszą często przygotowywać raporty dla sponsorów swoich badań. Rozmiar takich raportów może być bardzo różny. Przygotowując raport tego typu, powinieneś przede wszystkim pamiętać o odbiorcy – naukowcu lub laiku – i o powodach, dla których sponsoruje on twój projekt. Zanudzanie sponsora wynikami, które nie mają dla niego znaczenia, jest błędem strategicznym i przejawem złych manier. Niekiedy warto jednak krótko przedstawić ewentualny wkład twoich badań w dotychczasową wiedzę naukową.

Dokumenty robocze albo *monografie* to kolejna forma zdawania sprawy z przebiegu badań. Zwłaszcza w przypadku dużych i złożonych projektów badawczych przydadzą ci się uwagi i komentarze na temat przeprowadzonej przez ciebie analizy i interpretacji danych. Dokument roboczy nie jest ostateczną prezentacją i zaprasza do zgłaszania uwag. Dokument roboczy może również mieć różną objętość, może omawiać całość lub tylko część wyników badań. Dokument roboczy nie zagraża twojej reputacji naukowej, możesz więc zupełnie swobodnie prezentować nie do końca uzasadnione przypuszczenia, prosząc o ich ocenę.

Wiele projektów badawczych kończy się przedstawieniem referatu na konferencji naukowej. Często pełni on tę samą funkcję, co dokumenty robocze. Możesz prezentować pomysły i odkrycia interesujące twoich kolegów, prosząc o komentarze. Objętość takich *referatów naukowych* może być różna, zależnie od organizacji spotkania, zawsze jednak lepiej powiedzieć za mało niż za dużo. W dokumencie roboczym można swobodnie snuć przypuszczenia, dotyczące wstępnych wniosków, nie można jednak zmuszać słuchaczy konferencji, by siedzieli i słuchali, jak rozwijasz ustnie takie przypuszczenia. Zainteresowani słuchacze mogą potem zapytać o szczegóły, niezainteresowani zaś będą szczęśliwi, mogąc wymknąć się wcześniej.

Prawdopodobnie najbardziej powszechną formą raportu z badań jest *artykuł* w piśmie naukowym. Objętość może znowu być różna; powinieneś sprawdzić objętość artykułów publikowanych dotąd w tym czasopiśmie. Możesz przyjąć, że na ogół 25 stron znormalizowanego maszynopisu to dobra objętość. Dalsze wskazówki dotyczące organizacji raportu oparte są zasadniczo na strukturze artykułu, nie będę więc rozwodził się nad tym teraz; ograniczę się do uwagi, że prace semestralne studentów powinny być zbudowane według tego modelu. Dobra praca semestralna to praca, która byłaby też dobrym artykułem.

I w końcu *książka* – najbardziej prestiżowa forma raportu z badań. Jest długa i szczegółowa jak dokument roboczy, ale znacznie bardziej dopracowana. Publikacja wyników badań jako książki przydaje im więcej wagi i znaczenia, stwarza więc zarazem szczególne zobowiązania wobec czytelnika. Zdarza się oczywiście, że twoi koledzy zgłaszają uwagi, które skłonią cię do zmiany poglądów, inni czytelnicy mogą jednak bezkrytycznie zaakceptować twoje wyniki.

Cel raportu

Wcześniej mówiliśmy już co nieco na temat różnych celów projektów badawczych. Przygotowując raport, powinieneś pamiętać o różnorodności tych celów.

Niektóre raporty mogą koncentrować się na ogólnym zbadaniu przedmiotu, czyli *eksploracji*. Wnioski są w takich przypadkach wstępne i niekompletne. Powinieneś jasno poinformować czytelnika o eksploracyjnym celu badań, a jednocześnie wskazać na braki konkretnego projektu. Raport tego typu służy wskazaniu sposobów ściślejszego badania danego przedmiotu.

Większość raportów z badań zawiera element opisowy, odzwierciedlający cel badań, które dokumentują. Musisz starannie odróżnić od siebie opisy, które dotyczą tylko próby, i te, które dotyczą całej populacji. Podaj też czytelnikowi wskazówki dotyczące prawdopodobnego błędu w przypadku każdego opisu opartego na wnioskowaniu statystycznym.

Celem wielu raportów jest wyjaśnianie: wskazanie zależności przyczynowych między zmiennymi. Biorąc pod uwagę specyfikę twoich adresatów, powinieneś nakreślić dokładnie zasady wyjaśniania, które legły u podstaw twoich obliczeń i wniosków. Podobnie jak w przypadku opisu, powinieneś uświadomić czytelnikowi ograniczony zakres pewności twoich wniosków.

Niektóre raporty proponują też podjęcie określonych działań. Na przykład badacz zajmujący się uprzedzeniami może sugerować możliwości obniżenia poziomu uprzedzeń, pojawiające się na gruncie wyników jego badań. Cel ten okazuje się jednak często bardzo problematyczny, ponieważ wartości i poglądy badacza mogą wpływać na jego propozycje. Nie ma oczywiście przeciwwskazań dla oparcia tego typu propozycji na osobistych wartościach; propozycje działań zgłaszane przez badacza muszą jednak znajdować uzasadnienie w danych. Badacze powinni być zatem szczególnie ostrożni, wyjaśniając założenia, zgodnie z którymi przechodzą od danych empirycznych ku propozycjom działania.

Struktura raportu

Struktura raportu może się zmieniać w zależności od jego formy i celu, jednak przedstawienie ogólnego formatu prezentacji danych z badań może okazać się użyteczne. Dalsze uwagi stosują się przede wszystkim do artykułu w czasopiśmie, jednak po paru modyfikacjach mogą też znaleźć zastosowanie do innych form raportu z badań.

Cel i przegląd treści

Zawsze dobrze jest zacząć od krótkiego zdania na temat celów badania i głównych wyników analizy. W artykule taki przegląd treści przybiera niekiedy formę *streszczenia*.

Niektórzy badacze mają z tym trudności. Twoja analiza mogła na przykład angażować w znacznym stopniu czynności detektywistyczne, z których istotne wnioski ujawniają się dopiero w wyniku śmiałej dedukcji i manipulacji danymi. Być może chciałbyś więc, aby czytelnik przeszedł wraz z tobą ten pasjonujący proces, tak byś mógł odsłaniać przed nim swoje odkrycia, trzymając go w napięciu i zaskakując. Taka forma prezentacji może mieć istotną wartość edukacyjną, jeśli daje należyty obraz całości badań. Nie zapominaj jednak, że niektórzy czytelnicy mogą nie być zainteresowani śledzeniem całej twojej relacji z badań; jeśli zaś nie będą znali z góry celu i ogólnych wniosków, może być im trudno zrozumieć znaczenie całej pracy.

Stare sądowe porzekadło mówi: „Powiedz im, co masz zamiar im powiedzieć, powiedz im, a potem im powiedz, co im powiedziałeś". Dobrze byłoby, gdybyś stosował się do tych słów.

Przegląd literatury przedmiotu

Następną rzeczą jest wskazanie, w którym miejscu twoje badania wpisują się w kontekst całości wiedzy naukowej. Po przedstawieniu ogólnego celu badań powinieneś zapoznać czytelnika z dotychczasowymi pracami na ten temat, informując o punktach zgodnych i spornych wśród poprzednich badaczy.

W niektórych przypadkach możesz zdecydować się na krytykę akceptowanych dotąd poglądów. Dokonaj dokładnego przeglądu prac, które doprowadziły do ich przyjęcia, następnie wskaż czynniki, których dotąd nie uwzględniono, lub błędy logiczne poprzednich badań.

Jeśli zajmujesz się uzgodnieniem stanowisk dotychczasowych badaczy, powinieneś podsumować badania przemawiające na korzyść jednego poglądu, potem te, które wspierają drugi pogląd, na koniec zaś zasugerować wyjaśnienie przyczyn niezgodności.

Twój przegląd literatury przedmiotu pełni dla czytelnika funkcję wstępnej bibliografii. Jest on zestawieniem dotychczasowych badań na dany temat. Nie można jednak przesadzać – powinieneś unikać wstępu ciągnącego się przez trzy strony, zawierającego wyliczenie wszystkich prac w danej dziedzinie. Funkcję wyczerpującej informacji bibliograficznej najlepiej spełni bibliografia na końcu tekstu, przegląd literatury zaś powinien koncentrować się na tych pracach, które bezpośrednio dotyczą twojego tematu.

Unikanie plagiatu

Kiedy cytujesz cudze prace, musisz jasno informować, kto co powiedział. Oznacza to, że musisz unikać *plagiatu*, czyli przypadkowej lub celowej kradzieży cudzych słów i/lub pomysłów i prezentacji ich jako własnych. Ponieważ problem ten zdarza się studentom często i zawsze budzi wątpliwości, zajmijmy się nim bliżej. Oto kilka podstawowych reguł dotyczących plagiatu:

- Jeśli używasz dosłownie wypowiedzi innego autora, musisz użyć cudzysłowów i zamieścić dokładny przypis, wskazujący źródło cytatu, aby czytelnik mógł umieścić cytat w jego oryginalnym kontekście. Zaczerpnięcie z cudzej pracy fragmentu liczącego więcej niż osiem słów bez przypisu stanowi pogwałcenie federalnych ustaw o prawach autorskich.
- Nie wolno ci także zmieniać redakcji albo konkretnych słów innego autora i prezentować tak przetworzonej wersji jako własnej.
- Nie możesz wreszcie przedstawiać cudzych pomysłów jako własnych, nawet jeśli formułujesz je w zupełnie inny sposób.

Poniższe przykłady powinny objaśnić, co jest, a co nie jest dopuszczalne w posługiwaniu się cudzą pracą.

Dzieło oryginalne:

„Prawa wzrostu

Systemy są jak dzieci: skoro już się pojawią, są. Nie odchodzą. Wręcz przeciwnie, wykazują godną uwagi trwałość. I nie tylko trwają – rosną, a w miarę wzrostu rozszerzają się. Potencjał wzrostu systemów został odkryty – wstępnie i w sposób niepewny – przez Parkinsona, który stwierdził, że systemy administracyjne utrzymują stopę wzrostu na wysokości 5–6% rocznie, niezależnie od ilości pracy do wykonania. Parkinson miał w tym zakresie zupełną rację, a jego zasługą jest wszczęcie poważnych badań nad tym istotnym zagadnieniem. Teraz jednak ogłosimy to, czego Parkinson nie zauważył – odpowiednik prawa Parkinsona dla wszystkich systemów.

System sam z siebie wzrasta o 5–6% rocznie.

Prawo to jest tylko wstępem do najogólniejszego możliwego sformułowania – twierdzenia kosmologii systemów o wielkim wybuchu.

Systemy rozszerzają się, by wypełnić cały znany wszechświat".

(Gall 1975, s. 12–14)

Spójrzmy teraz na kilka *dopuszczalnych* sposobów wykorzystania pracy Galla w twoim referacie semestralnym.

- **Dopuszczalne**: John Gall, w swojej pracy *Systemantics*, rysuje humorystyczną analogię między systemami a dziećmi: „Systemy są jak dzieci: skoro już się pojawią, są. Nie odchodzą. Wręcz przeciwnie, wykazują godną uwagi trwałość. I nie tylko trwają – rosną, a w miarę wzrostu rozszerzają się"*.

* John Gall, *Systemantics: How Systems Work and Especially How They Fail*, Quadrangle, New York 1975, s. 12–14.

Dopuszczalne: John Gall ostrzega, że systemy są jak dzieci. Gdy stworzy się system, on już dalej po prostu jest. Co gorsza jednak, jak zauważa, systemy stale rosną*.

Dopuszczalne: Sugeruje się także, jakoby systemy miały naturalną tendencję do trwania, a nawet wzrostu i rozprzestrzeniania się (Gall 1975, s. 12).

Zauważ, że ta ostatnia forma wymaga, by w bibliografii znalazł się kompletny przypis – podobnie jak w tej książce. Dopuszczalne są także pełne przypisy na dole strony lub przypisy końcowe. Przejrzyj podręczniki edytorskie różnych organizacji, na przykład APA (*American Psychological Association*) lub ASA (*American Sociological Association*), albo *Chicago Manual of Style*, aby zapoznać się z odpowiednimi formatami przypisów.

Oto kilka *niedozwolonych* sposobów wykorzystania tego samego materiału, stanowiących przykład kilku powszechnych błędów.

Niedozwolone: W poniższej pracy chciałbym przyjrzeć się niektórym cechom systemów społecznych, które tworzymy w ramach naszych organizacji. Po pierwsze – systemy są jak dzieci: skoro już się pojawią, zostają na dłużej. Nie odchodzą. Wręcz przeciwnie, wykazują szczególną trwałość. Nie tylko trwają – rosną. [Niedozwolone jest używanie wprost słów kogoś innego bez cudzysłowów i kompletnego przypisu].

Niedozwolone: W poniższej pracy chciałbym przyjrzeć się niektórym cechom systemów społecznych, które tworzymy w ramach naszych organizacji. Po pierwsze, systemy są jak dzieci: kiedy się rodzą, zostają z nami na dłużej. Nie odchodzą, lecz trwają. W istocie zaś nie tylko trwają – rosną. [Niedopuszczalne jest przerabianie cudzej pracy i przedstawianie jej jako własnej].

Niedozwolone: W poniższej pracy chciałbym przyjrzeć się niektórym cechom systemów społecznych, które tworzymy w ramach naszych organizacji. Jedną z istotnych cech, które zauważyłem, jest to, że system, raz stworzony, zdaje się nigdy nie przemijać. Jest wręcz przeciwnie – wydaje się, że mają one tendencje wzrostowe. Pod tym względem można powiedzieć, że systemy bardzo przypominają dzieci. [Niedopuszczalne jest parafrazowanie myśli kogoś innego i przedstawianie ich jako twoich własnych].

Każdy z przytoczonych przykładów ilustruje zjawisko plagiatu, które jest poważnym wykroczeniem. Trzeba przyznać, że istnieją pewne „szare strefy". Niektóre idee należą mniej lub bardziej do domeny publicznej, nie są więc niczyją „własnością". Możesz też sam dojść do konkluzji, którą ktoś już wcześniej zapisał. Jeśli jakaś sytuacja budzi twoje wątpliwości, przedyskutuj ją zawczasu z prowadzącym zajęcia.

Omówiłem ten temat nieco dokładniej, ponieważ – mimo że musisz umieścić swoje badanie w kontekście tego, co inni już zrobili i powiedzieli – nieprawidłowe wykorzystanie ich prac stanowi poważne wykroczenie. Zdobycie umiejętności unikania popełniania plagiatu jest częścią twojego „dojrzewania" jako uczonego.

Plan badań i jego realizacja

Bywa, że raport z badań, zawierający ciekawe odkrycia i wnioski, staje się źródłem frustracji dla czytelnika, który nie jest w stanie zrozumieć metodologii badań. Wartość wszystkich odkryć naukowych zależy w znacznym stopniu od sposobu gromadzenia i analizy danych.

Referując plan badań i sposób przeprowadzenia sondażu zawsze zamieść w raporcie następujące dane: określenie populacji, metodę doboru próby, wielkość próby, metodę zbierania danych, odsetek braków odpowiedzi oraz metody obróbki i analizy danych. Jeśli stosujesz inne metody, podaj podobne szczegóły badania. Doświadczony badacz potrafi podać te informacje w stosunkowo niewielu słowach, nie pomijając zarazem żadnych danych istotnych dla oceny badań przez czytelnika.

Analiza i interpretacja

Umieściwszy swoje badania w kontekście dotychczasowej wiedzy na dany temat i opisawszy sposób, w jaki je zaprojektowano i przeprowa-

*John Gall, *Systemantics...*, s. 12.

dzono, powinieneś przystąpić do prezentacji danych. W niniejszym załączniku znajdziesz za chwilę dalsze wskazówki na ten temat, na razie jednak czas na kilka uwag ogólnych.

Prezentacja danych, ich przekształceń i twoich interpretacji powinny stanowić logiczną całość. Czytelnik popada we frustrację, zapoznając się ze zbiorem na pozór nie powiązanych ze sobą analiz i wniosków z obietnicą, że wszystko to zostanie uporządkowane dalej w raporcie. Każdy krok analizy powinien mieć jasne znaczenie w momencie, w którym go czynisz – powinieneś przedstawić uzasadnienie każdej poszczególnej analizy, podać adekwatne dane i zinterpretować wyniki, następnie zaś wskazać kolejne kroki, ku którym te wyniki prowadzą.

Podsumowanie i konkluzje

Zgodnie ze wspomnianym sądowym porzekadłem podsumowanie raportu jest niezbędne. Unikaj powtarzania każdego wniosku po kolei, ale dokonaj przeglądu najważniejszych z nich, zwracając raz jeszcze uwagę na ich ogólne znaczenie.

Raport powinien kończyć się stwierdzeniem odkryć, które poczyniłeś, oraz sugestiami co do kierunku przyszłych badań przedmiotu. Wiele artykułów kończy się zdaniem: „Oczywiste jest, że potrzeba jeszcze wielu badań". Wniosek ten jest najprawdopodobniej zawsze prawdziwy, zwykle jednak pozbawiony wartości, dopóki nie przedstawisz bardziej szczegółowych propozycji późniejszych badań. Powinieneś omówić braki twoich badań i wskazać, w jaki sposób braków tych można w przyszłości uniknąć.

Raport z analizy danych – wskazówki

Prezentacja przeprowadzonych analiz powinna dostarczać maksimum informacji, jednak bez nadmiernej szczegółowości. Najlepiej spełnisz to kryterium, jeśli będziesz nieustannie sprawdzał swój raport, by ustalić, czy udaje ci się osiągnąć cele, które omawiam poniżej.

Jeśli używasz danych ilościowych, zaprezentuj je tak, by czytelnik mógł na własną rękę dokonać wszystkich obliczeń. Na przykład przedstawiając tabelę z danymi procentowymi, zadbaj o to, by czytelnik był w stanie sam podsumować kategorie i przeliczyć odsetki. Czytelnicy powinni otrzymać dostatecznie wiele informacji, by mogli przeliczyć procenty w tabeli w odwrotnym kierunku, niż to uczyniłeś w swojej prezentacji.

Opisz wszystkie aspekty analizy ilościowej, by się upewnić, że ilość szczegółów umożliwi innemu analitykowi przeprowadzenie wtórnej analizy na podstawie tych samych danych. Oznacza to, że musi on być w stanie stworzyć te same indeksy i skale, tabele i równania regresji, otrzymać takie same czynniki itd. Rzadko oczywiście do tego dochodzi, jeśli jednak raport stworzy taką możliwość, czytelnik będzie znacznie lepiej przygotowany do oceny wyników badań.

Przedstawiaj szczegóły. Jeśli prowadzisz badania jakościowe, musisz podać dostatecznie dużo szczegółów, by czytelnik miał poczucie, że towarzyszył ci w czasie obserwacji. Nie wystarczy przedstawić dane, które przemawiają na korzyść twojej interpretacji – musisz też podzielić się wiedzą o danych, które jej przeczą. Powinieneś zapewnić dość informacji, żeby czytelnik mógł dojść do wniosków innych niż twoje – choć możesz mieć nadzieję, że twoja interpretacja jest najbardziej sensowna. Czytelnik powinien być w stanie powtórzyć całe badania na własną rękę, niezależnie od tego, czy wymagają one obserwacji uczestniczącej wśród fanek zespołów heavy-metalowych, eksperymentu dotyczącego obrad ławy przysięgłych czy jakiegokolwiek innego rodzaju badań. Przypomnij sobie, że powtarzalność jest zasadniczą normą w nauce. Pojedyncze badania nie dowodzą jeszcze niczego – trzeba do tego serii badań. Jeśli dane badania nie mogą zostać powtórzone, nie może też być sensownej serii badań.

Włącz do raportu materiał ilustracyjny. Wspominałem już wcześniej, jaką wagę ma łączenie w raporcie danych i interpretacji. Oto bardziej szczegółowe wskazówki, jak to zrobić. Tabele, wykresy i rysunki, jeśli takowe istnieją, powinny zostać włączone do tekstu ra-

portu w sąsiedztwie omawiającego je fragmentu tekstu. Studenci opisują niekiedy analizę w części głównej raportu, a wszystkie tabele zamieszczają w załącznikach. Procedura ta bardzo utrudnia lekturę. Na ogół najlepiej jest: (1) opisać powód, dla którego prezentuje się daną tabelę, (2) przedstawić ją oraz (3) podsumować ją i zinterpretować.

Wyciągaj jasne wnioski. Wprawdzie badania prowadzimy zazwyczaj, aby dojść do wniosków ogólnych, powinieneś jednak zawsze odnotować szczegółowe podstawy do tego typu konkluzji. W przeciwnym razie możesz sprawić, że twój czytelnik zaakceptuje wnioski pozbawione uzasadnienia.

Wskaż ograniczenia i warunki wpływające na ocenę wniosków. Na ogół sam jesteś najbardziej świadom braków i wstępnego charakteru twoich wniosków, powinieneś więc pozwolić czytelnikowi skorzystać z twojej wiedzy. Jeśli tego nie zrobisz, możesz spowodować, że przyszłe badania pójdą w złym kierunku i doprowadzą do marnotrawstwa funduszy przeznaczonych na badania.

Jak już nadmieniłem na początku tego załącznika, raport z badań powinien być napisany możliwie dobrym stylem. Niektórym ludziom łatwo jest pisać przejrzyście, innym przychodzi to z trudem – wszystkim jednak trudniej jest pisać dobrze niż kiepsko. Odsyłam cię raz jeszcze do książki Strunka i White'a. Każdemu badaczowi dobrze zrobi zastosowanie się do następującej procedury. Przeczytaj Strunka i White'a. Sprawdź. Znowu przeczytaj Strunka i White'a. Sprawdź raz jeszcze. Będzie to czynność trudna i czasochłonna, ale tak to już jest z nauką.

Bezbłędnie zaprojektowane, starannie przeprowadzone i błyskotliwie zanalizowane badanie będzie zupełnie bezwartościowe, jeśli nie potrafisz przekazać innym swoich wyników. Powyższy tekst miał ci dostarczyć kilku wskazówek w drodze do tego celu. Najlepszymi przewodnikami są logika, jasność i uczciwość. I wreszcie – nic nie zastąpi praktyki.

D

Kwestionariusz struktury gospodarstwa domowego z badania GSS

Krajowe Centrum Badania Opinii (*National Opinion Research Center – NORC*) przy Uniwersytecie Chicago jest jednym z czołowych centrów badań społecznych na świecie. Prowadzony przez nie Generalny Sondaż Społeczny (*General Social Survey* – GSS) stał się dla badaczy głównym źródłem danych z zakresu nauk społecznych. Być może zauważyłeś, że wiele danych w tej książce pochodzi z GSS. Wypada więc, abyśmy zaczerpnęli z GSS przykład kwestionariusza wywiadu. Na kilku następnych stronach zamieszczony został kwestionariusz, stosowany przez ankieterów w celu zebrania podstawowych danych demograficznych o gospodarstwach domowych, które znalazły się w próbie. Tego typu kwestionariusze są istotną częścią warsztatu badacza społecznego. Podobnie jak mikroskop jest narzędziem odpowiednim do obserwacji komórek i cząstek, teleskop zaś do obserwacji odległych planet i gwiazd, kwestionariusz tego typu okazuje się często najlepszym narzędziem obserwacji przedmiotu nauk społecznych.

GSS
AMERICA'S SOCIAL SURVEY

Nazwisko ankietera

Identyfikator ankietera

5325 NORC
Styczeń 1990

KWESTIONARIUSZ STRUKTURY GOSPODARSTWA DOMOWEGO KSGD

KOMPLETNY RAPORT Z JEDNOSTKOWEGO BADANIA

WPROWADZENIE:

Dzień dobry, nazywam się (*NAZWISKO*) i jestem z Krajowego Centrum Badania Opinii przy Uniwersytecie Chicago (POKAŻ IDENTYFIKATOR).

(Wysłaliśmy Państwu ostatnio list w wyjaśnieniem, że) Państwa gospodarstwo domowe zostało wybrane, by wziąć udział w tegorocznym GSS - Amerykańskim Sondażu Społecznym.

Prowadzimy te badania w całym kraju od ponad 15 lat, zdobywając wiedzę o opiniach ludzi na takie tematy, jak szkolnictwo, przestępczość, wydatki rządowe czy wojskowość. W tym roku kilka gospodarstw domowych w Państwa miejscowości weźmie udział w tych ważnych badaniach.

Po pierwsze, chciałbym się upewnić, że właściwie zapisałem Państwa adres - (PRZECZYTAJ Z KARTY PRZYDZIAŁU ALBO Z RAMKI PONIŻEJ: ULICĘ I NUMER DOMU, NUMER MIESZKANIA LUB INNY OPIS GD). JEŚLI SIĘ ZGADZA, ZAKREŚL KRATKĘ PONIŻEJ:

☐ *(PRZEJDŹ DO P. 1, STR. 2)*

JEŚLI WYSTĘPUJĄ JAKIEKOLWIEK RÓŻNICE, SPRAWDŹ DANE I UPEWNIJ SIĘ, ŻE SĄ WŁAŚCIWE. RÓŻNICE WYJAŚNIJ PONIŻEJ:

JEŚLI RÓŻNIC NIE DA SIĘ WYJAŚNIĆ, ZADZWOŃ DO KOORDYNATORA, ZANIM PRZEJDZIESZ DO DALSZEJ CZĘŚCI.

ANKIETER: TUTAJ PRZYKLEJ KARTĘ PRZYDZIAŁU
LUB
JEŚLI DODAJESZ TEN PRZYPADEK
JAKO UPRZEDNIO POMINIĘTE GD,
PRZEPISZ INFORMACJE Z ORYGINALNEJ KARTY

SONDAŻ 5325 NUMER PRZYPADKU

PSU

SEG PT 1 WIERSZ BLK

MIEJSCOWOŚĆ

ADRES GD

OPIS GD

A, B LUB C X LUB Y

(NR PRZYPADKU)	01-06/
(PSU)	07-09/
(SEG)	10-12/
(PT)	13/
(NR WIERSZA)	14-23/
(A,B LUB C)	24/
(X LUB Y)	25/
(INTID)	26-31/

1. Tylko jeden członek Państwa gospodarstwa domowego będzie uczestnikiem tych badań. Aby w sposób naukowy wybrać tę osobę, muszę najpierw zebrać dane wszystkich osób, które aktualnie tutaj mieszkają.

Godzina rozpocz

2. Proszę podać mi nazwiska osób, które aktualnie wchodzą w skład tego gospodarstwa domowego. Zacznijmy od głowy gospodarstwa domowego. WPISZ NAZWISKA W WIERSZACH 01-10 PONIŻEJ

3. Może zapomnieliśmy o kimś: o niemowlętach albo małych dzieciach, o sublokatorach, o ludziach, którzy mieszkają tutaj, a chwilowo są w delegacji, na wakacjach, w szkole, w szpitalu itp.?

Tak ☐ DODAJ POZOSTAŁE OSOBY W WIERSZACH 01-10 PONIŻEJ

Nie ☐ przejdź do P. 4

4. Czy przebywają tu obecnie jakieś osoby - goście, przyjaciele lub krewni - które na ogół tu nie mieszkają?

Tak ☐ WPISZ GOŚCI W WIERSZACH 11-14 PONIŻEJ

Nie ☐ PRZEJDŹ DO P. 5

PO PYTANIACH 1-4 ZADAJ KAŻDEJ OSOBIE PYTANIA 5-8

9. Czy jakakolwiek osoba, którą poprzednio wymieniliśmy, przebywa obecnie poza domem?
JEŚLI NIE, ZAKREŚL KRATKĘ I PRZEJD

DO P. 12 ☐

JEŚLI TAK, PRZEJDŹ DO P. 10 PONIŻEJ.

Imię Nazwisko	5. Jaka jest relacja między (OSOBĄ) a (GŁOWĄ RODZINY)?	6. ZAKODUJ PŁEĆ (UPEWNIJ SIĘ, JEŚLI NIE JEST OCZYWISTA) M K	7. Ile lat (OSOBA/GŁOWA RODZINY) skończyła w dniu swoich ostatnich urodzin?	8. JEŚLI 13 I WIĘCEJ, ZADAJ PYTANIE: Czy (OSOBA) pozostaje obecnie w związku małżeńskim, jest wdową/wdowcem, jest rozwiedziona, w separacji lub nigdy nie pozostawała w związku małżeńskim? Zm Wd Ro Se N	10. Kto przebywa obecnie poza domem? ZAZNACZ (✓) WIERSZ KAŻDEJ OSOBY, KTÓRA JEST POZA DOMEM.	11. ZAPYTAJ O KAŻDĄ OSOBĘ, KTÓ ZAZNACZYŁEŚ (✓) W P. 10: Gdzie jest teraz (OSOBA): jest w innym gospodarstwie domowym; czy podróżuje; czy przebywa w jakiejś instytucji- na przykład na studiach, w szpitalu lub co w tym rodzaju; czy też gdzi indziej? ZAKREŚL KOD ZGODNIE Z INSTRUKCJĄ Inne	Podróż	Instytucja	Ir
01	49 GŁOWA	50 1 2	51-52	53 1 2 3 4 5	54	1 WYKREŚL	2 ZOSTAW	3 WYKREŚL	4
02	56	57 1 2	58-59	60 1 2 3 4 5	61	1 WYKREŚL	2 ZOSTAW	3 WYKREŚL	4
03	63	64 1 2	65-66	67 1 2 3 4 5	68	1 WYKREŚL	2 ZOSTAW	3 WYKREŚL	4
04 POCZĄTEK ARKUSZA 03	07	08 1 2	09-10	11 1 2 3 4 5	12	1 WYKREŚL	2 ZOSTAW	3 WYKREŚL	4
05	14	15 1 2	16-17	18 1 2 3 4 5	19	1 WYKREŚL	2 ZOSTAW	3 WYKREŚL	2 4

| | | |

12. Przystąpię teraz do naukowego wyboru jednego członka tego gospodarstwa domowego, który będzie uczestniczył w badaniach. Prowadząc wywiady tylko z osobami dobranymi w ten sposób, mamy pewność, że opinie, które zbieramy, dokładnie odzwierciedlają poglądy w całym kraju.

KROK 1: UPRAWNIENI: 18 lat i więcej

IMIONA NIE WYKREŚLONE W ZESTAWIENIU CZŁONKÓW GOSPODARSTWA DOMOWEGO (P. 2)

JEŚLI JEST TYLKO JEDNA OSOBA UPRAWNIONA, PRZEJDŹ DO P. 13 I UMÓW SIĘ NA WYWIAD Z TĄ OSOBĄ.

JEŚLI WIĘCEJ NIŻ JEDNA, POSTĘPUJ DALEJ ZGODNIE ZE WSKAZANYMI PONIŻEJ KROKAMI 2-6.

KROK 2: WPISZ NAZWISKA UPRAWNIONYCH OSÓB W TABELĘ W KOLEJNOŚCI WIEKU MALEJĄCO.

ZESTAWIENIE OSÓB UPRAWNIONYCH

	NR WIERSZA	WIEK	NAZWISKA OSÓB UPRAWNIONYCH
NAJSTARSZY (niekoniecznie głowa gospodarstwa domowego)	1		
	2		
	3		
	4		
	5		
NAJMŁODSZY	6		

TABELA DOBORU

KROK 3: UŻYJ TABELI DOBORU PO PRAWEJ STRONIE, BY WYBRAĆ RESPONDENTA SPOŚRÓD OSÓB UPRAWNIONYCH

LICZBA OSÓB UPRAWNIONYCH WPISANYCH POWYŻEJ	PRZEPROWADŹ WYWIAD Z OSOBĄ WPISANĄ W WIERSZU O NUMERZE
DWIE	PRZYDZIAŁU, DOTYCZĄCĄ DOBORU RESPONDENTA. (Jeśli jest to „uprzednio pominięte g.d.", wpisz liczby z pierwotnego g.d.)
TRZY	
CZTERY	
PIĘĆ	
SZEŚĆ I WIĘCEJ	

KROK 4: OTOCZ OKRĘGIEM WIERSZ RESPONDENTA W ZESTAWIENIU OSÓB UPRAWNIONYCH

KROK 5: WPISZ TUTAJ IMIĘ I NAZWISKO WYBRANEGO RESPONDENTA

KROK 6: *PRZEPROWADŹ WYWIAD Z TĄ OSOBĄ. POPROŚ O TELEFON I ADRES DO KORESPONDENCJI (P. 13 I P. 14), ABY UŁATWIĆ KONTYNUACJĘ.*

13. *Gdybym chciał się skontaktować z (WYBRANYM RESPONDENTEM), pod jaki numer mogę zadzwonić?*

Podany numer telefonu: (.....) .....

KIERUNKOWY NUMER

A. Zakoduj lokalizację telefonu:

W domu 1 22/

U sąsiada 2

Inna (OKREŚL) 3

B. Jeśli nie podano numeru telefonu, zakoduj

Brak telefonu 4

Odmowa 5

14. Gdybym chciał coś przesłać (WYBRANEMU RESPONDENTOWI), na jaki adres mogę to zrobić?

ULICA I NUMER DOMU	
POCZTA, MIEJSCOWOŚĆ	
STAN	KOD

PODZIĘKUJ RESPONDENTOM ZA POŚWIĘCONY CZAS I ZA POMOC

GODZINA ZAKOŃCZENIA	
CZAS TRWANIa	

NATYCHMIAST PO OPUSZCZENIU GOSPODARSTWA DOMOWEGO WYPEŁNIJ POLA A-D PONIŻEJ I UZUPEŁNIJ RAPORT Z DOBORU RESPONDENTA:

A. Data wypełnienia KBGD ☐☐ ☐☐ ☐☐ 23-25
MIESIĄC DZIEŃ ROK

B. Imię i nazwisko informatora KBGD _____

C. Wpisz numer wiersza informatora KBGD z Kwestionariusza struktury gospodarstwa domowego (STR. 2, P. 4) ☐☐ 26-27/
(Jeśli informator pochodzi spoza gospodarstwa domowego, nr wiersza = 22)

D. Rasa mieszkańców gospodarstwa domowego (zaobserwowana)

Biała 1 28/

Czarna 2

Indianie 3

Azjaci, rasy wschodnie 4

Inna, mieszana, nieobserwowalna 5

E. Wpisz numer wiersza wybranego respondenta z Kwestionariusza struktury gospodarstw domowych (STR. 2, P. 4) ☐☐ 29-30

KSGD-4 KSGD-5

RAPORT W RAZIE BRAKU WYWIADU (RBW)

Nazwisko (jeśli znane):_____

Numer telefonu (jeśli dostępny):_____
Proszę zakreślić właściwy kod:

KSGD nie został wypełniony: 1 31/
KSGD został wypełniony, wywiad nie został przeprowadzony 2

1. Dlaczego nie udało się wypełnić KSGD/przeprowadzić wywiadu pod tym adresem?

BRAK GD 32-33/
Porzucone ..... (Odp. na P. 2). 01
Zniszczone ..... (Odp. na P. 2). 02
Lokal użytkowy ..... (Odp. na P. 2). 03
Brak takiego adresu/brak takiego GD ..... (Odp. na P. 2). 04
Kwatera zbiorowa ..... (Odp. na P. 2). 05
Domek letniskowy ..... (Odp. na P. 2). 06
Nie nadający się do użytku jako miejsce stałego zamieszkania ..... (Odp. na P. 2). 07
Czasowe użytkowanie (krócej niż 1 miesiąc) .... (Odp. na P. 2). 08
Brak GD z innych powodów ..... (Odp. na P. 2). 09
W trakcie budowy ..... (Odp. na P. 2). 10
PUSTE (Przejdź do P. 7). 11
ODMOWA (Przejdź do P. 4). 12
PRZERWANIE WYWIADU (Przejdź do P. 4). 13
PO CZTERECH TELEFONACH NIKOGO NIE UDAŁO SIĘ ZASTAĆ W DOMU (Przejdź do P. 3). 14
RESPONDENT POZOSTAJE POZA ZASIĘGIEM PRZEZ CAŁY CZAS TRWANIA BADAŃ (Przejdź do P. 5). 15
PROBLEM JĘZYKOWY (OKREŚL JĘZYK) (Przejdź do P. 7). 16
ZBYT CHORY (OPISZ SYTUACJĘ NA STR. 7) (Przejdź do P. 7). 17
INNE POWODY (OPISZ SYTUACJĘ NA STR. 7) (Przejdź do P. 7). 18

2. JEŚLI BRAK GD:
Opisz powody wyczerpująco w niniejszym RBW, następnie przejdź do P. 19 ..... 34-35/

3. JEŚLI NIE ZASTAŁEŚ NIKOGO W DOMU PO 4 TELEFONACH:
Dlaczego twoim zdaniem tak trudno było zastać mieszkańców tego domu? 36-37/

A. Opisz, jakie wysiłki podjąłeś w celu uzyskania informacji o mieszkańcach 38-39/

4. JEŚLI SPOTKAŁEŚ SIĘ Z ODMOWĄ LUB WYWIAD ZOSTAŁ PRZERWANY:

Czy nastąpiła odmowa ze strony respondenta?

Tak.....(Odp. na A)....1
40/
Nie.....(Odp. na B)....2
NW, KSGD nie został ukończony.....Przejdź do P. 6)....8
A. JEŚLI TAK: Dlaczego respondent (odmówił/przerwał wywiad)? (podaj dosłowne uwagi i powody)...41-4

(Przejdź do P. 6

B. JEŚLI NIE: Dlaczego nie udało ci się porozmawiać z respondentem? 43-44

(Przejdź do P. 6

5. JEŚLI CZASOWO NIEOSIĄGALNY PRZEZ CAŁY CZAS TRWANIA BADAŃ:
Z jakiego powodu? 45-46

A. Kiedy R będzie osiągalny? 47-48

6. Czy wysłano do respondenta list przypominający?
Tak ..... (Odp. na A) ...1 49
Nie ..... (Odp. na B) ...2
NW 8

A. JEŚLI TAK: Czy rozmawiałeś z R po otrzymaniu przez niego listu?

Tak 1 50
Nie ..... (Odp. na B) ...2
NW 8

B. JEŚLI NIE: Dlaczego nie?

ODPOWIEDZ NA P.7, CHYBA ŻE BRAK GD

7. Czy udało ci się w ogóle porozmawiać z kimś w tym gospodarstwi domowym (niekoniecznie z mieszkańcem)?
Tak.....(Przejdź do P. 8) 1 51/
Nie.....(Odp. na A) ...2
A. Dlaczego nie? 52-53/

8. *Czy zdarzyło się, że przy nieudanych próbach kontaktu w GD znajdowali się ludzie, którzy nie otwierali drzwi?*

Tak, zdecydowanie 1 54/
Tak podejrzewam 2
Nie mam powodu tak sądzić 3

ODPOWIEDZ NA PYTANIA 9-17 NA PODSTAWIE INFORMACJI I/LUB KONTAKTU Z MIESZKAŃCAMI GD, SĄSIADAMI LUB NA PODSTAWIE WIADOMOŚCI Z INNYCH ŹRÓDEŁ.

9. Jaki jest orientacyjny dochód rodziny respondenta? (UWAGA: JEŚLI JEST WIĘCEJ RODZIN, A R NIE ZOSTAŁ OKREŚLony, OCEŃ DLA RODZINY PIERWOTNEJ)
Niski 1 55/
Średni 2
Wysoki 3
Nie wiem/nie da się zaobserwować 8

10. Czy udało się wypełnić listę domowników?
Tak.....Przejdź do P. 16 .... 1 56/
Nie.....Odp. na P. 11-20 2

11. Jakiej rasy są domownicy?

Białej.....(zdecydowanie) 1 57/
Białej.....(prawdopodobnie) 2
Czarnej .....(zdecydowanie) 3
Czarnej.....(prawdopodobnie) 4
Latynoskiej 5
Żółtej/wschodniej 6
Innej (OPISZ) 7
Nie można określić.....

12. Oszacuj liczbę dorosłych mieszkających w GD.
liczba dorosłych ☐☐ 58-59/

13. Oszacuj liczbę dorosłych mężczyzn mieszkających w GD.
liczba dorosłych ☐☐ 60-61/

14. Czy w GD mieszka małżeństwo (albo para żyjąca jak małżeństwo)?
Tak.....1 62/
Nie.....2
Nie wiem.....3

15. Jak oceniasz wiek głowy GD?
Poniżej 30.....1 63/
30-64.....2
65 lub więcej.....3
Nie wiem.....8

16. Typ budowli 64-65/
Przyczepa kempingowa.....01
Domek jednorodzinny.....02
Dom dwurodzinny bliźniak.....03
Dom dwurodzinny podzielony piętrami.....04
Dom wolno stojący dla 3-4 rodzin.....05
Segment (3 lub więcej domów połączonych w rzędzie....06
Apartamentowiec (5 lub więcej mieszkań, do 3 pięter).07
Apartamentowiec (5 lub więcej mieszkań, powyżej 3 pięter).....08
Apartament w budynku z lokalami użytkowymi.....09
Inny (OKREŚL).....10

17. W porównaniu z innymi domami/mieszkaniami w sąsiedztwie, ten dom/mieszkanie był/o twoim zdaniem

Znacznie powyżej przeciętnej.....1 66/
Powyżej przeciętnej.....2
Przeciętne.....3
Poniżej przeciętnej.....4
Znacznie poniżej przeciętnej.....5

18. UWAGI DODATKOWE:

19. Nazwisko ankietera

_____

20. Numer ankietera ☐☐☐☐ – ☐

21. Wpisz nazwisko koordynatora.

KSGD-6

ZAPIS KONTAKTÓW Z GOSPODARSTWEM DOMOWYM

UŻYWAĆ OŁÓWKA

A. Dzień Pn. 1 Wt. 2 Śr. 3 Czw. 4 Pt. 5 Sb. 6 Nd. 7 B. Miesiąc C. Data	D. Godzina Przed 15.00 .....1 15.00-18.00 .....2 Po 18.00 .....3 E. Typ kontaktu Osob....1 Tel.....2 Poczt...3 F. Wynik kontaktu Jeśli 35, 37, 45, 47, 82, 92 wypełnij G	G. Opis odmowy	H. Wynik/powód odmowy. Podaj dosłownie powody i wyjaśnij wszystkie okoliczności.	I. Cel kontaktu	J. Inicjały ankietera
A. ☐ 32/R B. ☐☐ 33-34/R C. ☐☐ 35-36/R	D. ☐ 37/R E. ☐☐ 38/R F. ☐☐ 39-40/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 41/R KSGD Kwest.	J.
A. ☐ 42/R B. ☐☐ 43-44/R C. ☐☐ 45-46/R	D. ☐ 47/R E. ☐☐ 48/R F. ☐☐ 49-50/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 51/R KSGD Kwest.	J.
A. ☐ 52/R B. ☐☐ 53-54/R C. ☐☐ 55-56/R	D. ☐ 57/R E. ☐☐ 58/R F. ☐☐ 59-60/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 61/R KSGD Kwest.	J.
A. ☐ 62/R B. ☐☐ 63-64/R C. ☐☐ 65-66/R	D. ☐ 67/R E. ☐☐ 68/R F. ☐☐ 69-70/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 71/R KSGD Kwest.	J.

POCZĄTEK ARKUSZA 02

A. ☐ 07/R B. ☐☐ 08-09/R C. ☐☐ 10-11/R	D. ☐ 12/R E. ☐☐ 13/R F. ☐☐ 14-15/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 16/R FBGD Kwest.	J.
A. ☐ 17/R B. ☐☐ 18-19/R C. ☐☐ 20-21/R	D. ☐ 22/R E. ☐☐ 23/R F. ☐☐ 24-25/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 26/R FBGD Kwest.	J.
A. ☐ 27/R B. ☐☐ 28-29/R C. ☐☐ 30-31/R	D. ☐ 32/R E. ☐☐ 33/R F. ☐☐ 34-35/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 36/R FBGD Kwest.	J.
A. ☐ 37/R B. ☐☐ 38-39/R C. ☐☐ 40-41/R	D. ☐ 42/R E. ☐☐ 43/R F. ☐☐ 44-45/R	G. M.....K Wiek _____ Inne cechy: nr wiersza KSGD jeśli wypełniony _____	H.	I. 46/R FBGD Kwest.	J.

47-48

KODY STATUSU DLA DOCHODU (kolumna F)

TYMCZASOWY

KSGD	KWEST.
31 Działań nie podjęto	43 R nie ma w domu/Nie można się skontaktować*
32 Nikogo nie ma w domu/Brak odpowiedzi/Zajęty	45 Czas. odmowa/Przerwanie wywiadu
35 Czasowa odmowa/Przerwanie wywiadu	46 Umówiono spotkanie
36 Spotkanie	47 Odwołano spotkanie
37 Odwołano spotkanie	49 Inne
38 Głowa GD nieuchwytna	
39 Inne	

OSTATECZNY (wymaga zatwierdzenia przez koordynatora)

KSGD	KWEST.
10 Brak GD*	13 R nie mówi po angielsku*
11 Puste*	92 Odmowa/Przerwanie wywiadu*
12 Głowa GD nie mówi po angielsku*	94 R nieobecny przez cały okres trwania badań*
82 Odmowa/Przerwanie wywiadu*	97 Inne*
84 Nikogo nie ma w domu/Nie można się skontaktować*	60 Tylko kwest.
85 Głowa GD nieuchwytna*	65 Kwest. i niektóre, nie wszystkie, dokumenty
87 Inne*	67 Kwest. i wszystkie dokumenty

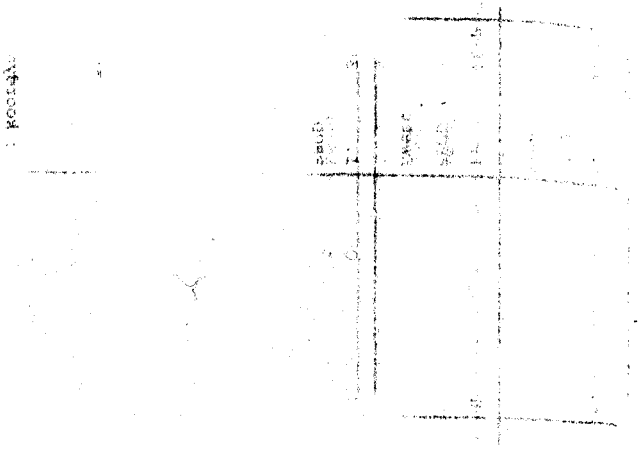
WYPEŁNIJ RAPORT Z WYBORU RESPONDENTA

W RAZIE KONIECZNOŚCI KONTYNUOWAĆ NA NOWYM ARKUSZU

* WYPEŁNIJ RBW

Załącznik
E

Liczby losowe



10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	36188	18510
02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	79666	95725
29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760	64584	96096	98253
05366	04213	25669	26422	44407	44048	37397	63904	45766	66134	75470	66520	34693	90449
91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618
00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808	76655	62028	76630
00725	69884	62797	56170	86324	88072	76222	36086	84637	93161	76038	65855	77919	88006
69011	65795	95876	55293	18988	27354	26575	08625	40801	59920	29841	80150	12777	48501
25976	57948	29888	88604	67917	48708	18912	82271	65424	69774	33611	54262	85963	03547
09763	83473	73577	12908	30883	18317	28290	35797	05998	41688	34952	37888	38917	88050
91567	42595	27958	30134	04024	86385	29880	99730	55536	84855	29080	09250	79656	73211
17955	56349	90999	49127	20044	59931	06115	20542	18059	02008	73708	83517	36103	42791
46503	18584	18845	49618	02304	51038	20655	58727	28168	15475	56942	53389	20562	87338
92157	89634	94824	78171	84610	82834	09922	25417	44137	48413	25555	21246	35509	20468
14577	62765	35605	81263	39667	47358	56873	56307	61607	49518	89656	20103	77490	18062
98427	07523	33362	64270	01638	92477	66969	98420	04880	45585	46565	04102	46880	45709
34914	63976	88720	82765	34476	17032	87589	40836	32427	70002	70663	88863	77775	69348
70060	28277	39475	46473	23219	53416	94970	25832	69975	94884	19661	72828	00102	66794
53976	54914	06990	67245	68350	82948	11398	42878	80287	88267	47363	46634	06541	97809
76072	29515	40980	07391	58745	25774	22987	80059	39911	96189	41151	14222	60697	59583
90725	52210	83974	29992	65831	38857	50490	83765	55657	14361	31720	57375	56228	41546
64364	67412	33339	31926	14883	24413	59744	92351	97473	89286	35931	04110	23726	51900
08962	00358	31662	25388	61642	34072	81249	35648	56891	69352	48373	45578	78547	81788
95012	68379	93526	70765	10592	04542	76463	54328	02349	17247	28865	14777	62730	92277
15664	10493	20492	38391	91132	21999	59516	81652	27195	48223	46751	22923	32261	85653

16408	81899	04153	53381	79401	21438	83035	92350	36693	31238	59649	91754	72772	02338
18629	81953	05520	91962	04739	13092	97662	24822	94730	06496	35090	04822	86774	98289
73115	35101	47498	87637	99016	71060	88824	71013	18735	20286	23153	72924	35165	43040
57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812	12515	98931	91202
30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68668	30429	70735	25499
16631	35006	85900	98275	32388	52390	16815	69298	82732	38480	73817	32523	41961	44437
96773	20206	42559	78985	05300	22164	24369	54224	35083	19687	11052	91491	60383	19746
38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318	29686	03387	59846
31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325
78919	19474	23632	27889	47914	02584	37680	20801	72152	39339	34806	08930	85001	87820
03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833	25570	38818	46920
74426	33278	43972	10119	89917	15665	52872	73823	73144	88662	88970	74492	51805	99378
09066	00903	20795	95452	92648	45454	09552	88815	16553	51125	79375	97596	16296	66092
42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689	05974	52468	16834
16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510	70625	28725	34191
21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240	15957	16572	06004
21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995	43805	33386	21597
55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525	42786	05269	92532
44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911	25650	12682	73572
91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203	71795	99533	50501
91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867	14951	91696	85065
50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405	56087	82790	70925
65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202	94617	23772	07896
27504	96131	83944	41575	10573	08619	64482	73923	36152	05184	94142	25299	84387	34925
37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330	74301	00275	48280
11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84081	66938	93654	59894
37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651	50245	34971	52924
46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345	81073	49106	79860
30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339	58861	74818	46942
63798	64995	46583	09785	44160	78128	83991	42865	92520	83531	80377	35909	81250	54238
82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870	52689	51275	83556
21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194	52799	28225	85762
60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535	12133	14645	23541
43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434	98227	21824	19585
97656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534	03862	78095	50136
03299	01221	05418	38982	55758	92237	26759	86367	21216	98442	08303	56613	91511	75928
79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076	72811	22717	50585
85636	68335	47539	03129	65651	11977	02510	26113	99447	68645	34327	15152	55230	93448
18039	14367	61337	06177	12143	46609	32989	74014	64708	00533	35398	58408	13261	47908
08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639	82163	60859	75567
79556	29068	04142	16268	15387	12856	66227	38358	22478	73373	88732	09443	82558	05250
92608	82674	27072	32534	17075	27698	98204	63863	11951	34648	88022	56148	34925	57031
23982	25835	40055	67006	12293	02753	14827	23235	35071	99704	37543	11601	35503	85171
09915	96306	05908	97901	28395	14186	00821	80703	70426	75647	76310	88717	37890	40129
59037	33300	26695	62247	69927	76123	50842	43834	86654	70959	79725	93872	28117	19233
42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	80799	76536	71255	64239
46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	53203	18098	47625	88684
03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	06216	95787	42579	90730
86591	81482	52667	61582	14972	90053	89534	76036	49199	43716	97548	04379	46370	28672
38534	01715	94964	87288	65680	43772	39560	12918	86537	62738	19636	51132	25739	56947

Skrót na podstawie *Handbook of Tables for Probability Statistics*, wyd. 2, William H. Beyer (red.), The Chemical Rubber Company, Cleveland 1968. Wykorzystano za zgodą The Chemical Rubber Company.

Załącznik
F

Rozkład chi-kwadrat

df	Prawdopodobieństwo													
	,99	,98	,95	,90	,80	,70	,50	,30	,20	,10	,05	,02	,01	,001
1	,$0^3$157	,$0^3$628	,00393	,0158	,0642	,148	,455	1,074	1,642	2,706	3,841	5,412	6,635	10,827
2	,0201	,0404	,103	,211	,446	,713	1,386	2,408	3,219	4,605	5,991	7,824	9,210	13,815
3	,115	,185	,352	,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,341	16,268
4	,297	,429	,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277	18,465
5	,554	,752	1,145	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,070	13,388	15,086	20,517
6	,872	1,134	1,635	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812	22,457
7	1,239	1,564	2,167	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475	24,322
8	1,646	2,032	2,733	3,490	4,594	5,528	7,344	9,524	11,030	13,362	15,507	18,168	20,090	29,125
9	2,088	2,532	3,325	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666	27,877
10	2,558	3,059	3,940	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209	29,588
11	3,053	3,609	4,575	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725	31,264
12	3,571	4,178	5,226	6,304	7,807	9,034	11,340	14,011	15,812	18,549	21,026	24,054	26,217	32,909
13	4,107	4,765	5,892	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	25,472	27,688	34,528
14	4,660	5,368	6,571	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141	36,123
15	5,229	5,985	7,261	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578	37,697
16	5,812	6,614	7,962	9,312	11,152	12,624	15,338	18,841	20,465	23,542	26,296	29,633	32,000	39,252
17	6,408	7,255	8,672	10,085	12,002	13,531	16,338	15,511	21,615	24,769	27,587	30,995	33,409	40,790
18	7,015	7,906	9,390	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	32,346	34,805	42,312
19	7,633	8,567	10,117	11,651	13,716	15,352	18,338	21,689	23,900	27,204	30,144	33,687	36,191	43,820
20	8,260	9,237	10,851	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,410	35,020	37,566	45,315
21	8,897	9,915	11,591	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932	46,797
22	9,542	10,600	12,338	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289	48,268
23	10,196	11,293	13,091	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638	49,728
24	10,856	11,992	13,848	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,270	42,980	51,179
25	11,524	12,697	14,611	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314	52,620
26	12,198	13,409	15,379	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642	54,052
27	12,879	14,125	16,151	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,140	46,963	55,476
28	13,565	14,847	16,928	18,939	21,588	23,647	27,336	31,391	34,027	37,916	41,337	45,419	48,278	56,893
29	14,256	15,574	17,708	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588	58,302
30	14,953	16,306	18,493	20,599	23,364	25,508	29,336	35,530	36,250	40,256	43,773	47,962	50,892	59,703

Dla dużych wartości df, wyrażenie $\sqrt{2\chi^2} - \sqrt{2df} - 1$ można traktować jako zmienną o rozkładzie normalnym standaryzowanym; trzeba tylko pamiętać, że test χ^2 jest zawsze jednostronny.

Źródło: Jestem zobowiązany wykonawcy testamentu Sir Ronalda A. Fishera, Dr. Frankowi Yatesowi oraz Grupie Longman Ltd. Londyn za zezwolenie na przedruk tabeli nr IV z ich książki *Statistical Tables for Biological, Agricultural and Medical Research*, wyd. 6, 1974.

Załącznik
G

Powierzchnie pod krzywą normalną

z	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,0	,0000	,0040	,0080	,0120	,0160	,0199	,0239	,0279	,0319	,0359
0,1	,0398	,0438	,0478	,0517	,0557	,0596	,0636	,0675	,0714	,0753
0,2	,0793	,0832	,0871	,0910	,0948	,0987	,1026	,1064	,1103	,1141
0,3	,1179	,1217	,1255	,1293	,1331	,1368	,1406	,1443	,1480	,1517
0,4	,1554	,1591	,1628	,1664	,1700	,1736	,1772	,1808	,1844	,1879
0,5	,1915	,1950	,1985	,2019	,2054	,2088	,2123	,2157	,2190	,2224
0,6	,2257	,2291	,2324	,2357	,2389	,2422	,2454	,2486	,2517	,2549
0,7	,2580	,2611	,2642	,2673	,2704	,2734	,2764	,2794	,2823	,2852
0,8	,2881	,2910	,2939	,2967	,2995	,3023	,3051	,3078	,3106	,3133
0,9	,3159	,3186	,3212	,3238	,3264	,3289	,3315	,3340	,3365	,3389
1,0	,3413	,3438	,3461	,3485	,3508	,3531	,3554	,3577	,3599	,3621
1,1	,3643	,3665	,3686	,3708	,3729	,3749	,3770	,3790	,3810	,3830
1,2	,3849	,3869	,3888	,3907	,3925	,3944	,3962	,3980	,3997	,4015
1,3	,4032	,4049	,4066	,4082	,4099	,4115	,4131	,4147	,4162	,4177
1,4	,4192	,4207	,4222	,4236	,4251	,4265	,4279	,4292	,4306	,4319
1,5	,4332	,4345	,4357	,4370	,4382	,4394	,4406	,4418	,4429	,4441
1,6	,4452	,4463	,4474	,4484	,4495	,4505	,4515	,4525	,4535	,4545
1,7	,4554	,4564	,4573	,4582	,4591	,4599	,4608	,4616	,4625	,4633
1,8	,4641	,4649	,4656	,4664	,4671	,4678	,4686	,4693	,4699	,4706
1,9	,4713	,4719	,4726	,4732	,4738	,4744	,4750	,4756	,4761	,4767
2,0	,4772	,4778	,4783	,4788	,4793	,4798	,4803	,4808	,4812	,4817
2,1	,4821	,4826	,4830	,4834	,4838	,4842	,4846	,4850	,4854	,4857
2,2	,4861	,4864	,4868	,4871	,4875	,4878	,4881	,4884	,4887	,4890
2,3	,4893	,4896	,4898	,4901	,4904	,4906	,4909	,4911	,4913	,4916
2,4	,4918	,4920	,4922	,4925	,4927	,4929	,4931	,4932	,4934	,4936
2,5	,4938	,4940	,4941	,4943	,4945	,4946	,4948	,4949	,4951	,4952
2,6	,4953	,4955	,4956	,4957	,4959	,4960	,4961	,4962	,4963	,4964
2,7	,4965	,4966	,4967	,4968	,4969	,4970	,4971	,4972	,4973	,4974
2,8	,4974	,4975	,4976	,4977	,4977	,4978	,4979	,4979	,4980	,4981
2,9	,4981	,4982	,4982	,4983	,4984	,4984	,4985	,4985	,4986	,4986
3,0	,4987	,4987	,4987	,4988	,4988	,4989	,4989	,4989	,4990	,4990

Skrót na podstawie tabeli I zawartej w: A. Hald, *Statistical Tables and Formulas*, John Wiley & Sons, Inc., New York 1952. Wykorzystano za zgodą wydawnictwa John Wiley & Sons, Inc.

Szacunkowy błąd z próby

Oto jak należy korzystać z tej tabeli: odszukaj punkt przecięcia dla wielkości próby i przybliżonego rozkładu zmiennej dwuwartościowej w tej próbie. Liczba w tym punkcie przecięcia oznacza szacunkowy błąd z próby, przy poziomie ufności 95%, wyrażony w procentach (plus lub minus).

Przykładowo: w próbie 400 respondentów 60% odpowiedziało „tak", a 40% „nie". Błąd z próby oszacowano na plus minus 4,9 pkt. proc. Przedział ufności mieści się więc między 55,1% a 64,9%. Możemy oszacować (z 95-procentową pewnością), że dla całej populacji odsetek osób, które odpowiedziałyby „tak", mieści się w granicach tego przedziału.

Rozkład zmiennej dwuwartościowej

Wielkość próby	50/50	60/40	70/30	80/20	90/10
100	10	9,8	9,2	8	6
200	7,1	6,9	6,5	5,7	4,2
300	5,8	5,7	5,3	4,6	3,5
400	5	4,9	4,6	4	3
500	4,5	4,4	4,1	3,6	2,7
600	4,1	4	3,7	3,3	2,4
700	3,8	3,7	3,5	3	2,3
800	3,5	3,5	3,2	2,8	2,1
900	3,3	3,3	3,1	2,7	2
1000	3,2	3,1	2,9	2,5	1,9
1100	3	3	2,8	2,4	1,8
1200	2,9	2,8	2,6	2,3	1,7
1300	2,8	2,7	2,5	2,2	1,7
1400	2,7	2,6	2,4	2,1	1,6
1500	2,6	2,5	2,4	2,1	1,5
1600	2,5	2,4	2,3	2	1,5
1700	2,4	2,4	2,2	1,9	1,5
1800	2,4	2,3	2,2	1,9	1,4
1900	2,3	2,2	2,1	1,8	1,4
2000	2,2	2,2	2	1,8	1,3

Załącznik

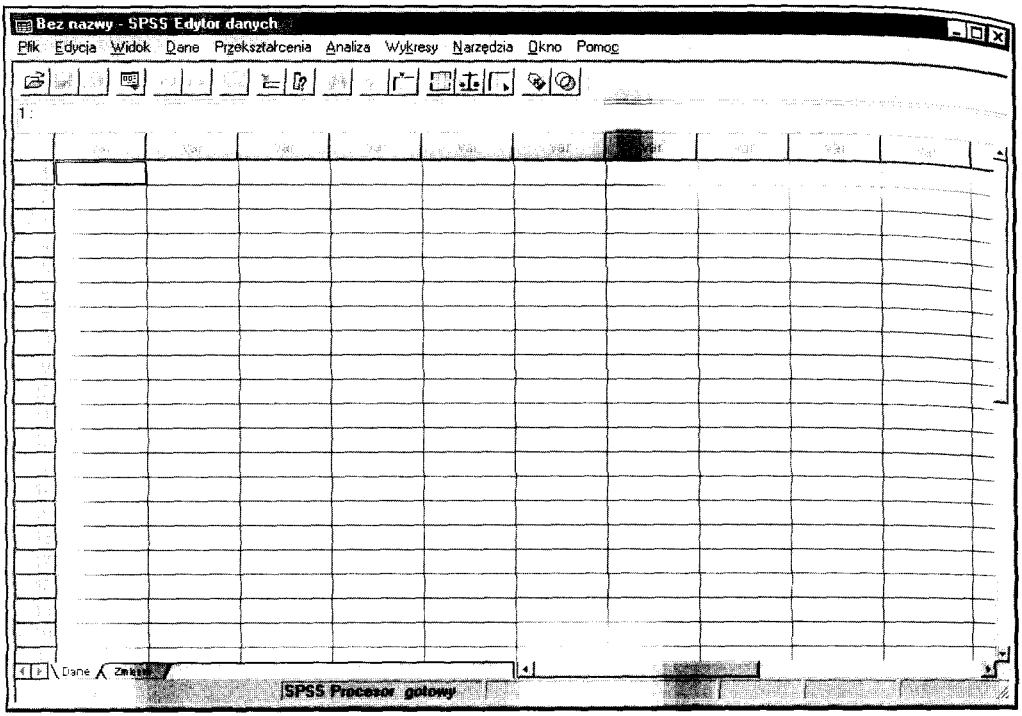
Podręcznik SPSS



www.spss.pl
www.analizadanych.pl
info@spss.pl

Zaczynamy
Otwieranie pliku z danymi
Rozkłady częstości
Tabele krzyżowe
Rekodowanie zmiennych
Zapisywanie zmian
Tabele wielozmiennowe
Testy istotności statystycznej
Korelacja i regresja
Tworzenie indeksów
Grafika
Kopiowanie wyników do dokumentu
Zamykanie programu

Rysunek I.1
Otwieranie SPSS



Niniejszy załącznik jest krótkim przewodnikiem po programie SPSS (*Statistical Package for Social Sciences*™). Od wielu lat SPSS jest najczęściej używanym programem do ilościowej analizy danych w naukach społecznych. Ewolucja programu obejmuje wiele wersji zarówno dla komputerów Macintosh, jak i dla środowiska Windows. W niniejszym załączniku będziemy wykorzystywać SPSS 11.0 w polskiej wersji językowej oraz danych z Generalnego Sondażu Społecznego (*General Social Survey*) z 1996 roku[1]. Jeśli używasz innej wersji SPSS lub innych danych, będziesz musiał dokonać kilku zmian. Niniejszy załącznik wprowadzi cię w ogólną logikę i możliwe zastosowania SPSS. Niezależnie od tego z jakiej wersji SPSS korzystasz, zaglądaj do podręcznika zawsze, gdy potrzebujesz dodatkowych wskazówek.

Ostatnio ukazało się kilka książek mających na celu zapoznanie badacza społecznego z pakietem SPSS. Jedną z nich jest *Adventures in Social Research* Earla Babbiego, Freda Halleya, Jeannie Zaino, wydana przez Pine Forge Press, Thousand Oaks, Cal. 2000[2].

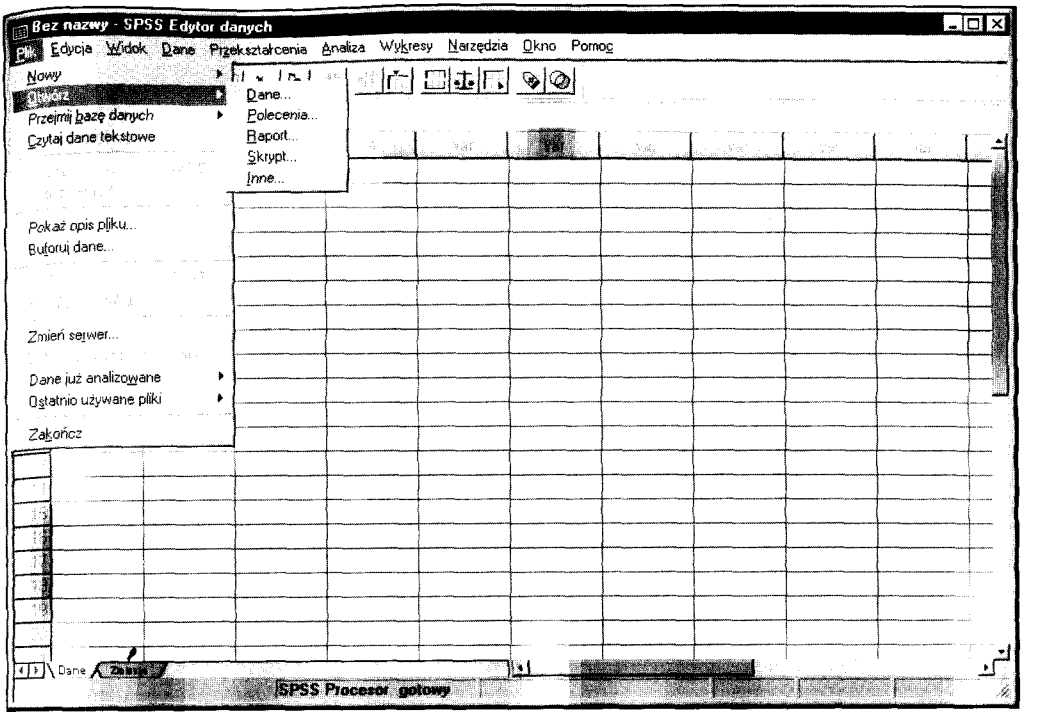
Zaczynamy

Gdy pierwszy raz oglądasz na komputerze zbiór plików SPSS, nie jest od razu oczywiste, który z nich jest programem. W wersji dla Windows jest to **SPSSWIN.EXE**. Po dwukrotnym

[1] Czytelników polskiego wydania zachęcamy do samodzielnego zapoznania się z danymi Polskiego Generalnego Sondażu Społecznego (PGSS). Informacje o tym badaniu znajdują się na stronie internetowej PGSS pod adresem http://www.iss.uw.edu.pl/osrodki/obs/pgss/pl (przyp. red. nauk.).

[2] W języku polskim dostępny jest prosty podręcznik, napisany przez Jarosława Górniaka i Janusza Wachnickiego *Pierwsze kroki w analizie danych*, wyd. SPSS Polska, Kraków 2000 (przyp. red. nauk.).

Rysunek I.2
Otwieranie menu Plik



kliknięciu na plik zobaczysz w końcu tabelę, taką jak na rysunku I.1.

Każda z kolumn reprezentuje zmienną, taką jak płeć respondenta, wiek czy postawy wobec aborcji. Każdy rząd komórek reprezentuje pojedynczego respondenta. Zatem w każdej komórce tabeli (macierzy danych) przechowywana jest jakaś informacja o pewnej osobie. Na rysunku I.1 wszystkie komórki są puste. Następnym krokiem jest wczytanie danych.

Otwieranie pliku z danymi

Spójrz na pierwszy rząd na rysunku I.1. Składa się on z kilku menu: **Plik**, **Edycja**, **Widok**, **Dane**, **Przekształcenia**, **Analiza**, **Wykresy**, **Narzędzia**, **Okno** oraz **Pomoc**. Będziemy korzystać z tych menu w całym niniejszym załączniku.

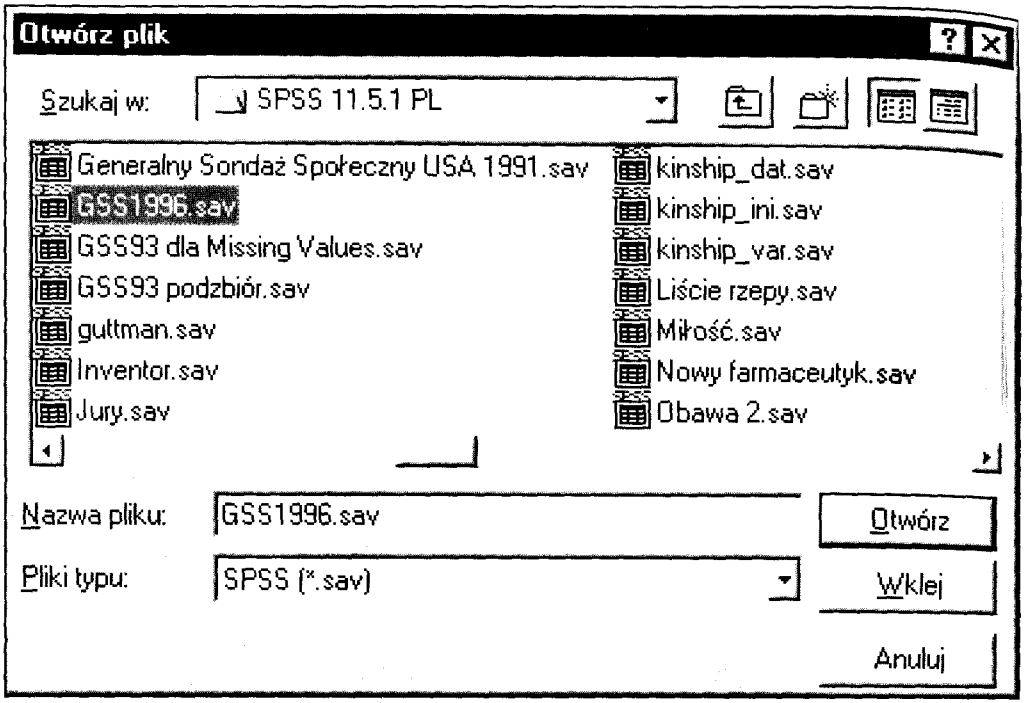
Zauważ, że pierwsza litera nazwy każdego menu jest podkreślona (na przykład **Plik**). Oznacza to, że można je wywołać poprzez jednoczesne przyciśnięcie klawisza **ALT** oraz podkreślonej litery. Zatem jeśli naciśniesz **ALT+P**, to otworzysz menu **Plik**. Można to samo uzyskać po prostu klikając w nazwę menu. Bez względu na sposób, w końcu zobaczysz ekran taki jak na rysunku I.2.

Jak widzisz, w menu **Plik** dostępnych jest kilka różnych opcji. Teraz jednak jesteśmy zainteresowani tylko otwieraniem pliku. Kliknij na **Otwórz**, a następnie z ramki, która otworzy się po prawej stronie, wybierz **Dane**, co będzie oznaczać, iż chcesz wczytać do programu plik z danymi.

Zobaczysz potem okno dialogowe, zapraszające do wybrania zbioru danych, który ma zostać otworzony. Twój ekran będzie prawdopodobnie wyglądał zupełnie inaczej niż ten z rysunku I.3, ponieważ masz inny zbiór plików do wyboru (chyba że mamy synchroniczność okoliczności, jakiej Guinness nie zna nawet w przybliżeniu).

Wybierz zbiór danych przez podwójne kliknięcie na nim albo przez pojedyncze klik-

Rysunek I.3
Pobieranie zbioru danych



nięcie oraz późniejsze kliknięcie na **Otwórz**. Wybraliśmy zestaw danych z Generalnego Sondażu Społecznego (GSS) z 1996 roku, który zawiera kilkaset zmiennych, pochodzących od 2090 respondentów. Badanie wykonało Krajowe Centrum Badania Opinii (*National Opinion Research Center*) przy Uniwersytecie Chicago. Jest ono reprezentatywne dla mieszkańców USA w wieku 18 lat i więcej. Ty oczywiście możesz pracować z innym zbiorem danych.

Gdy już wydasz programowi polecenie otwarcia zbioru danych, twoja pierwotna macierz zapełni się danymi, podobnie jak na rysunku I.4. Zauważ, że w rzędzie znajdującym się tuż powyżej macierzy pojawiły się teraz nazwy zmiennych – **abany**, **abdefect** itd. SPSS używa skróconych nazw zmiennych, nie dłuższych niż osiem znaków.

Na rysunku I.4 wybrana jest komórka w lewym górnym rogu (widać wokół niej czarną obwódkę), a w rzędzie powyżej nazw zmiennych wyświetliła się nazwa wybranej zmiennej oraz wartość, jaką przybiera ona dla wybranego przypadku. Innymi słowy, w przypadku pierwszego respondenta zmienna **abany** ma wartość 1. Co to jednak właściwie znaczy?

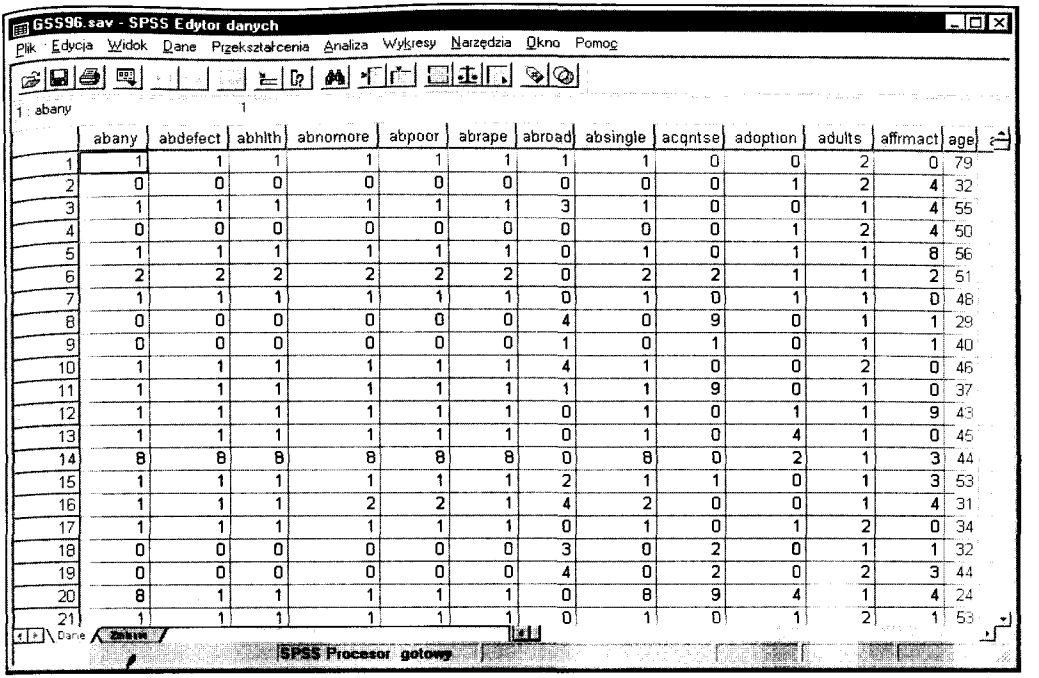
Znaczenie konkretnej nazwy zmiennej można znaleźć na kilka sposobów. Pierwszy, najłatwiejszy – zatrzymanie kursora na nazwie zmiennej na szczycie kolumny. Pojawi się wtedy dłuższa etykieta zmiennej. Można też dwukrotnie kliknąć nazwę zmiennej w nagłówku, co spowoduje przejście do zakładki **Zmienne**. Na rysunku I.5 widać podświetlony wiersz, który zawiera informacje na temat zmiennej **abany**. W innych wierszach tej zakładki zdefiniowane są pozostałe zmienne.

Zwróć uwagę na to, co jest napisane w kolumnie **Etykieta**: „ABORTION IF WOMAN WANTS FOR ANY REASON[3]". Mimo że cały czas jest to tylko skrót, można domyślić się, że reprezentuje poglądy na to, czy kobieta powinna mieć możliwość przeprowadzenia aborcji z dowolnego powodu. Pełen opis zmiennej najłat-

[3] Aborcja na życzenie kobiety (przyp. tłum.).

Rysunek I.4

Wypełniona macierz danych



GSS96.sav - SPSS Edytor danych

Plik Edycja Widok Dane Przekształcenia Analiza Wykresy Narzędzia Okno Pomoc

1 : abany 1

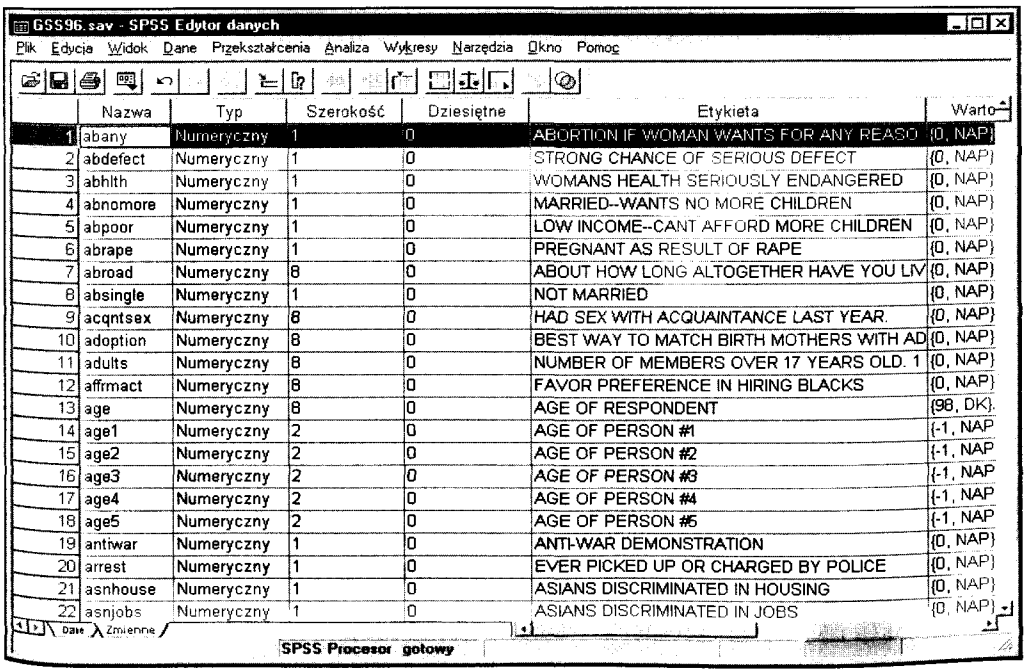
	abany	abdefect	abhlth	abnomore	abpoor	abrape	abroad	absingle	acqntsex	adoption	adults	affrmact	age
1	1	1	1	1	1	1	1	1	0	0	2	0	79
2	0	0	0	0	0	0	0	0	0	1	2	4	32
3	1	1	1	1	1	1	3	1	0	0	1	4	55
4	0	0	0	0	0	0	0	0	0	1	2	4	50
5	1	1	1	1	1	1	0	1	0	1	1	8	56
6	2	2	2	2	2	2	0	2	2	1	1	2	51
7	1	1	1	1	1	1	0	1	0	1	1	0	48
8	0	0	0	0	0	0	4	0	9	0	1	1	29
9	0	0	0	0	0	0	1	0	1	0	1	1	40
10	1	1	1	1	1	1	4	1	0	0	2	0	46
11	1	1	1	1	1	1	1	1	9	0	1	0	37
12	1	1	1	1	1	1	0	1	0	1	1	9	43
13	1	1	1	1	1	1	0	1	0	4	1	0	45
14	8	8	8	8	8	8	0	8	0	2	1	3	44
15	1	1	1	1	1	1	2	1	1	0	1	3	53
16	1	1	1	2	2	1	4	2	0	0	1	4	31
17	1	1	1	1	1	1	0	1	0	1	2	0	34
18	0	0	0	0	0	0	3	0	2	0	1	1	32
19	0	0	0	0	0	0	4	0	2	0	2	3	44
20	8	1	1	1	1	1	0	8	9	4	1	4	24
21	1	1	1	1	1	1	0	1	0	1	2	1	53

Dane Zmienne

SPSS Procesor gotowy

Rysunek I.5

Opis zmiennej **abany**



GSS96.sav - SPSS Edytor danych

Plik Edycja Widok Dane Przekształcenia Analiza Wykresy Narzędzia Okno Pomoc

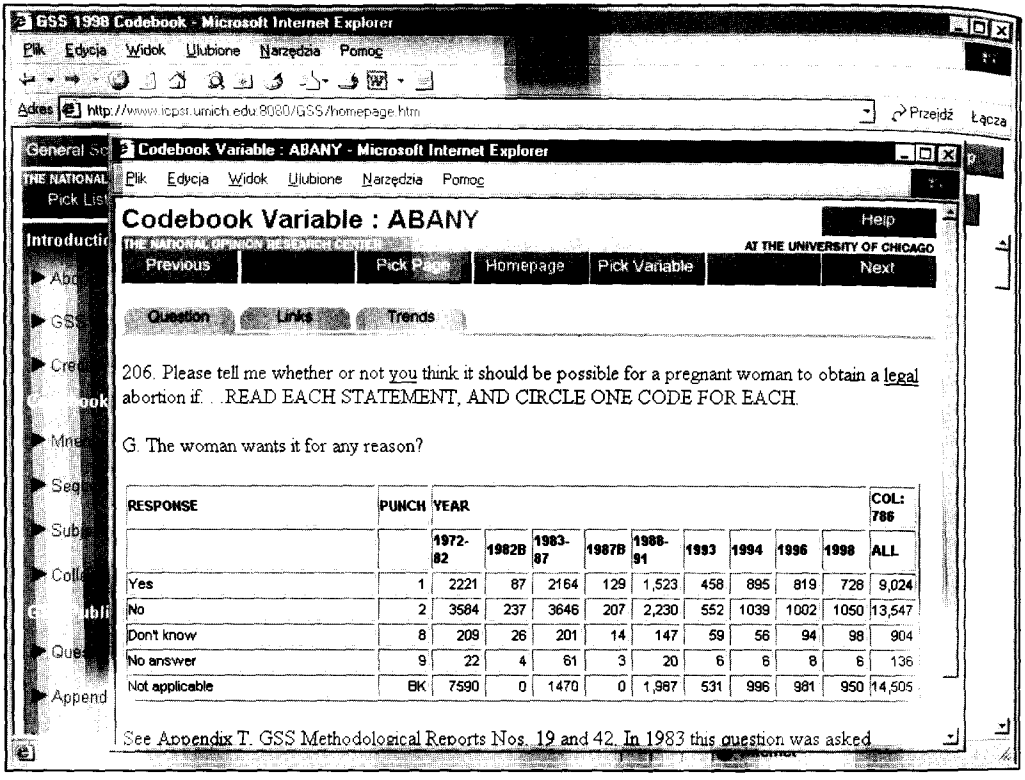
	Nazwa	Typ	Szerokość	Dziesiętne	Etykieta	Warto
1	abany	Numeryczny	1	0	ABORTION IF WOMAN WANTS FOR ANY REASO	{0, NAP}
2	abdefect	Numeryczny	1	0	STRONG CHANCE OF SERIOUS DEFECT	{0, NAP}
3	abhlth	Numeryczny	1	0	WOMANS HEALTH SERIOUSLY ENDANGERED	{0, NAP}
4	abnomore	Numeryczny	1	0	MARRIED--WANTS NO MORE CHILDREN	{0, NAP}
5	abpoor	Numeryczny	1	0	LOW INCOME--CANT AFFORD MORE CHILDREN	{0, NAP}
6	abrape	Numeryczny	1	0	PREGNANT AS RESULT OF RAPE	{0, NAP}
7	abroad	Numeryczny	8	0	ABOUT HOW LONG ALTOGETHER HAVE YOU LIV	{0, NAP}
8	absingle	Numeryczny	1	0	NOT MARRIED	{0, NAP}
9	acqntsex	Numeryczny	8	0	HAD SEX WITH ACQUAINTANCE LAST YEAR.	{0, NAP}
10	adoption	Numeryczny	8	0	BEST WAY TO MATCH BIRTH MOTHERS WITH AD	{0, NAP}
11	adults	Numeryczny	8	0	NUMBER OF MEMBERS OVER 17 YEARS OLD. 1	{0, NAP}
12	affrmact	Numeryczny	8	0	FAVOR PREFERENCE IN HIRING BLACKS	{0, NAP}
13	age	Numeryczny	8	0	AGE OF RESPONDENT	{98, DK}.
14	age1	Numeryczny	2	0	AGE OF PERSON #1	{-1, NAP
15	age2	Numeryczny	2	0	AGE OF PERSON #2	{-1, NAP
16	age3	Numeryczny	2	0	AGE OF PERSON #3	{-1, NAP
17	age4	Numeryczny	2	0	AGE OF PERSON #4	{-1, NAP
18	age5	Numeryczny	2	0	AGE OF PERSON #5	{-1, NAP
19	antiwar	Numeryczny	1	0	ANTI-WAR DEMONSTRATION	{0, NAP}
20	arrest	Numeryczny	1	0	EVER PICKED UP OR CHARGED BY POLICE	{0, NAP}
21	asnhouse	Numeryczny	1	0	ASIANS DISCRIMINATED IN HOUSING	{0, NAP}
22	asnjobs	Numeryczny	1	0	ASIANS DISCRIMINATED IN JOBS	{0, NAP}

Dane Zmienne

SPSS Procesor gotowy

Rysunek I.6

Opis zmiennej **abany** z książki kodowej GSS



RESPONSE	PUNCH	YEAR									COL: 786
		1972-82	1982B	1983-87	1987B	1988-91	1993	1994	1996	1998	ALL
Yes	1	2221	87	2164	129	1,523	458	895	819	728	9,024
No	2	3584	237	3646	207	2,230	552	1039	1002	1050	13,547
Don't know	8	209	26	201	14	147	59	56	94	98	904
No answer	9	22	4	61	3	20	6	6	8	6	136
Not applicable	BK	7590	0	1470	0	1,987	531	996	981	950	14,505

wiej znaleźć na stronie internetowej Generalnego Sondażu Społecznego. Wykaz poszczególnych kategorii kodowych (według nazw zmiennych) znajduje się pod adresem http://www.icpsr.umich.edu:8080/GSS/homepage.htm. Wejdź tam i kliknij na pierwszą literę nazwy zmiennej. Otrzymasz wtedy listę zmiennych zaczynających się na *a*, wśród których znajduje się nasze tajemnicze **abany**. Jeśli klikniesz na nazwę zmiennej, to ukaże się okienko, takie jak na rysunku I.6.

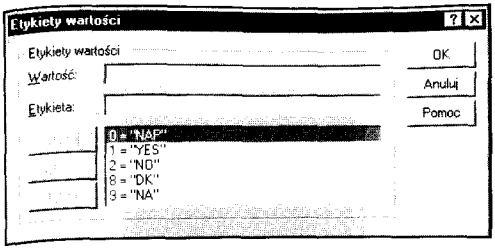
Teraz możesz dokładniej zobaczyć, co ta zmienna reprezentuje. Respondentom zadawano serię pytań dotyczących ich postaw wobec aborcji – a w szczególności warunków (na przykład gwałt, ryzyko wad wrodzonych), pod jakimi uznaliby, że kobieta powinna mieć prawo do legalnej aborcji. W tym przypadku zapytano badanych o to, czy popierają prawo kobiety do aborcji jako czysto osobistego wyboru: „z dowolnego powodu".

Oprócz dosłownego brzmienia pytania na stronie znajdują się także kategorie odpowiedzi oraz wyniki wielu sondaży zawierających w kolejnych latach to pytanie. Zauważ, że kod „1" oznacza odpowiedź „tak". Teraz wiemy, że pierwsza osoba w tym zbiorze danych uważa, że kobieta powinna mieć możliwość przeprowadzenia aborcji z dowolnego powodu.

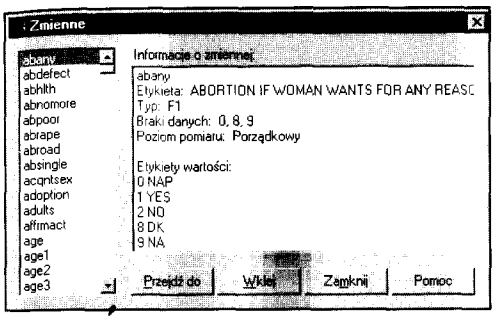
Jeśli powrócisz do macierzy danych z rysunku I.4, zauważysz, że osoba szósta nie przyznałaby tego prawa. W tym przypadku „2" oznacza „nie". Zauważ także, że wartość tej zmiennej (a także innych zmiennych) dla niektórych respondentów wynosi „0". Strona internetowa GSS nie wyjaśnia tego kodu, gdyż został on przypisany, gdy zbiór danych zapisano w formacie SPSS.

Aby dowiedzieć się, co oznacza „0", musimy nauczyć się najpierw, jak sprawdzać kody zmiennych w SPSS. Podwójne kliknięcie na nazwę zmiennej **abany** w nagłówku kolumny w zbiorze danych przenosi nas do zakładki **Zmienne**. Na-

Rysunek I.7
Etykiety kodów zmiennej **abany**



Rysunek I.8
Narzędzia/Zmienne



stępnie należy dwukrotnie kliknąć na komórkę odpowiadającą zmiennej **abany** w kolumnie **Wartości**. Jak widać z rysunku I.7, „0" oznacza „NAP", czyli „not applicable"[4]. Innymi słowy, to konkretne pytanie nie zostało zadane niektórym respondentom. Na przykład drugiemu z respondentów tego pytania nie postawiono. Poprzez zadawanie różnych zestawów pytań różnym respondentom badacze mogą zebrać dane dla kilkuset zmiennych bez doprowadzania ankietowanego do samobójstwa bądź zabójstwa.

Jest też inny sposób znalezienia informacji na temat zmiennych występujących w zbiorze danych. Wejdź do menu **Narzędzia** powyżej macierzy danych i wybierz pierwszą opcję – **Zmienne**. Rysunek I.8 pokazuje efekt takiego działania.

Przekonasz się przede wszystkim, że mamy z grubsza te same informacje, które otrzymaliśmy poprzednio. Zauważ jednak kolumnę po lewej stronie. Jest to początek listy wszystkich zmiennych występujących w tym zbiorze danych (możesz zobaczyć resztę listy, przewijając pasek z prawej strony listy). Znajdź nazwę zmiennej, którą jesteś zainteresowany, i kliknij na nią. Natychmiast wyświetli się na ekranie jej etykieta oraz etykiety jej wartości. *Uwaga:* lista z lewej strony składa się z nazw zmiennych zamiast etykiet (opisów tekstowych), ale można to łatwo zmienić. W menu **Edycja** wybierz **Opcje**. W zakładce **Ogólne** znajdź w kolumnie po prawej stronie sekcję **Listy zmiennych**. Kliknij na **Pokaż etykiety**. Aby zobaczyć efekt, będziesz musiał jeszcze raz wczytać dane, ale jest to tego warte, ponieważ korzystając z etykiet, będziesz mógł w prosty sposób odnaleźć zmienne.

Jeszcze innym wyjściem jest kliknięcie na menu **Widok** i wybranie **Etykiety wartości**. Pozwolę ci samemu sprawdzić, co się stanie, gdy to zrobisz. (Aby to wyłączyć, otwórz jeszcze raz **Widok** i kliknij na **Etykiety wartości**. Zauważ, iż pojawiający się ptaszek sygnalizuje, czy funkcja jest włączona, czy wyłączona).

Rozkłady częstości

Teraz, gdy już przekonaliśmy się, co oznaczają skrótowe nazwy zmiennych oraz poszczególne kategorie kodów liczbowych, jesteśmy gotowi do badania opinii publicznej. Pomyśl o pytaniu, które analizowaliśmy do tej pory. Jak przypuszczasz, co sądzą ludzie w Stanach Zjednoczonych na temat prawa kobiety do aborcji? Innymi słowy, jak twoim zdaniem wygląda procentowy rozkład odpowiedzi „tak" i „nie"? Aby to stwierdzić, wybierz z menu **Analiza** opcję **Opis statystyczny**, a spośród dostępnych wariantów zaznacz **Częstości** (rysunek I.9).

Po wykonaniu tej komendy na ekranie pojawi się lista zmiennych, taka jak na rysunku I.10.

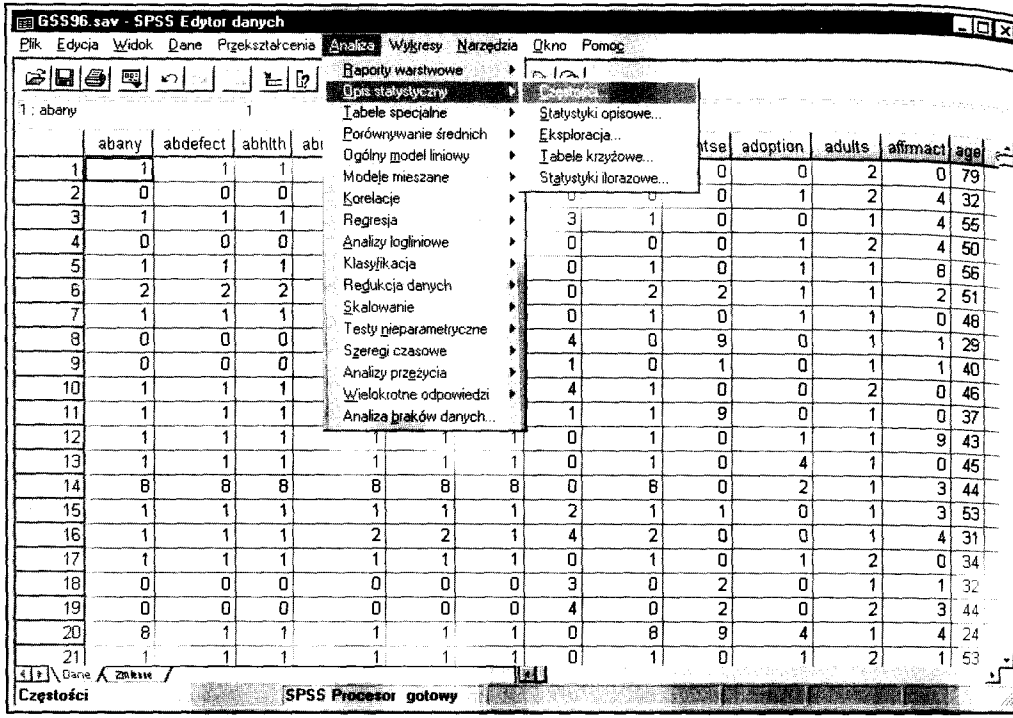
Teraz możesz wybrać zmienną poprzez dwukrotne kliknięcie na jej nazwę bądź pojedyncze kliknięcie na nazwę i kliknięcie na skierowaną w prawo trójkątną strzałkę. Każde z tych działań przeniesie zmienną z lewej do prawej kolumny. Na rysunku I.11 przedstawiono widok okna po wybraniu w ten sposób trzech zmiennych.

Następnie naciskamy przycisk **OK**. SPSS przez chwilę będzie wykonywał polecenie, zanim ostatecznie określi rozkłady zadanych

[4] Nie dotyczy (przyp. tłum.).

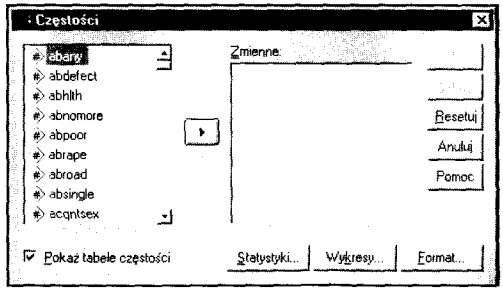
Rysunek I.9

Otrzymywanie rozkładu częstości



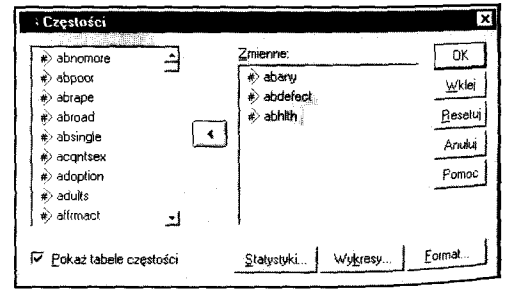
Rysunek I.10

Wybieranie zmiennej do rozkładu częstości



Rysunek I.11

Wybieranie zmiennych do rozkładu częstości



zmiennych. W wyniku tego utworzy tabele częstości, pokazane na rysunku I.12.

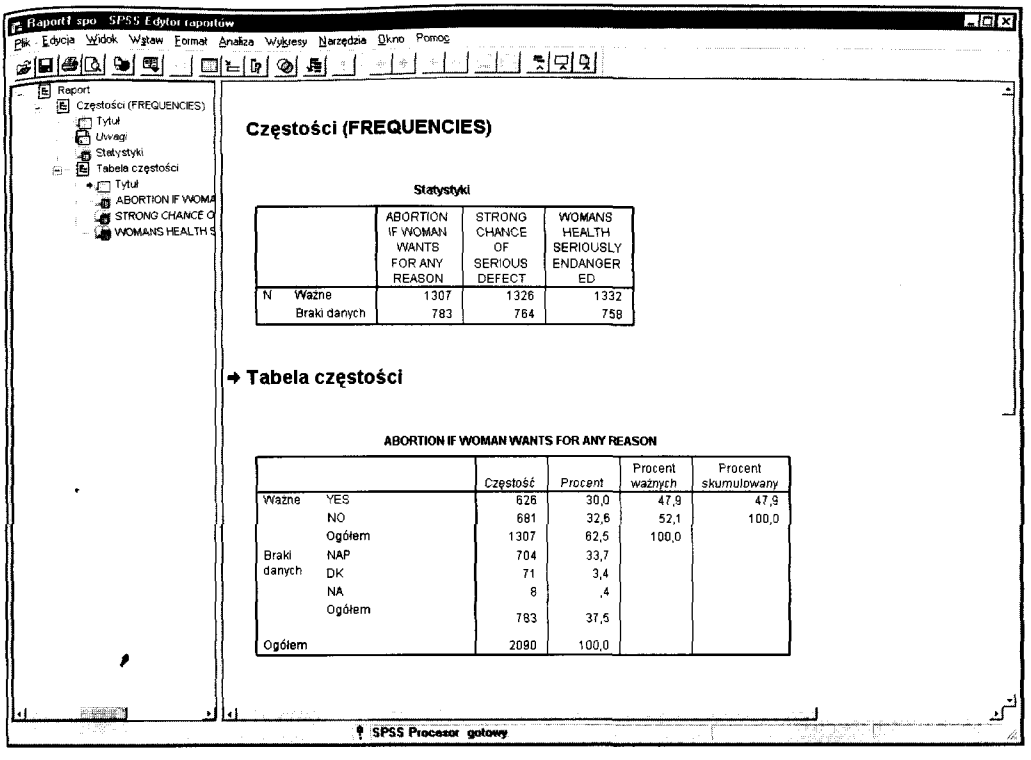
Zauważ, że SPSS otworzył teraz nowe okno. Pierwsze było oknem edytora danych, nowe jest oknem edytora raportów. W trakcie dalszej pracy z SPSS będziesz często wykonywał różne operacje, korzystając z tych dwóch okien. Program często zmienia je automatycznie.

Mniejsze okno po lewej stronie rysunku I.12 pokazuje strukturę wyników. Kliknij na dowolny element tej struktury, a w prawym oknie pokażą się żądane dane.

Po prawej stronie rysunku I.12 widzimy dwie tabele. Pierwsza z nich jest podsumowaniem trzech wybranych zmiennych. Jednak wszystko, czego możemy się z niej dowiedzieć, to ilu respondentów udzieliło odpowiedzi na pytanie, a ilu nie. Druga natomiast przedstawia rozkład odpowiedzi na pytanie, czy kobieta powinna mieć prawo do aborcji z dowolnego po-

Rysunek I.12

Tabele rozkładu częstości



wodu. Oprócz odpowiedzi „tak" i „nie" tabela podaje trzy inne możliwości:

NAP: „not applicable"[5] (pytanie nie zostało zadane),
DK: „don't know"[6],
NO: „no answer"[7] (respondentowi zadano to pytanie, ale nie udzielił odpowiedzi).

Z drugiej kolumny tabeli **Częstości** możemy się dowiedzieć, ilu respondentów wybrało poszczególne kategorie odpowiedzi. W kolumnie **Procent** znajdziemy bardziej przydatną formę tej informacji, pokazującą procentowe udziały poszczególnych kategorii. Najbardziej użyteczna kolumna to **Procent ważnych**. Mówi nam ona, iż spośród 1307 respondentów, którzy udzielili ważnych odpowiedzi, 47,9% wybrało odpowiedź „tak", a 52,1% „nie". Interpretacja tego wyniku może być taka, że opinia publiczna w tej sprawie jest prawie równo podzielona.

Przewijając okno w dół za pomocą paska z prawej strony lub posługując się strukturą wyników w lewym oknie, można sprawdzić wyniki dla innych zmiennych. Teraz jednak przejdźmy do bardziej złożonych analiz.

Tabele krzyżowe

Rozkłady częstości, które właśnie przećwiczyliśmy, to przykład analizy jednozmiennowej (analizowana jest jedna zmienna). Przejdziemy teraz do analizy dwuzmiennowej (dwie zmienne na raz).

Pozostańmy przy temacie „aborcji z dowolnego powodu". Właśnie zobaczyliśmy, iż amerykańska opinia publiczna jest prawie równo podzielona w tej sprawie. Co twoim zdaniem decyduje o takim podziale? Ludzie częs-

[5] Nie dotyczy (przyp. tłum.).
[6] Nie wiem (przyp. tłum.).
[7] Brak odpowiedzi (przyp. tłum.).

to domniemują, że kobiety będą skłonne bardziej niż mężczyźni do popierania aborcji jako prawa kobiety. Zobaczmy, jak można sprawdzić słuszność tego przypuszczenia.

Powróć do menu **Opis statystyczny**, ale wybierz teraz **Tabele krzyżowe**. Tym razem pojawi się na ekranie nieco inne okno, a mianowicie takie jak na rysunku 1.13.

Zbudujemy teraz procentową tabelę obejmującą dwie zmienne: **abany** oraz **sex**. Tabela będzie miała kolumny i rzędy. Jest wiele sposobów skonstruowania takiej tabeli, ale my w kolumnach umieścimy zmienną **sex** (czyli płeć: kobieta lub mężczyzna). Potem zaś przyjrzymy się rozkładowi zmiennej **abany** w ramach każdej z tych kategorii. Zgodnie z logiką i językiem SPSS, przyporządkowujemy zmienną **abany** rzędom, a zmienną **sex** kolumnom. Aby to zrobić, wybierz nazwy zmiennych z listy i przenieś je do właściwych okien po prawej stronie. Jest to pokazane na rysunku 1.14.

Aby znaleźć na liście nazwę jakiejś zmiennej, można przewijać listę zmiennych albo kliknąć na dowolną i następnie wpisać nazwę potrzebnej zmiennej. Trzeba trochę popróbować, żeby odkryć, w jakim tempie należy pisać, aby to działało.

Do tej pory kazaliśmy programowi SPSS zaprojektować taką oto tabelę:

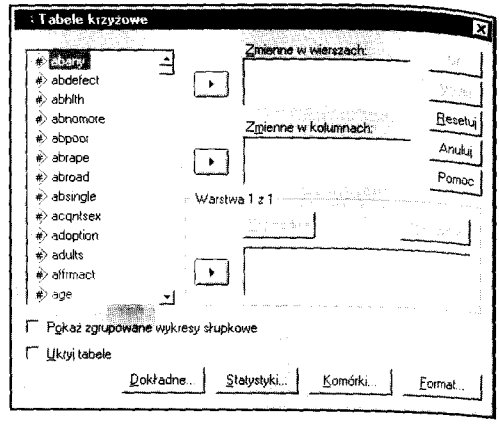
	Mężczyźni	Kobiety
Popierają		
Nie popierają		

Aby zakończyć to polecenie, trzeba wskazać, czy procenty mają być sumowane w rzędach, czy w kolumnach. W naszym przypadku będziemy chcieli uzyskać dane dotyczące tego, jaki procent mężczyzn popiera aborcję oraz jaki procent jej nie popiera, przy czym obie wielkości powinny się sumować do 100%. Podobne dane zechcemy uzyskać także dla kobiet. Innymi słowy, ustawiamy w SPSS procentowanie w kolumnach. Opcja **Tabele krzyżowe** wyposażona jest w narzędzie do wskazania naszych preferencji w tym zakresie. Kliknij przycisk **Komórki** w oknie dialogowym (rysunek 1.15).

W otwartym oknie dialogowym będzie już zaznaczona opcja **Liczebności Obserwowane**.

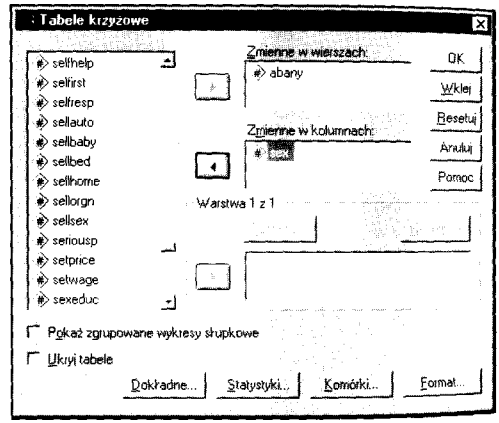
Rysunek 1.13

Okno dialogowe tabel krzyżowych



Rysunek 1.14

Wybieranie zmiennych do tabel krzyżowych

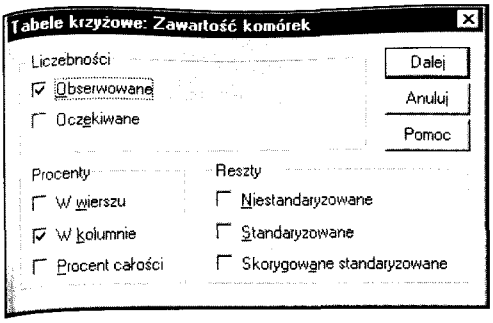


Nie należy jej zmieniać. W sekcji **Procenty** kliknij na kratkę **W kolumnie**. Jest to wskazówka dla SPSS, że ma procentować w kolumnach. Kliknij na **Dalej**, aby zakończyć pracę z tym oknem dialogowym. Potem zaś kliknij na **OK**, aby uruchomić polecenie utworzenia tabeli krzyżowej. Gdy SPSS skończy tabelę, zostaniemy przeniesieni do okna edytora raportów jak na rysunku 1.16.

Zobaczmy, co nam mówi tabela. Chcieliśmy sprawdzić, czy kobiety i mężczyźni różnią się w swych opiniach, dotyczących prawa kobiety do przeprowadzenia aborcji na życzenie. Tabela sugeruje, że nie ma między nimi znaczącej różnicy. Tylko trochę mniej mężczyzn (46,5%) niż kobiet (49%) jest zdania, że kobieta powinna mieć prawo do aborcji z dowolnego powodu.

Rysunek I.15

Określanie sposobu procentowania



Sprawdźmy inną zmienną, która może mieć wpływ na postawy wobec aborcji, *poglądy polityczne*. W GSS poglądy polityczne reprezentowane są przez zmienną **polviews**. Reprezentuje ona standardowe pytanie zadawane respondentom. Są oni mianowicie proszeni o określenie swoich poglądów politycznych gdzieś pomiędzy parą określeń „skrajnie liberalne" i „skrajnie konserwatywne". Rysunek I.17 pokazuje wpływ tej zmiennej na postawy wobec aborcji.

Ze względu na dużą liczbę kategorii poglądów politycznych będziesz prawdopodobnie musiał użyć paska na dole ekranu do poruszania się wzdłuż tabeli. Zauważ, że na rysunku I.17 przesunęliśmy obraz jak najbardziej w prawą stronę.

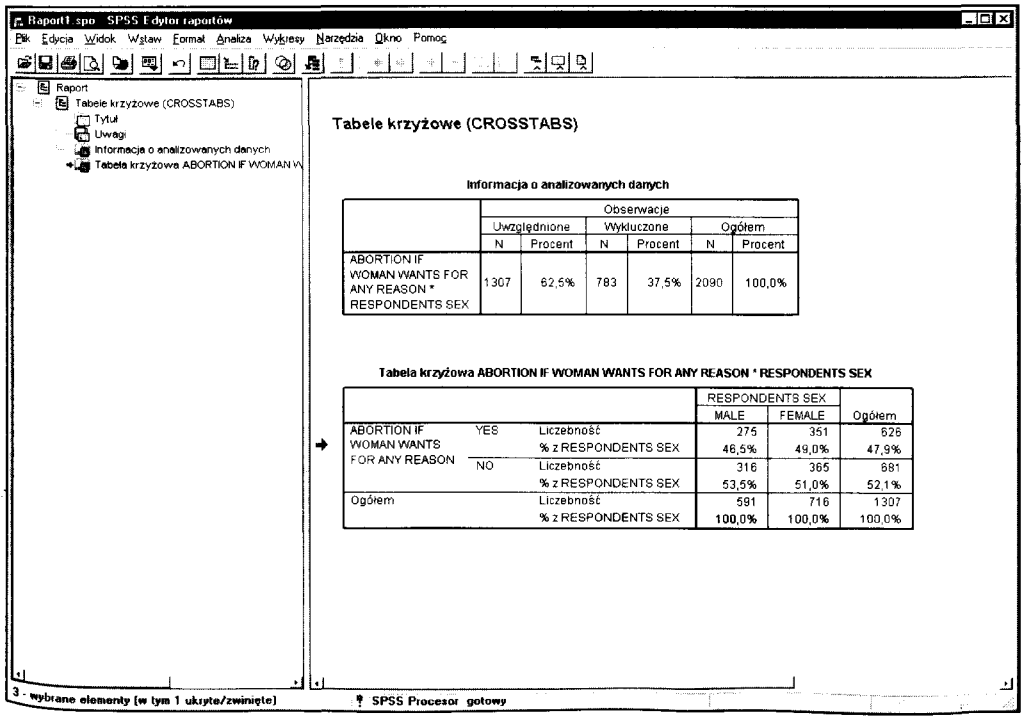
Wpływ poglądów politycznych na postawy wobec aborcji jest zupełnie jasny. Generalnie rzecz biorąc, liberałowie częściej niż konserwatyści popierają prawo do aborcji. Jedynym wyjątkiem od tego wzorca jest to, iż ludzie o poglądach „skrajnie liberalnych" rzadziej popierają prawo do aborcji niż ci o poglądach „liberalnych". Różnica wydaje się znaczna, a jej przyczyną mogą być odmienne interpretacje tych terminów politycznych.

Rekodowanie zmiennych

Często użyteczne jest rekodowanie zmiennych, mających liczne kategorie, tak aby łatwiej pod-

Rysunek I.16

Tabela krzyżowa dla zmiennych **abany** i **sex**



Rysunek I.17

Tabela krzyżowa dla zmiennych **abany** i **polviews**

Raport1.spo - SPSS Edytor raportów

Plik Edycja Widok Wstaw Format Analiza Wykresy Narzędzia Okno Pomoc

Tabela krzyżowa ABORTION IF WOMAN WANTS FOR ANY REASON * THINK OF SELF AS LIBERAL OR CONSERVATIVE

		THINK OF SELF AS LIBERAL OR CONSERVATIVE							
		EXTREMLY LIBERAL	LIBERAL	SLIGHTLY LIBERAL	MODERATE	SLIGHTLY CONSERVATIVE	CONSERVATIVE	EXTREMLY CONSERVATIVE	Ogółem
YES	Liczebność	14	108	94	217	85	71	9	598
	% z THINK OF SELF AS LIBERAL OR CONSERVATIVE	60,9%	72,5%	56,3%	47,7%	42,7%	35,1%	20,9%	48,3%
NO	Liczebność	9	41	73	238	114	131	34	640
	% z THINK OF SELF AS LIBERAL OR CONSERVATIVE	39,1%	27,5%	43,7%	52,3%	57,3%	64,9%	79,1%	51,7%
	Liczebność	23	149	167	455	199	202	43	1238
	% z THINK OF SELF AS LIBERAL OR CONSERVATIVE	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

1 - wybrane elementy (w tym 0 ukryte/usunięte) SPSS Processor gotowy H: 224 , W: 928 pkt

dawały się one analizie. W naszym przypadku moglibyśmy spróbować połączyć kategorie zmiennej **polviews** w trzy kategorie poglądów: „liberalne", „umiarkowane" i „konserwatywne".

Możemy ręcznie połączyć kategorie, korzystając z danych zawartych w tabeli na rysunku I.17. Z łatwością policzymy na przykład, że 339 respondentów uważa się za liberałów (23 + 149 + 167). Wśród nich 216 popiera prawo do aborcji na życzenie kobiety (14 + 108 + 94). Jeśli podzielimy te dwie liczby, okaże się, że 64% liberałów popiera prawo do aborcji. Z podobnych wyliczeń dotyczących konserwatystów wynika, że 165 z 444, czyli 37% popiera prawo do aborcji. 48-procentowe poparcie wśród respondentów o umiarkowanych poglądach sytuuje ich niemal pośrodku między liberałami a konserwatystami.

Takie łączenie kategorii ułatwia korzystanie ze zmiennej w dalszych analizach. Trzeba jednak sprawić, by SPSS stworzył całkiem nową zrekodowaną zmienną, tak aby za każdym razem nie trzeba było ręcznie wykonywać obliczeń. Aby to zrobić, musimy najpierw wrócić do okna edytora danych. Jeśli jesteś teraz w oknie edytora raportów, możesz po prostu kliknąć ikonę okna edytora danych w pasku zadań na dole ekranu. Gdy już powrócisz do okna edytora danych, kliknij na menu **Przekształcenia** i przesuń kursor na opcję **Rekoduj**. Gdy to zrobisz, zobaczysz na ekranie dwie inne opcje do wyboru, tak jak na rysunku I.18.

SPSS oferuje dwa tryby rekodowania. Może zmodyfikować dane i zapisać je pod tą samą nazwą zmiennej – **Na te same zmienne** lub może stworzyć nową zmienną dla zmodyfikowanych danych – **Na inne zmienne**. Wybierz **Na inne zmienne**, gdyż wybór pierwszego trybu spowoduje utratę oryginalnych danych.

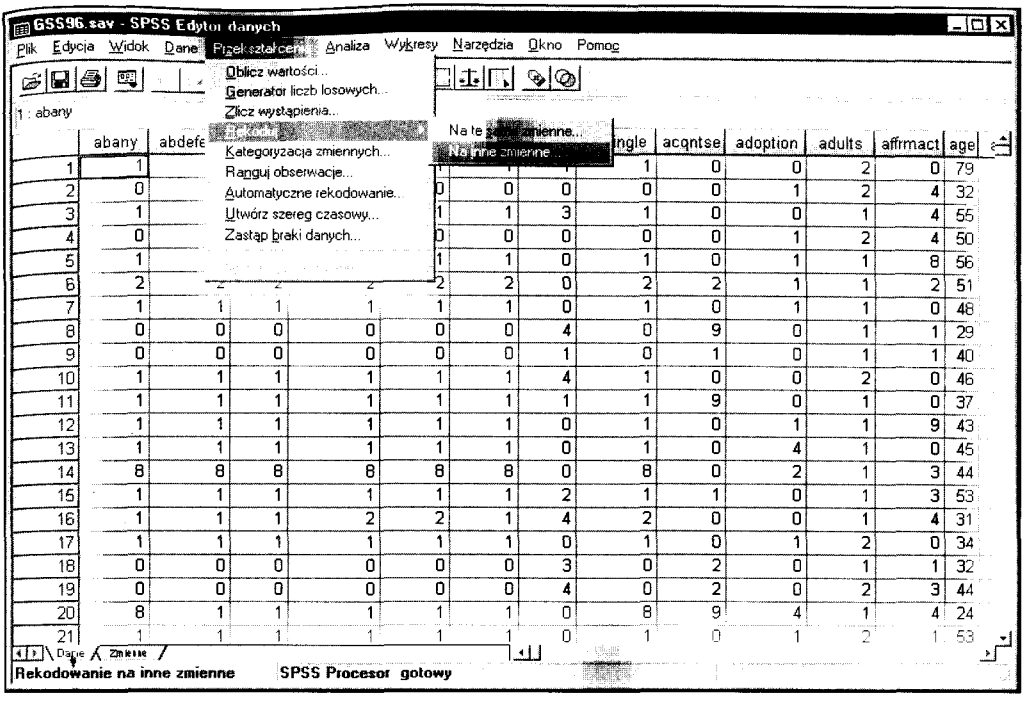
Następnie pokaże się duże okno dialogowe, takie jak na rysunku I.19.

Początkowo ramka z prawej strony będzie pusta. Aby uzyskać wygląd, taki jak na rysunku I.19, należy:

1. Wybrać **polviews** z listy zmiennych i przenieść je do środkowej ramki poprzez dwu-

Rysunek I.18

Polecenie rekodowania



krotne kliknięcie lub za pomocą trójkątnej strzałki.

2. Wpisać z klawiatury **poviewr** w wolnym miejscu w oknie **Zmienna wynikowa** → **Nazwa** i kliknąć na przycisk **Zmień**.
3. Wpisać opisową etykietę w celu określenia, co reprezentuje zmienna **polviewr**.

Aby przejść dalej, należy kliknąć przycisk **Wartości źródłowe i wynikowe**. Przeniesie to nas do okna dialogowego, jak na rysunku I.20.

Musimy teraz wskazać programowi, w jaki sposób ma stworzyć zmienną **polviewr** ze zmiennej **polviews**. W tym celu wskazujemy, jakie wartości zmiennej **polviews** będą reprezentowane przez określone wartości zmiennej **polviewr**. Zacznijmy od stworzenia kategorii „liberalne", do której wejdą wszyscy respondenci, dla których wartość zmiennej **polviews** wynosiła „1", „2" lub „3". Przypiszemy tej nowej kategorii wartość „1".

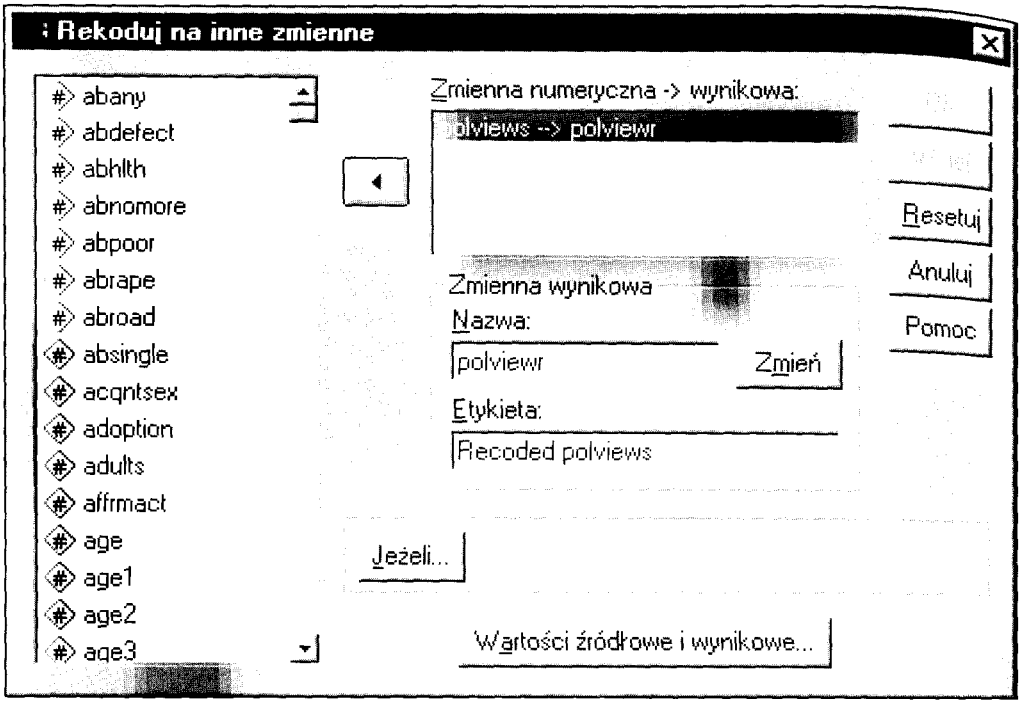
Na rysunku I.21 wybraliśmy opcję **Zakres od**, wskazując, że respondentom, dla których zmienna **polviews** przybiera wartość z zakresu od „1" do „3", przypisano wartość „1" zmiennej **polviewr**. Upewnij się, czy widzisz w oknie dialogowym miejsca, gdzie wprowadzono te instrukcje.

Gdy klikniesz na przycisk **Dodaj**, instrukcje dotyczące transformacji zmiennej są przenoszone do pola po prawej stronie okna dialogowego, tak jak to widać na rysunku I.22.

Użyjemy teraz innej opcji, aby stworzyć nową kategorię „umiarkowane". Jak pamiętasz, byli to respondenci, dla których wartość zmiennej **polviews** wynosiła „4". Przypiszemy im wartość nowej zmiennej **polviewr** równą „2". Zrobimy to, wpisując źródłową i wynikową wartość zmiennej w polach **Wartość źródłowa** oraz **Wartość wynikowa**. Ponownie klikając na **Dodaj**, wprowadzamy w tym polu nową instrukcję. A teraz poświęć chwilę na samodzielne stworzenie nowej kategorii „konserwatywne" w zmiennej **polviewr** z kategorii „5", „6" oraz „7" zmiennej **polviews**. Gdy już się z tym uporasz, powinieneś otrzymać okno dialogowe, takie jak na rysunku I.23. Pozostaje teraz

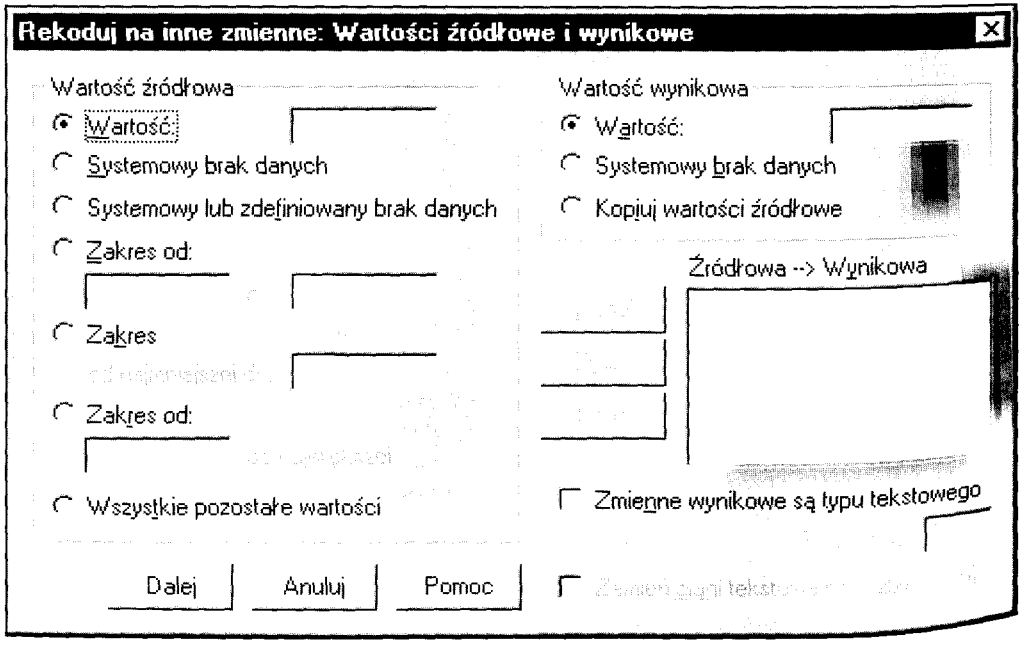
Rysunek I.19

Okno dialogowe **Rekoduj**



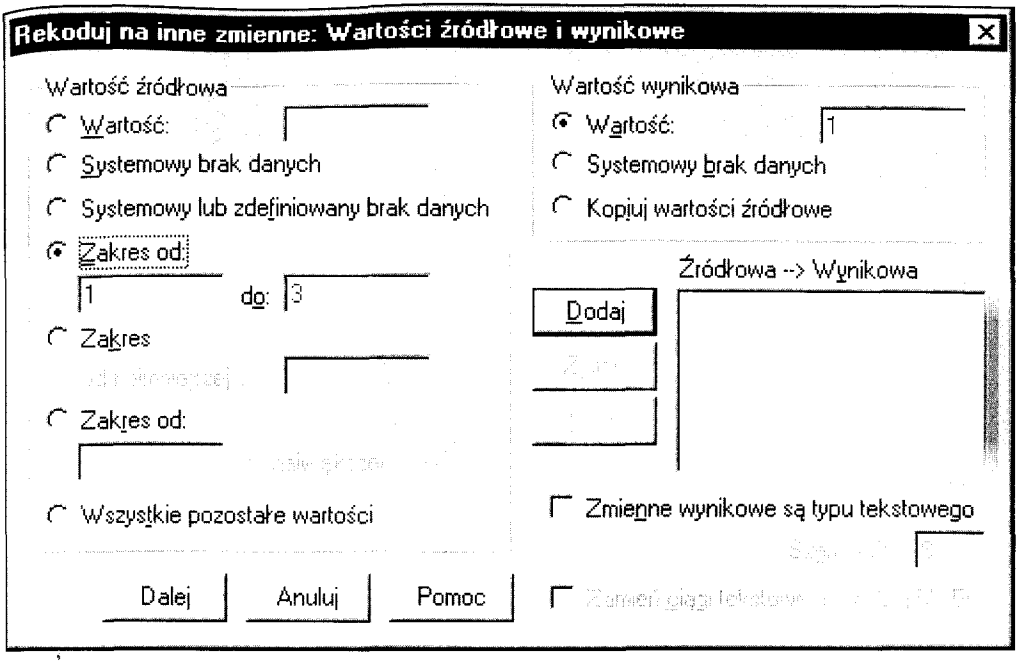
Rysunek I.20

Określenie sposobu rekodowania kategorii



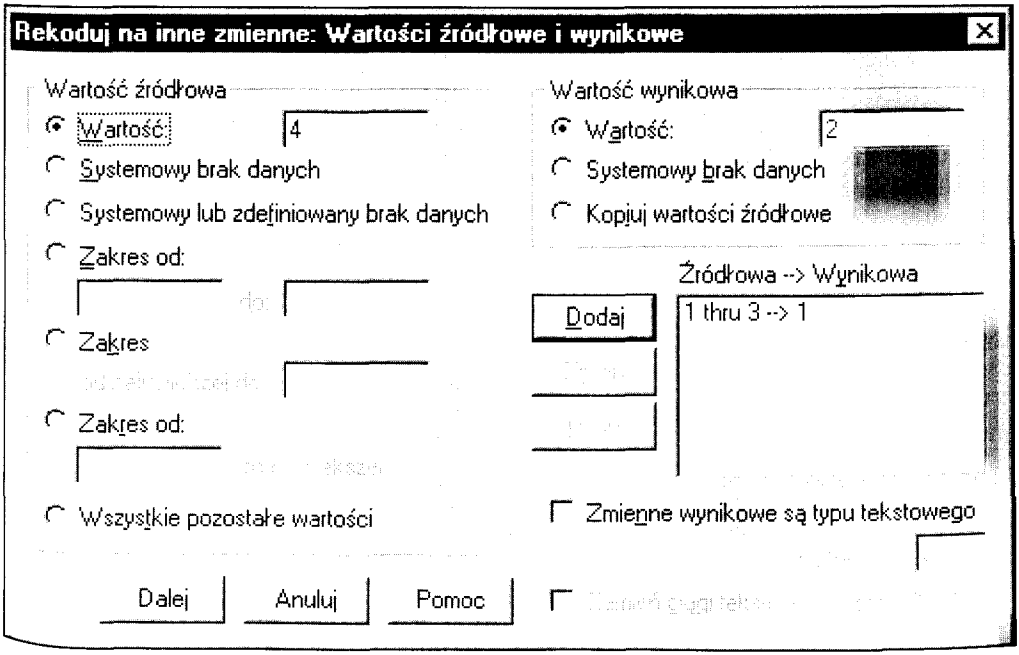
Rysunek I.21

Tworzenie pojedynczej kategorii „liberalne"

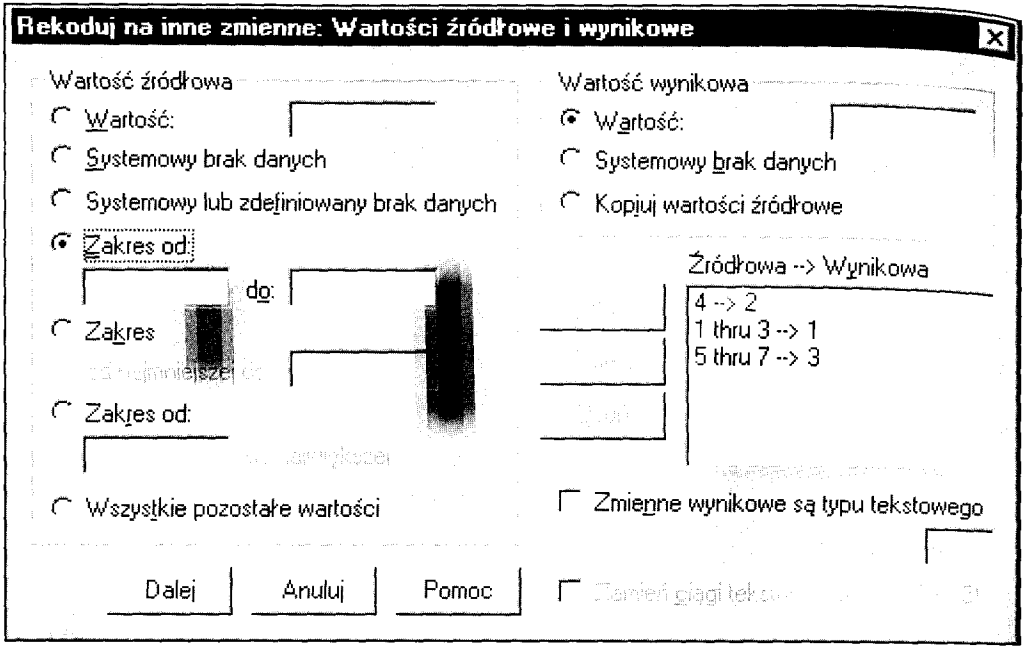


Rysunek I.22

Zmienianie kodu liczbowego kategorii „umiarkowane"



Rysunek I.23
Zakończone instrukcje rekodowania



tylko kliknięcie na **Dalej**, co cofnie cię do poprzedniego okna dialogowego. Naciśnij teraz **OK**.

Doszlifujmy naszą nową zmienną. Powróć najpierw do okna edytora danych. Następnie przewiń listę zmiennych do końca w prawo za pomocą znajdującego się na dole paska. Odszukaj **polviewr** i kliknij dwukrotnie na nazwę zmiennej w nagłówku kolumny. Program przeniesie nas do zakładki **Zmienne** w edytorze danych. Kliknij teraz dwukrotnie na komórkę odpowiadającą naszej nowej zmiennej (która najprawdopodobniej została dopisana na końcu zbioru) w kolumnie **Wartości**. Otworzy się wówczas okno **Etykiety wartości**, w którym możesz nazwać poszczególne kategorie.

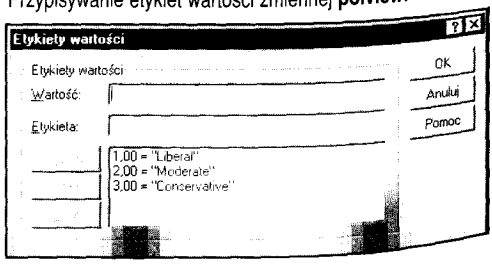
1. Wpisz „1" w polu **Wartość**.
2. Wpisz „liberalne" (*Liberal*) w polu **Etykieta**.
3. Kliknij na przycisk **Dodaj**.

Powtórz te czynności, przypisując wartości „2" określenie „umiarkowane" (*Moderate*), wartości „3" zaś określenie „konserwatywne" (*Conservative*). Pamiętaj, aby za każdym razem nacisnąć przycisk **Dodaj**. Gdy skończysz, okno dialogowe powinno wyglądać tak jak na rysunku I.24. Naciśnij **OK**.

Gdy wybierzesz teraz **Analiza/Opis statystyczny/Częstości** i przewiniesz listę w dół, znajdziesz na niej nową pozycję: **polviewr**. Wybierz ją, aby zobaczyć rozkład częstości naszych nowych kategorii (rysunek I.25).

Skoro przeszliśmy przez to wszystko po to, aby nasza analiza była prostsza, zobaczmy, czy się to opłaciło. Użyjmy zmiennej **polviewr**, aby ponownie zbadać związek między poglądami politycznymi a postawami wobec aborcji. Użyj menu **Analiza/Opis statystyczny/Tabele krzyżowe**, aby stworzyć tabelę krzyżową dla zmiennych **abany**

Rysunek I.24
Przypisywanie etykiet wartości zmiennej **polviewr**



Rysunek I.25

Rozkład częstości zmiennej **polviewr**

Political Views (Recoded)

		Częstość	Procent	Procent ważnych	Procent skumulowany
Ważne	Liberal	535	25,6	27,0	27,0
	Moderate	745	35,6	37,6	64,6
	Conservative	702	94,8	35,4	100,0
	Ogółem	1982	,9	100,0	
Braki danych	Systemowe braki danych	108	5,2		
Ogółem		2090	100,0		

Rysunek I.26

Tabela krzyżowa dla zmiennych **abany** i **polviewr**

Tabela krzyżowa ABORTION IF WOMAN WANTS FOR ANY REASON * Political Views (Recoded)

			Political Views (Recoded)			
			Liberal	Moderate	Conservative	Ogółem
ABORTION IF WOMAN WANTS FOR ANY REASON	YES	Liczebność	216	217	165	598
		% z Political Views (Recoded)	63,7%	47,7%	37,2%	48,3%
	NO	Liczebność	123	238	279	640
		% z Political Views (Recoded)	36,3%	52,3%	62,8%	51,7%
Ogółem		Liczebność	339	455	444	1238
		% z Political Views (Recoded)	100,0%	100,0%	100,0%	100,0%

oraz **polviewr**. Powinieneś otrzymać wynik podobny do przedstawionego na rysunku I.26.

Zauważ, o ile łatwiej, w porównaniu z rysunkiem I.17, czyta się tę tabelę. Widzimy, że 64% liberałów, 48% umiarkowanych oraz 37% konserwatystów popiera prawo do aborcji na życzenie kobiety. (Dobrze jest zaokrąglać miejsca dziesiętne w takich procentowaniach, gdyż bazują one na próbach, które przede wszystkim dają jedynie oszacowanie odpowiednich wartości dla populacji).

Zapisywanie zmian

Gdy modyfikujesz zbiór danych, na przykład tworząc nowe, zrekodowane zmienne, zechcesz prawdopodobnie zapisać te zmiany, aby móc w przyszłości z nich korzystać.

Trzeba zauważyć, że wszystkie wprowadzone zmiany będą działały tak długo, jak trwa sesja robocza programu SPSS, ale gdy zechcesz wyjść z programu (zobacz opis w ostatnim fragmencie tego załącznika), wszystkie zmiany zostaną utracone. Mądrze jest zatem zapisywać zmiany od razu, gdy nabierzesz przekonania, że chcesz to uczynić.

Zapisywanie zmienionego pliku na dysk jest dość proste. Najpierw wybierz okno edytora danych (nie okno edytora raportów). Potem wejdź do menu **Plik** i wybierz **Zapisz**. Możesz też zrobić to w inny sposób – naciskając **CTRL+S**. Od tej pory plik ten po otwarciu będzie zawierał wprowadzone zmiany.

Zauważ, że gdy zapisujesz plik w ten sposób, zastępuje on wcześniejszą wersję. Jeśli więc skasowałbyś dane lub przekształciłbyś zmienne używając ich oryginalnych nazw, na zawsze pozbawiłbyś się dostępu do oryginalnych danych.

Jeśli chcesz zachować zarówno oryginalne dane, jak i plik ze zmienionymi danymi, z menu **Plik** wybierz **Zapisz jako**. W tym przypadku SPSS zapyta, jaką nazwę ma mieć plik ze zmienionymi danymi. Użyj innej nazwy niż nazwa oryginalnego zbioru danych. Zwróć też uwagę, w jakim miejscu na dysku zostały zapisane dane, byś mógł je później znaleźć.

Tabele wielozmiennowe

Dwuzmiennowe tabele są zazwyczaj dopiero początkiem ilościowej analizy danych. Na przykład możesz chcieć zbadać, czy zaobserwowana zależność między poglądami politycznymi a postawami wobec aborcji jest taka sama w przypadku mężczyzn i kobiet. Dzięki SPSS zaspokojenie twojej ciekawości jest dość proste.

Powróć do **Analiza/Opis statystyczny/Tabele krzyżowe** i wybierz trzecią zmienną, tak jak na rysunku I.27.

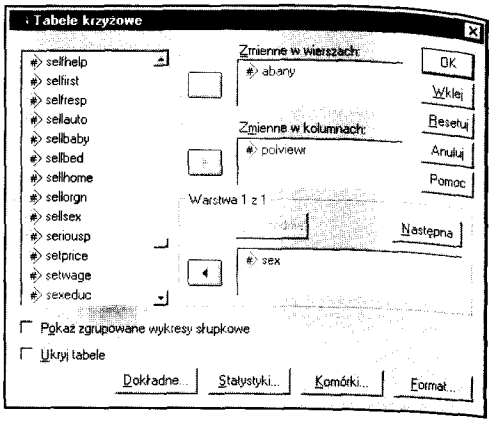
Zauważ, że po prostu przesunęliśmy trzecią zmienną, **sex**, do dolnego pola w oknie dialogowym. Naciśnij **OK**, aby zobaczyć wynik, który powinien być taki jak na rysunku I.28.

Właściwie ta nowa tabela jest podzieloną na dwie części tabelą z rysunku I.26. Górna połowa ukazuje zależność między **polviewr** a **abany** dla mężczyzn, a dolna tę samą zależność dla kobiet. Widzimy od razu, że pierwotna zależność jest powtórzona dla każdej z kategorii płci.

Pierwsza od prawej kolumna ukazuje ogólną zależność między płcią a poparciem dla aborcji. Ogólnie (czyli zapominając o poglądach politycznych) różnice w kwestii aborcji między kobietami a mężczyznami są nikłe, gdyż 47% mężczyzn i 50% kobiet popiera prawo do aborcji na życzenie kobiety. (Zauważ, że zaokrągliłem podane w tabeli 46,8% oraz 49,6%).

Rysunek I.27

Polecenie utworzenia tabeli trójzmiennowej



Rysunek I.28

Tabela krzyżowa dla zmiennych **abany** i **polviewr** w zależności od wartości zmiennej **sex**

Tabela krzyżowa ABORTION IF WOMAN WANTS FOR ANY REASON * Political Views (Recoded) * RESPONDENTS SEX

RESPONDENTS SEX				Political Views (Recoded)			Ogółem
				Liberal	Moderate	Conservative	
MALE	ABORTION IF WOMAN WANTS FOR ANY REASON	YES	Liczebność	87	87	90	264
			% z Political Views (Recoded)	63,5%	44,2%	39,1%	46,8%
		NO	Liczebność	50	110	140	300
			% z Political Views (Recoded)	36,5%	55,8%	60,9%	53,2%
	Ogółem		Liczebność	137	197	230	564
			% z Political Views (Recoded)	100,0%	100,0%	100,0%	100,0%
FEMALE	ABORTION IF WOMAN WANTS FOR ANY REASON	YES	Liczebność	129	130	75	334
			% z Political Views (Recoded)	63,9%	50,4%	35,0%	49,6%
		NO	Liczebność	73	128	139	340
			% z Political Views (Recoded)	36,1%	49,6%	65,0%	50,4%
	Ogółem		Liczebność	202	258	214	674
			% z Political Views (Recoded)	100,0%	100,0%	100,0%	100,0%

Porównując wyniki dla kobiet i mężczyzn w innych kolumnach, należy wnioskować, że płeć, bez względu na poglądy polityczne, ma mały wpływ na postawy wobec aborcji. Wśród liberałów kobiety i mężczyźni mają praktycznie takie same postawy. Wśród umiarkowanych kobiety nieco częściej niż mężczyźni popierają prawo do aborcji, a wśród konserwatystów mężczyźni są nieco bardziej przychylni aborcji. Jednak żadna z różnic nie jest zbyt duża.

SPSS pozwala wyjść dalej poza trójzmienne tabele, mimo że stają się one wówczas coraz trudniejsze do czytania i analizy. Aby trochę z tym poeksperymentować, kliknij na przycisk **Następna** w oknie dialogowym **Tabele krzyżowe** i dodaj nowe warstwy zmiennych do tabeli.

Testy istotności statystycznej

W poprzednim przykładzie mimochodem zauważyłem, że różnice procentowe nie są zbyt duże. Była to subiektywna ocena istotności różnic.

Jak wiesz, testy istotności statystycznej określają prawdopodobieństwo tego, że zależności zaobserwowane w próbie są jedynie wynikiem błędu z próby, nie zaś odbiciem rzeczywistych różnic w populacji, z której dobrano próbę. Zobaczmy, jak za pomocą SPSS korzystać z takich testów.

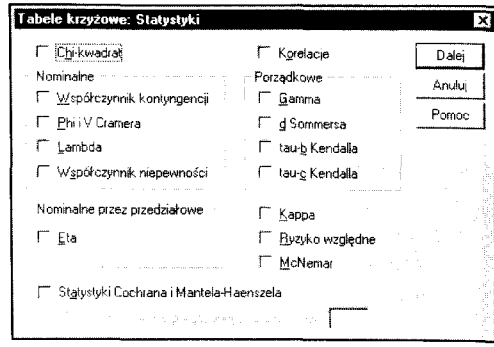
Powróć do okna dialogowego **Tabele krzyżowe** poprzez **Analiza/Opis statystyczny**. Kliknij na znajdujący się na dole przycisk **Statystyki**. Otrzymasz obraz, taki jak na rysunku I.29.

Jak widzisz, SPSS daje możliwość wyliczenia wielu statystyk opisowych. Z trzema z nich zetknąłeś się w tej książce – chi-kwadrat, lambda oraz gamma. Przypominasz sobie może, że test chi-kwadrat jest odpowiedni dla zmiennych nominalnych, takich jak **abany** oraz **sex**. Użyjmy zatem tych zmiennych i sprawdźmy, jak można wykorzystać SPSS do testu chi-kwadrat.

Zaznacz okienko **Chi-kwadrat**. Kliknij potem na **Dalej** i wpisz **abany** oraz **sex** w odpowiednich miejscach w oknie dialogowym **Tabele krzyżowe**. Oprócz normalnej tabeli z procentami SPSS stworzy dodatkową tabelę, tak jak na rysunku I.30.

Rysunek I.29

Wybór statystyk w oknie dialogowym tabel krzyżowych



Jeśli przeszedłeś kurs statystyki, nieobcych ci będzie wiele testów podanych w tej tabeli. Dla naszych celów skoncentrujmy się na pierwszym rzędzie tabeli – „Chi-kwadrat Pearsona". Trzecia kolumna określa prawdopodobieństwo tego, że sam tylko błąd z próby mógłby być przyczyną zależności o obserwowanym natężeniu, gdyby mężczyźni i kobiety w całej populacji byli dokładnie tacy sami pod względem postaw wobec aborcji. Konkretnie, wynik testu mówi nam, że prawdopodobieństwo to wynosi 0,37, czyli więcej niż jeden do trzech. Przez to test chi-kwadrat potwierdza to, co stwierdziliśmy subiektywnie: kobiety i mężczyźni nie różnią się bardzo, jeśli w ogóle, w kwestii poglądów na prawo kobiety do aborcji.

Zależność między **abany** a **polviewr** była znacznie większa. Przekonajmy się, jak SPSS ocenia tę zależność. Powtórz opisaną wyżej procedurę, zastępując zmienną **sex** przez zmienną **polviewr**. Zauważ, że nie musisz ponownie wybierać **Chi-kwadrat**, musisz jedynie zmienić zawartość pola **Kolumny**. SPSS zachowuje te ustawienia aż do wyjścia z programu. Zaczynając nową sesję SPSS, będziesz musiał od nowa określić tego rodzaju ustawienia. Możesz też oczywiście zmienić je zawsze, gdy nie są ci już potrzebne.

Rysunek I.31 przedstawia ocenę zależności między **abany** a **polviewr** za pomocą testu chi-kwadrat. Zauważ, że obliczony poziom istotności wynosi w tym przypadku ,000. SPSS wyświetla tylko trzy pierwsze miejsca po przecinku. Zatem prawdopodobieństwo, że zaobserwowana przez nas zależność jest po prostu

Rysunek I.30

Test chi-kwadrat dla zmiennych **abany** i **sex**

Testy Chi-kwadrat

	Wartość	df	Istotność asymptotyczna (dwustronna)	Istotność dokładna (dwustronna)	Istotność dokładna (jednostronna)
Chi-kwadrat Pearsona	,805[b]	1	,370		
Poprawka na ciągłość[a]	,708	1	,400		
Iloraz wiarygodności	,805	1	,370		
Dokładny test Fishera				,374	,200
Test związku liniowego	,804	1	,370		
N Ważnych obserwacji	1307				

a. Obliczone wyłącznie dla tabeli 2x2.

b. ,0% komórek (0) ma liczebność oczekiwaną mniejszą niż 5. Minimalna liczebność oczekiwana wynosi 283,07.

Rysunek I.31

Test chi-kwadrat dla zmiennych **abany** i **polviewr**

Testy Chi-kwadrat

	Wartość	df	Istotność asymptotyczna (dwustronna)
Chi-kwadrat Pearsona	54,390[b]	2	,000
Iloraz wiarygodności	54,988	2	,000
Test związku liniowego	53,483	1	,000
N Ważnych obserwacji	1238		

b. ,0% komórek (0) ma liczebność oczekiwaną mniejszą niż 5. Minimalna liczebność oczekiwana wynosi 163,75.

wynikiem błędu z próby, nie jest równe dokładnie 0 – co mogłoby mieć miejsce, ale szanse na to są niewielkie. Konkretnie ujmując, prawdopodobieństwo jest mniejsze niż 0,001, czy inaczej: jeden do tysiąca, co jest często używanym standardem istotności statystycznej. Wnioskujemy zatem, że zaobserwowana w tej starannie dobranej próbie zależność z dużym prawdopodobieństwem jest odzwierciedleniem czegoś, co istnieje w szerszej populacji.

Korelacja i regresja

Do tej pory badaliśmy dane nominalne i porządkowe, które stanowią zasadniczą część danych w naukach społecznych. SPSS umożliwia także pracę z danymi mierzonymi na skali interwałowej (przedziałowej) i ilorazowej.

Może na przykład słyszałeś, że ludzie z wyższym wykształceniem mają na ogół mniej dzieci niż ludzie o niższym poziomie wykształcenia. Użyjmy zatem SPSS, aby to spraw-

Rysunek I.32

Korelacja według momentu iloczynowego Pearsona między zmiennymi **educ** i **childs**

Korelacje

		NUMBER OF CHILDREN	HIGHEST YEAR OF SCHOOL COMPLETED
NUMBER OF CHILDREN	Korelacja Pearsona	1	-,216**
	Istotność (dwustronna)	.	,000
	N	2078	2070
HIGHEST YEAR OF SCHOOL COMPLETED	Korelacja Pearsona	-,216**	1
	Istotność (dwustronna)	,000	.
	N	2070	2082

**. Korelacja jest istotna na poziomie 0.01 (dwustronnie).

dzić. W Generalnym Sondażu Społecznym (GSS) zmienne te nazywają się **educ** oraz **childs**. Wybierz **Korelacje** z menu **Analiza** i gdy pojawi się kolejny wybór, zaznacz **Parami** (SPSS może wykonywać także bardziej złożone analizy korelacji, ale w ramach tego wprowadzenia poprzestańmy na prostszej). W oknie dialogowym **Korelacje** zaznacz **educ** oraz **childs** i kliknij na **OK**. Efekt tego działania został pokazany na rysunku I.32.

Jak widzisz, istnieje korelacja o wartości –0,216 (lub inaczej: „ujemna korelacja o wartości 0,216") między tymi dwiema zmiennymi. Ujemna korelacja oznacza, że wraz ze wzrostem liczby lat przeznaczonych na edukację, zmniejsza się liczba posiadanych dzieci.

Oczywiście analiza ta nie może określić kierunku związku przyczynowego. Można więc także powiedzieć, że gdy zwiększa się liczba posiadanych dzieci, to liczba lat przeznaczonych na edukację maleje. Obie interpretacje mają sens i prawdopodobnie sprawdzają się w pewnych przypadkach. Niektórzy młodzi ludzie przymusowo skracają liczbę lat nauki ze względu na własne obowiązki rodzicielskie. Z kolei ci, którzy kontynuują naukę, mogą być zmuszeni do odłożenia w czasie decyzji o rodzicielstwie, a gdy już się zdecydują, mają mniej dzieci.

Jakkolwiek będziemy tłumaczyć ten związek, SPSS informuje nas, że korelacja jest istotna na poziomie 0,01. Innymi słowy, błąd z próby mógłby wygenerować taką korelację mniej niż raz na sto.

Mimo że w tej analizie wybraliśmy tylko dwie zmienne, SPSS przyjmie naraz tyle zmiennych, ile tylko chcesz i stworzy macierz korelacji każdej zmiennej z każdą inną. Spróbuj poeksperymentować z tą możliwością.

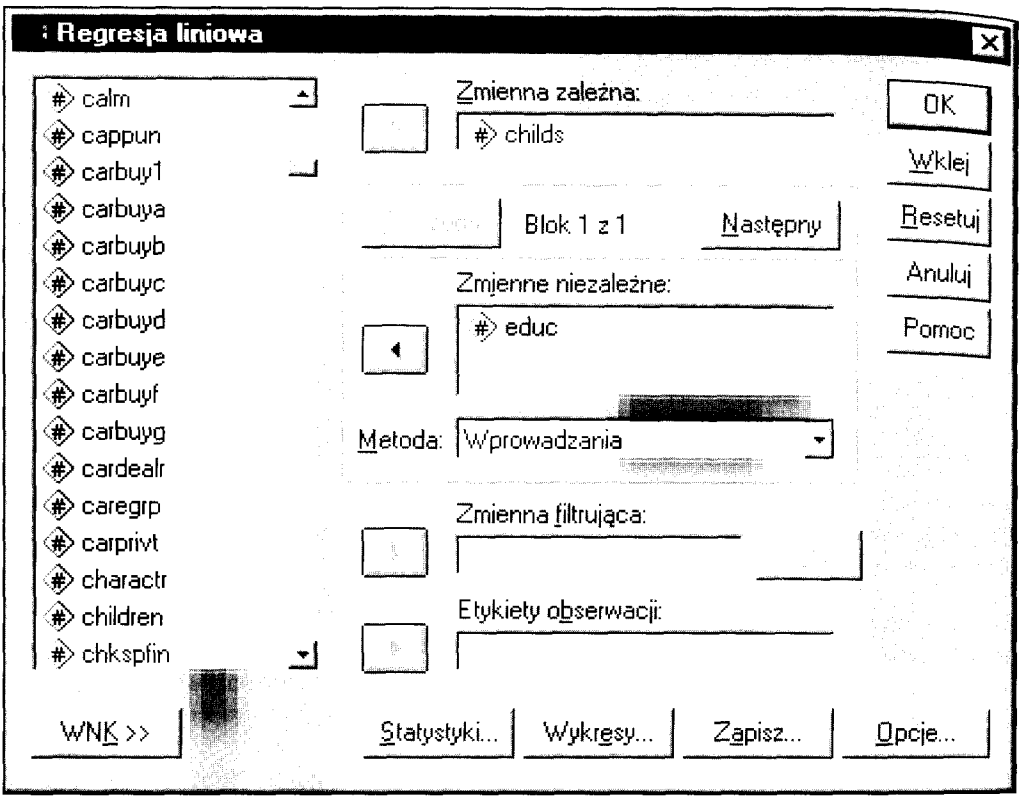
Analiza regresji opiera się na logice korelacji i tworzy równania, które przewidują wartość jakiejś zmiennej na podstawie wartości innych zmiennych. Oto jak można przedstawić zależność między zmiennymi **educ** oraz **childs** w postaci równania regresji. Wybierz **Regresja** z menu **Analiza**, a potem spośród możliwych opcji zaznacz **Liniowa**. Na ekranie ukaże się okno dialogowe, takie jak na rysunku I.33.

Zastosujmy logikę służącą wyjaśnieniu liczby posiadanych dzieci, a zatem **childs** jest naszą zmienną zależną, a **educ** zmienną niezależną, która będzie wyjaśniała różnice w liczbie dzieci. Wpisz dwie zmienne w okno dialogowe, tak jak u góry rysunku I.33. Kliknij **OK**, a otrzymasz wynik, taki jak na rysunku I.34.

SPSS zaprezentuje trzy tabele wyników, ale my w tej chwili jesteśmy zainteresowani tylko trzecią tabelą – **Współczynniki**. Tak naprawdę interesuje nas tylko pierwsza kolumna tej tabeli – **Współczynniki niestandaryzowane**. Pierwszy z tych współczynników – **Stała** – jest wartością zmiennej zależnej (liczba dzieci), gdy wartość zmiennej niezależnej (liczba lat nauki)

Rysunek I.33

Okno dialogowe regresji



wynosi zero. Statystycy często określają ją jako *odcinek y* lub jako punkt, w którym na wykresie linia regresji przecina oś *y*.

Wartość B, związana ze zmienną niezależną (–0,122), wskazuje, o ile zmienia się wartość zmiennej zależnej, gdy zmienna niezależna zmienia się o 1. W naszym przykładzie oznacza to oczekiwaną zmianę liczby posiadanych dzieci wraz z ukończeniem dodatkowego roku nauki. Regresja, wyrażona w postaci równania, wygląda tak:

childs = 3,41 – (0,122 x **educ**)

Załóżmy, że jakaś osoba ma za sobą 10 lat nauki. Przewidujemy, że będzie mieć ona

3,41 – 1,22 = 2,09 dzieci

Rysunek I.34

Przewidywanie wartości zmiennej **childs** na podstawie zmiennej **educ** za pomocą regresji liniowej

Współczynniki[a]

Model		Współczynniki niestandaryzowane		Współczynniki standaryzowane	t	Istotność
		B	Błąd standardowy	Beta		
1	(Stała)	3,410	,168		20,266	,000
	HIGHEST YEAR OF SCHOOL COMPLETED	-,122	,012	-,216	-10,077	,000

a. Zmienna zależna: NUMBER OF CHILDREN

Dla osoby z wyższym wykształceniem, po 16 latach nauki, przewidujemy, że będzie miała

$$3{,}41 - 1{,}952 = 1{,}458 \text{ dzieci}$$

Rzecz jasna, te szacunki obrazują średnie statystyczne, gdyż nikt nie może mieć 2,09 lub 1,458 dziecka. Gdyby jednak chcieć, tylko na podstawie liczby lat spędzonych w szkole, przewidywać, ile dzieci będzie mieć dana osoba, to równanie to byłoby najlepszym do tego narzędziem. Gdybyś przyjmował liczne zakłady tego rodzaju, na pewno byś wygrywał.

Aby dowiedzieć się o regresji czegoś więcej, spróbuj dodać jeszcze jedną zmienną niezależną. SPSS pokaże nową wartość odcinka *y* i nowe współczynniki dla każdej ze zmiennych niezależnych. Upewnij się, że właściwie interpretujesz dodatnie i ujemne znaki.

Tworzenie indeksów

W rozdziale 6 omawialiśmy tworzenie złożonych mierników, takich jak indeksy oraz skale. Teraz zobaczymy, jak można wykorzystać SPSS do stworzenia prostego indeksu.

Bez odświeżania wiedzy na temat logiki tworzenia indeksów, stwórzmy indeks tolerancji seksualnej, do którego włączymy następujące zmienne z Generalnego Sondażu Społecznego:

premarsx:	seks przedmałżeński
xmarsex:	seks pozamałżeński
homosex:	stosunki homoseksualne

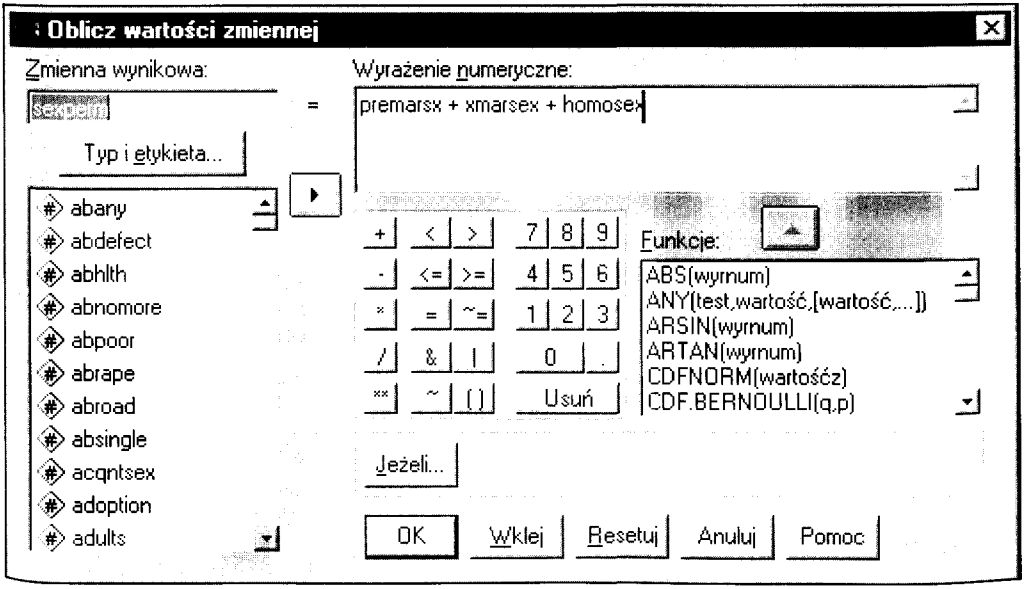
W przypadku każdej z tych zmiennych respondent był pytany, czy takie działania są:

1) zawsze niewłaściwe,
2) prawie zawsze niewłaściwe,
3) czasem niewłaściwe,
4) nie ma w nich nic niewłaściwego.

Ze względu na identyczną formę każdego z trzech pytań możemy dość łatwo zbudować indeks złożony. Pomimo że wartości 1–4 użyte do oznaczania odpowiedzi są tylko etykietami – tak jak wcześniej „1" dla mężczyzny oraz „2" dla kobiety – możemy skorzystać z nich jako liczb. W każdym z tych pytań wyższa wartość liczbowa oznacza wyższy poziom tolerancji seksualnej. Jeśli dodamy do siebie liczby, oznaczające odpowiedzi respondenta na każde z tych trzech pytań, możliwe do uzyskania sumy będą mieścić

Rysunek I.35

Sumowanie wartości zmiennych **premarsx**, **xmarsex**, **homosex**



Rysunek I.36
Rozkład częstości zmiennej **sexperm**

SEXPERM

		Częstość	Procent	Procent ważnych	Procent skumulowany
Ważne	3,00	137	6,6	21,7	21,7
	4,00	42	2,0	6,6	28,3
	5,00	70	3,3	11,1	39,4
	6,00	117	5,6	18,5	57,9
	7,00	45	2,2	7,1	65,0
	8,00	42	2,0	6,6	71,7
	9,00	113	5,4	17,9	89,6
	10,00	43	2,1	6,8	96,4
	11,00	13	,6	2,1	98,4
	12,00	10	,5	1,6	100,0
	Ogółem	632	30,2	100,0	
Braki danych	Systemowe braki danych	1458	69,8		
Ogółem		2090	100,0		

się w przedziale od 3 do 12. Najwyższy stopień tolerancji seksualnej będzie oznaczony przez wartość 12, a najniższy przez wartość 3.

Możemy teraz wygenerować indeks przez użycie **Oblicz wartości** z menu **Przekształcenia**, tak jak pokazano na rysunku I.35. Wprowadź informacje, wpisując nazwy zmiennych albo wybierając z listy i klikając znak „+" na małej klawiaturze, zamieszczonej w oknie dialogowym (rysunek I.35)

Gdy skończysz, naciśnij **OK** w oknie dialogowym. SPSS stworzy w zbiorze danych nową zmienną – **sexperm** i każdemu z respondentów przypisze odpowiednią jej wartość. Jeśli przesuniesz się w oknie edytora danych do końca w prawo, zobaczysz nową zmienną w ostatniej kolumnie. Przesuwając się w górę i w dół obejrzyj jej wartości przypisane poszczególnym respondentom. Ci, którzy nie mają przypisanej żadnej wartości w nowej zmiennej, musieli nie odpowiedzieć na żadne z pytań użytych do jej konstrukcji.

Aby mieć bardziej całościowy ogląd nowej zmiennej, zobacz, jak wygląda jej rozkład częstości (rysunek I.36).

Po skonstruowaniu takiego złożonego miernika warto zawsze, jeśli można, dokonać jego walidacji. To znaczy, jeśli punktacja indeksu ukazuje prawdziwe różnice w poziomie tolerancji seksualnej, powinna ona umożliwić przewidywanie odpowiedzi respondentów na inne pytania. Na przykład można sprawdzić, czy postawy wobec aborcji mają związek z tolerancją seksualną. Możemy to zbadać, tworząc tabelę krzyżową dla naszego indeksu i, powiedzmy, zmiennej **abany**.

Rezultat takiej walidacji jest pokazany na rysunku I.37. Zauważ, że otrzymana tabela ma nieco inny format niż tworzone poprzednio. Z powodu dużej ilości kategorii zmiennej **sexperm** byłoby trudno zmieścić całą tabelę na ekranie komputera (i w tej książce). Dlatego zdecydowałem się umieścić zmienną **sexperm** w rzędach, **abany** w kolumnach, a procentowanie tabeli ustawiłem w rzędach, a nie w kolumnach. Przez to czytamy teraz tabelę „w dół", podczas gdy wcześniej czytaliśmy ją „w poprzek".

Związek między tymi dwiema zmiennymi jest bardzo silny i jednoznaczny. Spośród osób, dla których wartość indeksu wynosiła 3 (co oznacza najniższy poziom tolerancji seksualnej), tylko 18% popiera prawo do aborcji na życzenie kobiety. Procent ten konsekwentnie rośnie wraz ze wzrostem wartości indeksu i osiąga 90% przy wartości indeksu równej 12.

Tworzenie indeksu ze zmiennych, których wartości nie poddają się sumowaniu w taki

Rysunek I.37
Walidacja indeksu **sexperm**

Tabela krzyżowa SEXPERM * ABORTION IF WOMAN WANTS FOR ANY REASON

			ABORTION IF WOMAN WANTS FOR ANY REASON		
			YES	NO	Ogółem
SEXPERM	3,00	Liczebność	23	108	131
		% z SEXPERM	17,6%	82,4%	100,0%
	4,00	Liczebność	11	29	40
		% z SEXPERM	27,5%	72,5%	100,0%
	5,00	Liczebność	30	35	65
		% z SEXPERM	46,2%	53,8%	100,0%
	6,00	Liczebność	59	54	113
		% z SEXPERM	52,2%	47,8%	100,0%
	7,00	Liczebność	24	18	42
		% z SEXPERM	57,1%	42,9%	100,0%
	8,00	Liczebność	25	16	41
		% z SEXPERM	61,0%	39,0%	100,0%
	9,00	Liczebność	80	30	110
		% z SEXPERM	72,7%	27,3%	100,0%
	10,00	Liczebność	28	10	38
		% z SEXPERM	73,7%	26,3%	100,0%
	11,00	Liczebność	11	2	13
		% z SEXPERM	84,6%	15,4%	100,0%
	12,00	Liczebność	9	1	10
		% z SEXPERM	90,0%	10,0%	100,0%
Ogółem		Liczebność	300	303	603
		% z SEXPERM	49,8%	50,2%	100,0%

łatwy sposób, jest nieco bardziej skomplikowane. Aby to zilustrować, stwórzmy indeks obrazujący stosunek respondenta do kwestii broni palnej. Odnoszą się do niej dwa pytania GSS:

gunlaw: czy jest za, czy przeciw zezwoleniom na broń
1) za,
2) przeciw;

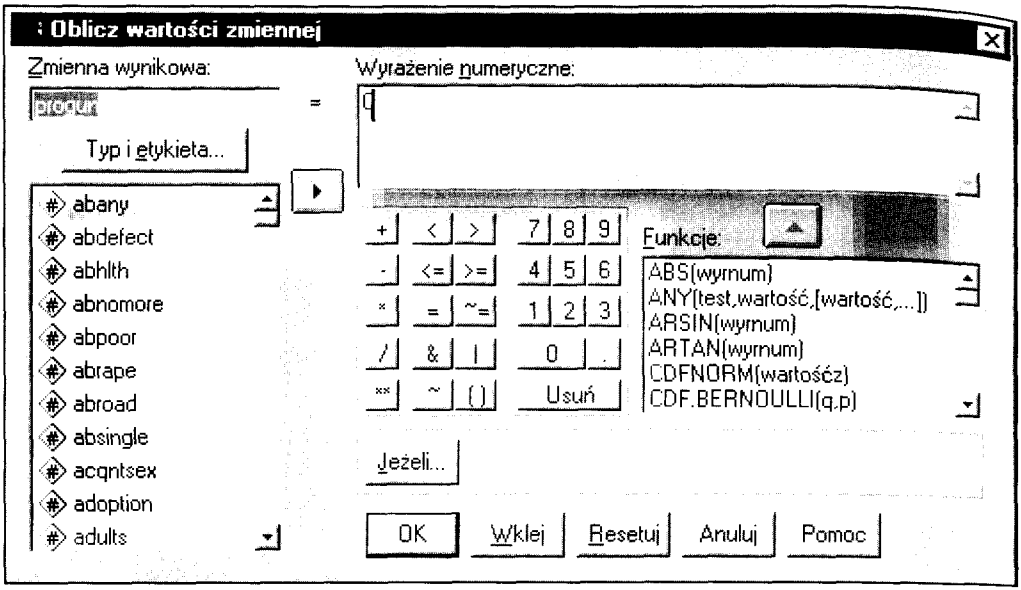
owngun: czy ma broń w domu
1) tak,
2) nie.

Wydaje się logiczne, że ci, którzy mają broń w domu i są przeciwni zezwoleniom na broń, są największymi zwolennikami broni palnej. Natomiast ci, którzy nie mają broni na własność i są zwolennikami zezwoleń na broń, byliby największymi przeciwnikami broni palnej. Zauważ, że poglądy zwolenników broni palnej są reprezentowane przez „2" w zmiennej **gunlaw** oraz „1" w **owngun**. Nie możemy zatem po prostu dodać do siebie tych wartości. Można utworzyć prosty indeks z tych dwóch pytań w inny sposób.

Stwórzmy nową zmienną **progun**, dla której wyższe wartości oznaczają większe poparcie dla broni palnej. Aby zacząć tę procedurę, powróć do menu **Przekształcenia/Oblicz wartości**. W oknie **Zmienna wynikowa** wpisz **progun**, a w oknie obok wpisz „0" jako początkową wartość dla wszystkich (rysunek I.38). Naciśnij **OK**, aby stworzyć nową zmienną. Powróć do **Prze-**

Rysunek I.38

Rozpoczęcie procesu tworzenia zmiennej **progun**

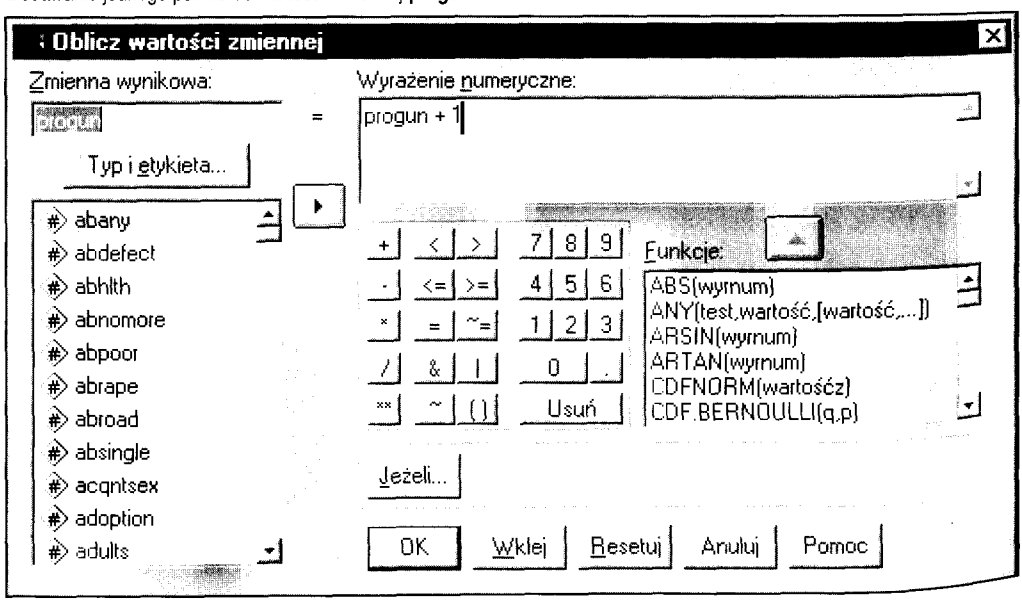


kształcenia/Oblicz wartości i zmień „0" na „progun +1", tak jak pokazano na rysunku I.39.

Jednak nie mamy wcale zamiaru dodawać wszystkim jednego punktu do wartości indeksu. Naciśnij znajdujący się u dołu okna dialogowego przycisk **Jeżeli**. Będziemy wtedy mogli ustalić warunki, pod którymi indeks zostanie zwiększony o 1. Zaznacz opcję „**Uwzględnij, jeśli obserwacja spełnia warunek:**". Potem utwórz warunek pokazany na rysunku I.40. W ten spo-

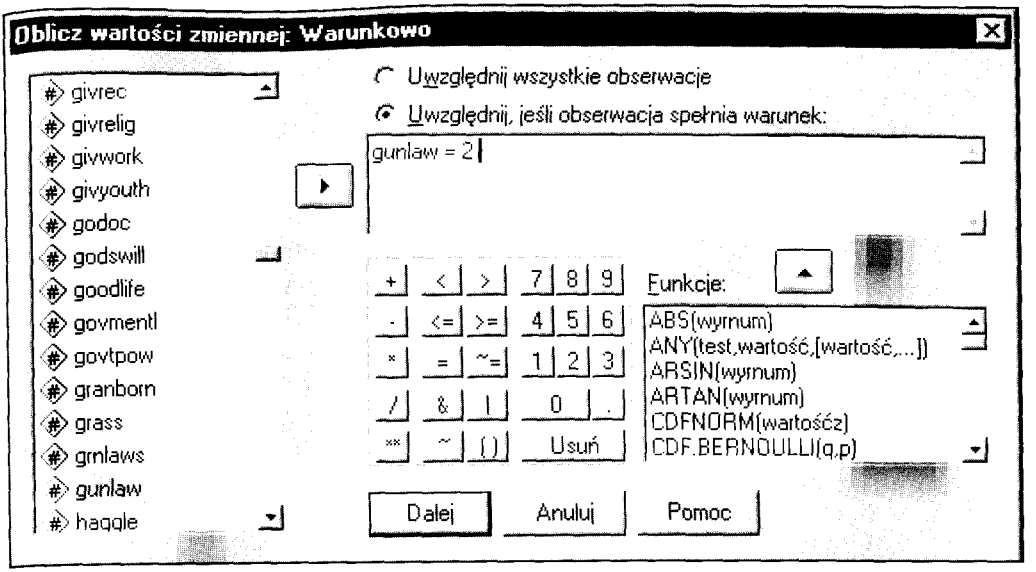
Rysunek I.39

Dodawanie jednego punktu do wartości zmiennej **progun**



Rysunek I.40

Dodawanie jednego punktu przeciwnikom zezwoleń na broń



sób zażądaliśmy, aby SPSS dodał punkt każdemu, kto opowiada się przeciw zezwoleniom na broń (a zatem ma „2" w **gunlaw**).

Naciśnij **Dalej**, aby powrócić do wcześniejszego okna dialogowego. Następnie kliknij na **OK**, aby SPSS przyjął wpisany warunek. Gdy SPSS zapyta, czy chcesz zmienić istniejącą zmienną, wybierz **OK**.

Wybierz jeszcze raz **Przekształcenia/Oblicz wartości**. Zauważ, że wcześniejsze polecenie dodania punktu jest tam cały czas wpisane. Zostaw je, ale naciśnij **Jeżeli**, aby zmodyfikować warunek. Zmień go na tych, którzy zadeklarowali posiadanie broni palnej (a zatem mieli „1" w **owngun**), wpisując „owngun=1" jako warunek. Potem, tak jak poprzednio, kliknij po kolei na **Dalej**, **OK** oraz **OK**. Teraz ci, którzy mieli wartość „0" z powodu popierania zezwoleń na broń, będą mieli „1", jeśli posiadają broń, lub nadal pozostaną z wartością „0", jeśli broni nie mają. Ci, którzy dostali punkt za sprzeciw wobec zezwoleń na broń, dostaną jeszcze jeden, jeśli mają broń (i będą razem mieli „2"), lub pozostaną w punkcie 1, jeśli jej nie mają. Końcowy indeks ma trzy możliwe wartości „0", „1" oraz „2".

Jest jeszcze tylko jeden problem z taką formą indeksu. Jako że wszyscy zaczynali z wartością „0", to ci, którzy nie odpowiedzieli na jedno lub oba pytania, też będą mieli wartość indeksu „0", co mogłoby sugerować, że są przeciwni broni palnej. Ostatni krok w tworzeniu indeksu polega na wyeliminowaniu respondentów z brakami danych.

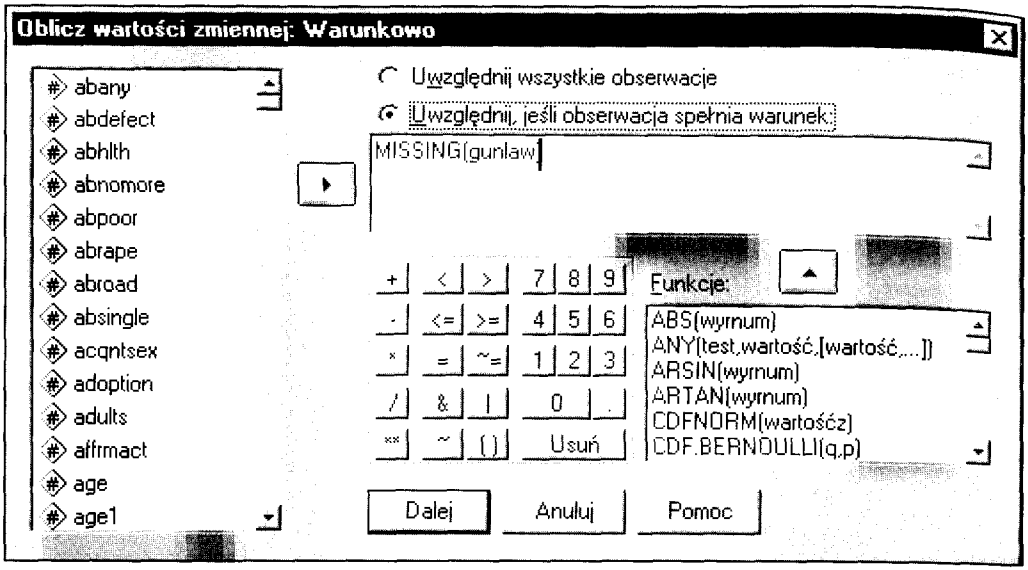
Stwórzmy najpierw kod dla „braku danych". Użyjemy do tego „99". Powróć do **Przekształcenia/Oblicz wartości**. W pierwszym oknie dialogowym wpisz „**progun**=99". Naciśnij **Jeżeli**, aby określić warunek: „MISSING(**gunlaw**)", tak jak pokazano na rysunku I.41. Naciśnij **Dalej** oraz **OK** i powtórz całą procedurę dla „MISSING (**owngun**)".

Jeśli jednak przyjrzysz się możliwym odpowiedziom dla zmiennej **owngun**, zobaczysz, że sześć osób odmówiło odpowiedzi, co zostało zakodowane jako „3". Powróć zatem do **Przekształcenia/Oblicz wartości** i przypisz wartość indeksu 99 każdemu, dla kogo „**owngun**=3".

Ostatnim krokiem jest zrekodowanie wartości 99. Wybierz **Przekształcenia/Rekoduj**, ale tym razem zaznacz **Na te same zmienne**. Gdy pokaże się okno dialogowe, przekształć 99 na **SYSMIS**, tak jak pokazano na rysunku I.42. Wpisz 99 jako wartość źródłową, naciśnij „**Systemowy brak danych**", a potem kliknij na przycisk **Dodaj**.

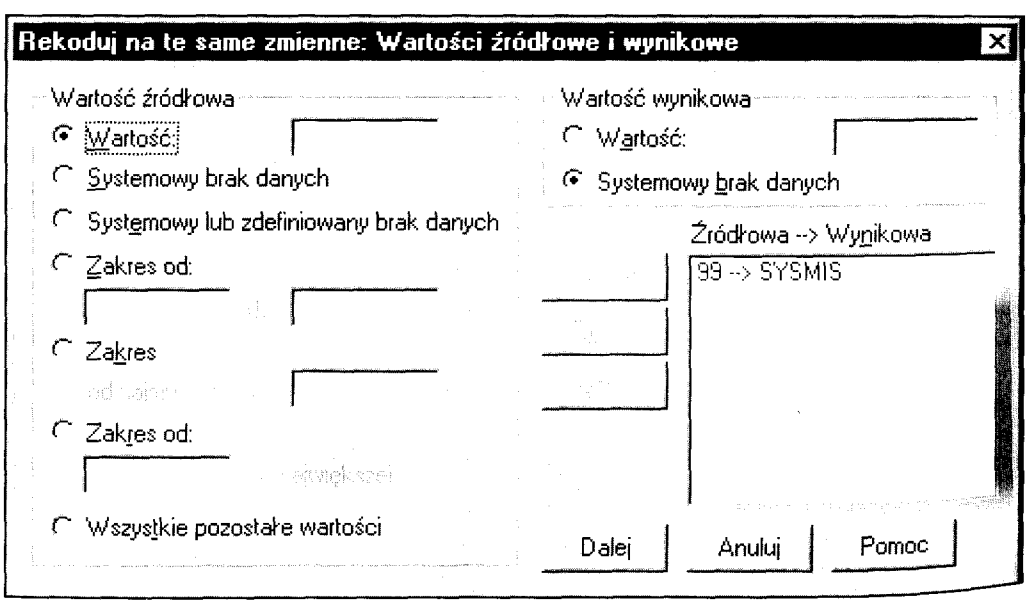
Rysunek I.41

Braki danych jako warunek



Rysunek I.42

Przekształcanie 99 na **SYSMIS**



Aby zakończyć to działanie, kliknij na **Dalej**, a potem na **OK**. Indeks jest już gotowy. Możesz to sprawdzić, uruchamiając **Analiza/Opis statystyczny/Częstości**. Dla większej pewności utwórz tabelę krzyżową dla tych dwóch zmiennych. Pozwoli ona sprawdzić poprawność liczby respondentów dla odpowiednich wartości punktowych.

Grafika

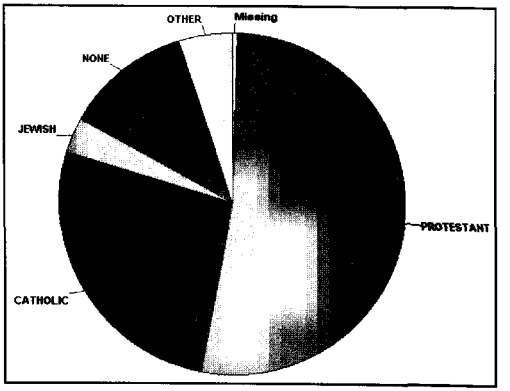
Wraz z postępem w dziedzinie grafiki komputerowej SPSS oferuje wiele rodzajów prezentacji danych w formie innej niż tabelaryczna. Przejrzyjmy kilka z nich, zaczynając od prostego rozkładu częstości.

Na rysunku I.43 pokazany jest w formie wykresu kołowego rozkład danych GSS, dotyczących tożsamości religijnej (**relig**). Można taki wykres uzyskać, wybierając z menu **Wykresy** opcję **Kołowy**, zaznaczając **Opisy dla grup obserwacji** i wskazując na **relig** jako zmienną, której rozkład chcemy zobaczyć. Można też rozkład tej samej zmiennej przedstawić w formie wykresu słupkowego, co pokazano na rysunku I.44. Zmienne ilorazowe, takie jak lata nauki, można przedstawić za pomocą wykresu liniowego. Jest to pokazane na rysunku I.45.

Są to tylko niektóre z możliwości graficznych SPSS. Poeksperymentuj z nimi, aż znajdziesz formę prezentacji graficznej najwłaściwszą dla twoich celów.

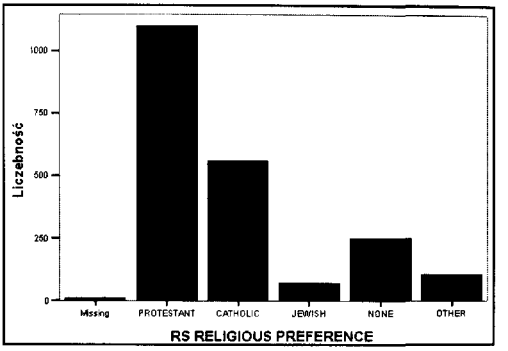
Rysunek I.43

Wykres kołowy dla tożsamości religijnej



Rysunek I.44

Wykres słupkowy zmiennej **relig**



Kopiowanie wyników do dokumentu

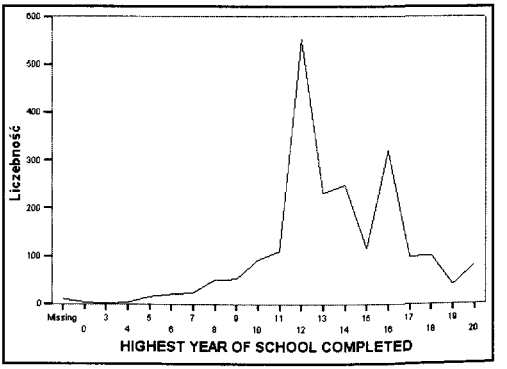
Będziesz często używać SPSS do ilościowych analiz, związanych z przygotowaniem referatu semestralnego, pracy zaliczeniowej czy innego projektu. Uzyskane za pomocą SPSS wyniki można oczywiście ręcznie przepisać do referatu, ale można też skorzystać z oszczędzających energię ułatwień. Zależnie od tego, z jakiego edytora tekstu korzystasz, będziesz być może musiał nieco poeksperymentować.

Spróbuj zrobić kopię tabeli obrazującej rozkład częstości zmiennej **polviewr**. Kliknij na tabelę. Zobaczysz, że została wzięta w ramkę. Oznacza to, że została ona zaznaczona przez komputer. Potem w menu **Plik** zaznacz **Export**. Na rysunku I.46 przedstawiono okno dialogowe, które powinno pokazać się na ekranie.

Masz do wyboru kilka opcji. Dla naszych celów wybierz **Dokument wynikowy**. Możesz teraz albo zmienić nazwę tworzonego pliku

Rysunek I.45

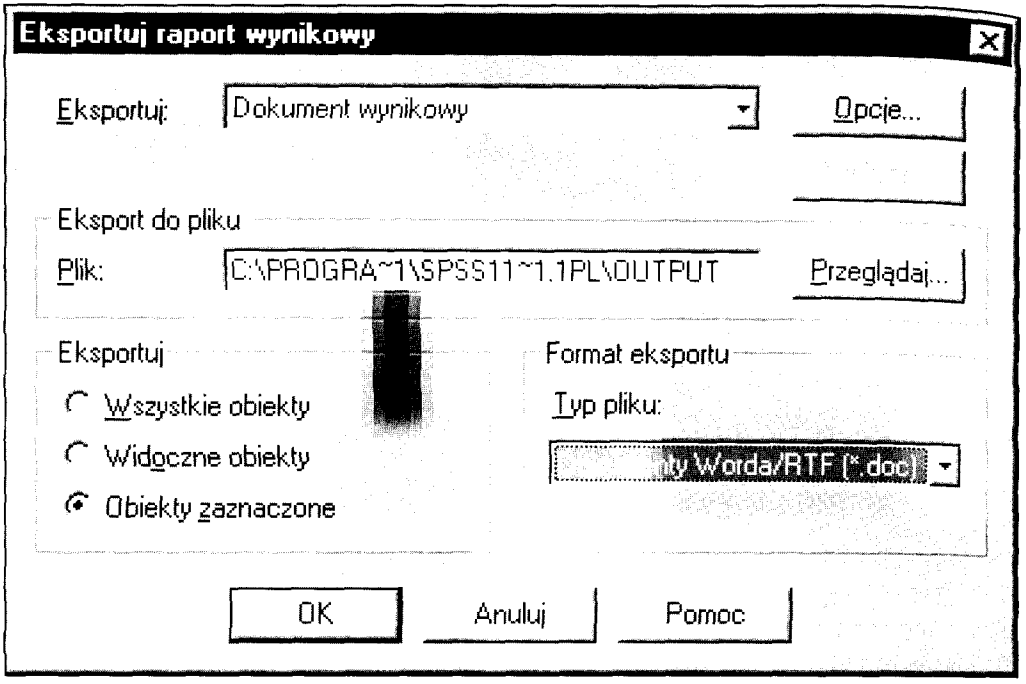
Wykres liniowy zmiennej **educ**



eksportowego, albo zaakceptować (i zapamiętać) nazwę i lokalizację zaproponowane przez SPSS. Potem wybierz do wyeksportowania **Obiekty zaznaczone** oraz wybierz **Dokumenty**

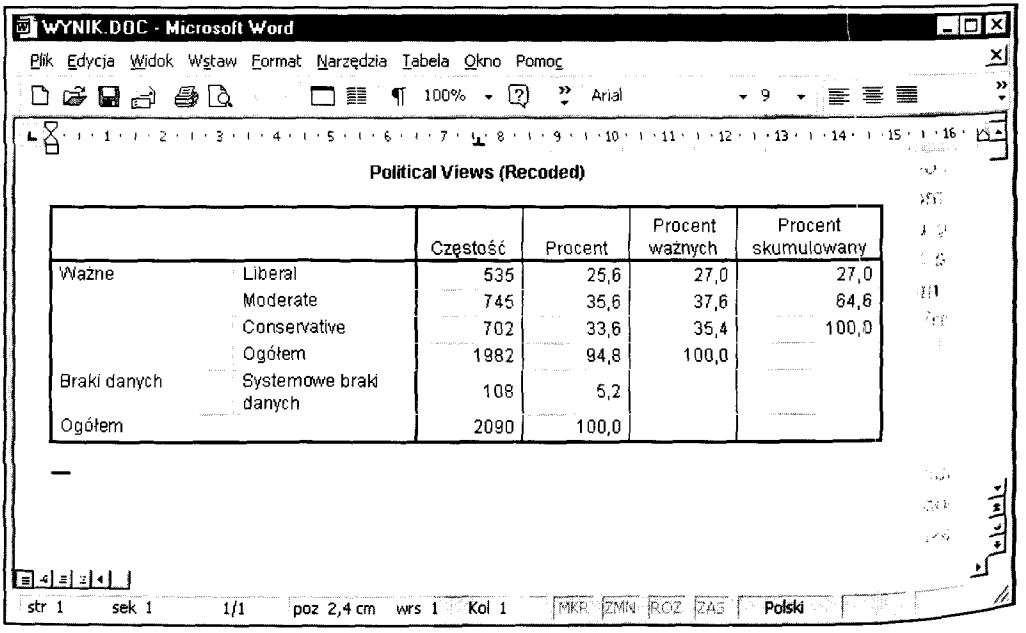
Rysunek I.46

Okno dialogowe eksportu



Rysunek I.47

Tabela rozkładu częstości zmiennej **polviewr** umieszczona w edytorze tekstu



Worda/RTF(*.doc) jako format eksportu. Po wybraniu tych ustawień kliknij **OK**.

Uruchom teraz edytor tekstu i otwórz dokument, który rozpaczliwie domaga się tabeli. Ustaw kursor w miejscu, gdzie chcesz umieścić tabelę, i wybierz **Wklej** (najprawdopodobniej w menu **Edycja**). Powinieneś zostać nagrodzony czymś w rodzaju tabeli, którą zamieściłem na rysunku 1.47.

Spróbujmy bardziej wyrafinowanego sposobu. Stwórz jeszcze raz tabelę rozkładu częstości. Gdy masz ją na ekranie, wciśnij klawisz **Print Screen** na swoim komputerze (być może będziesz musiał zrobić to, trzymając jednocześnie klawisz **Shift**, **ALT** lub **CTRL**. Poeksperymentuj). Obraz całego ekranu zostanie wtedy skopiowany do schowka w twoim komputerze.

Masz w tym momencie dwie drogi do wyboru. Możesz wrócić do docelowego dokumentu i użyć komendy **Wklej**, aby wkleić wszystko, co przed chwilą skopiowałeś, w wybranym przez ciebie miejscu. Inna metoda to skorzystanie z programu graficznego, takiego jak na przykład Adobe Photoshop. Uruchom program i utwórz pusty plik. Wklej do niego zawartość schowka. Możesz teraz usunąć niechciane fragmenty obrazu lub zmienić plik w inny sposób (to właśnie dzięki tej metodzie stworzono rysunki w tym załączniku).

Można też wykorzystać możliwość obsługi licznych innych formatów tabel, oferowaną przez SPSS. Aby je przejrzeć, wybierz w SPSS **Edycja/Opcje** i kliknij na **Tabele przestawne**. Wybierając potem różne opcje z okna **Szablon**, możesz z prawej strony zobaczyć próbkę danego układu tabeli. Jeśli któryś z formatów ci się spodoba, zaznacz go. Następnie zamknij okno dialogowe i utwórz nową tabelę. Będzie miała wybrany przez ciebie format.

Zamykanie programu

Bez względu na to jak bardzo polubisz SPSS, przyjdzie moment, gdy trzeba będzie program zamknąć. Wejdź do menu **Plik** i wybierz opcję **Zakończ**. SPSS zapyta, czy chcesz zachować **zawartość edytora raportów**. Jeśli jakoś go nazwiesz i wskażesz, gdzie SPSS ma go zapisać, to będziesz mógł go później otworzyć. Jeśli tylko ćwiczyłeś z SPSS, najprawdopodobniej wybierzesz „Nie".

Jeśli w czasie pracy zmieniłeś zbiór danych, na przykład tworząc nową zmienną, SPSS zapyta, czy zmiany mają zostać zapisane. Jeśli nie chcesz się ich pozbywać, wybierz „Tak". Powinieneś jednak zmienić zbiór danych tylko wtedy, gdy wolno ci to zrobić. Jeśli jesteś współużytkownikiem pliku, na przykład dzielisz go z innymi studentami z twojej grupy, zapisywanie przez ciebie zmian może być niedopuszczalne. Jeśli masz jakieś wątpliwości, zapytaj o to prowadzącego zajęcia.

SPSS zostanie teraz zamknięty, podobnie jak ten dodatek. Miłej zabawy.

Dwadzieścia pytań, które dziennikarz powinien zadać, zapoznając się z wynikami sondażu opinii*

Sheldon R. Gawiser
G. Evans Witt

Zarówno dla dziennikarzy, jak i dla autorów sondaży opinii pytania są głównym narzędziem zdobywania informacji. Oto 20 pytań, które dziennikarz powinien zadać autorowi sondażu, zanim napisze o jego wynikach. Tekst ten ma pomóc dziennikarzom w przygotowaniu starannych, profesjonalnych materiałów, dotyczących badań opinii. Nie należy go traktować jako wprowadzenia do metodologii badań sondażowych.

Będziemy tutaj pisać o sondażach „naukowych". Pytania pomogą rozstrzygnąć, czy dane badania są wartym zainteresowania sondażem „naukowym", czy też sondażem nienaukowym, być może zabawnym, ale nie niosącym żadnych treści.

Spośród sondaży naukowych chyba najbardziej znane i najstarsze są sondaże Gallupa, ale znakomite sondaże przeprowadza także wiele innych renomowanych firm.

Sondaże nienaukowe są mniej znane, ale szeroko rozpowszechnione: „sondaże audiotele", „sondy uliczne", sondaże w centrach handlowych, a nawet klasyczne *toilet tissue poll*, w których opinię wyraża się poprzez wybór papieru toaletowego, ozdobionego portretami kandydatów na poszczególnych listkach.

Główna różnica między sondażami naukowymi a nienaukowymi polega na tym, kto dobiera respondentów. W sondażach naukowych respondentów wyznacza i wyszukuje osoba prowadząca badania. W sondażach nienaukowych badani z reguły wybierają się sami i z własnej inicjatywy wygłaszają swoje opinie.

Wyniki dobrze przeprowadzonych naukowych badań sondażowych mogą być wiarygodnym źródłem informacji o opiniach znacznie większej liczby ludzi niż tylko ci, z którymi przeprowadzano wywiady – nawet wszystkich Amerykanów. Wyniki sondaży nienaukowych mówią tylko o tym, co powiedzieli biorący w nich udział respondenci.

Wykorzystując poniższe 20 pytań, dziennikarz może przyjrzeć się faktom i zadecydować, jak należy potraktować każdy z sondaży, które codziennie trafiają na redakcyjne biurka.

1. Kto przeprowadził sondaż? Pierwsze pytanie powinno zawsze brzmieć: jaka agencja badawcza, organizacja polityczna, firma czy też inna grupa przeprowadziła dane badania?

Jeśli nie będziesz wiedzieć, kto przeprowadził sondaż, nie uda ci się uzyskać odpowiedzi na pozostałe pytania. A jeśli osoba udostępniająca wyniki sondażu nie umie lub nie chce powiedzieć, kto go przeprowadził, są powody, by podać w wątpliwość rzetelność i wiarygodność jego wyników.

W większości przypadków renomowane firmy badawcze udostępniają informacje pozwalające ocenić wartość sondażu. A ponieważ renoma ma dla porządnej firmy sporą wartość,

* Pierwotnie opublikowane w postaci broszury przez *National Council on Public Polls* (Krajową Radę do Spraw Badania Opinii Publicznej). Przedruk za pozwoleniem.

w sondażach przeprowadzanych przez profesjonalistów udaje się uniknąć wielu błędów.

2. Kto zapłacił za sondaż i dlaczego został on przeprowadzony? Musisz wiedzieć, kto sfinansował badania, ponieważ dowiesz się z tego – i dowiedzą się twoi czytelnicy lub słuchacze – kto uznał, że problematyka sondażu jest na tyle ważna, by wydawać pieniądze na sprawdzenie, co ludzie o tym myślą. A stąd już tylko krok do bardziej generalnego pytania: dlaczego w ogóle przeprowadzono ten sondaż?

Sondaży z reguły nie przeprowadza się po prostu dla dobra ludzkości. Przeprowadza się je z jakiegoś powodu: albo by zdobyć użyteczne informacje, albo by uzyskać poparcie dla jakiejś sprawy.

Może chodzić o to, by gazeta mogła opublikować dobry tekst, albo o to, by jakiś polityk ponownie wygrał wybory, bądź też jakaś firma chce wprowadzić na rynek nowy produkt. Może jakaś grupa interesów chciała sprawdzić, czy jej poglądy pokrywają się z poglądami społeczeństwa.

Wszystko to może być dobrym powodem do przeprowadzenia sondażu.

Ważnym zadaniem dziennikarza jest ocena, czy konkretny motyw nie wywołuje wątpliwości co do rzetelności sondażu na tyle poważnych, by wykluczały one celowość opublikowania jego wyników.

Przykładem podejrzanych sondaży są sondaże prowadzone przez komitety wyborcze. Ich celem jest wyłącznie ułatwienie zwycięstwa własnemu kandydatowi – i nic poza tym. W takim sondażu mogą być zadawane skrajnie tendencyjne pytania albo stosowane bardzo dziwne techniki doboru próby – wszystko po to, by zrealizować taktyczne cele kampanii wyborczej. Celem takiego sondażu może być na przykład testowanie nowych haseł, nowego stanowiska w jakiejś ważnej kwestii publicznej albo nowego zarzutu przeciwko konkurentom.

Ponieważ jednak precyzyjna ocena odczuć ogółu społeczeństwa nie jest celem takiego sondażu, do publikowania jego wyników należy podchodzić z dużą ostrożnością.

Ryzykowne mogą być także sondaże zlecone przez grupy interesu. Na przykład organizacje ekologiczne mogą ogłosić wyniki sondażu, według których Amerykanie popierają podjęcie zdecydowanych działań na rzecz ochrony środowiska. Być może tak jest, ale zleceniodawcy sondażu mieli w tej sprawie bardzo ściśle określone poglądy. Mogły one wpłynąć na sformułowanie pytań, moment przeprowadzenia sondażu, dobór badanych, kolejność pytań. Powinieneś przeanalizować sondaż, by mieć pewność, że dokładnie testuje on opinię publiczną, a nie po prostu forsuje jeden punkt widzenia.

3. Z iloma osobami przeprowadzono wywiady? To również jest podstawowa informacja, którą należy uzyskać. Sondaże dają jedynie przybliżone wyniki – im więcej osób liczy próba w sondażu naukowym, tym mniejszy, przy niezmienionych innych warunkach, będzie błąd związany z wielkością próby.

Trzeba się jednak ustrzec pułapki, tkwiącej w powszechnym przekonaniu, że „więcej automatycznie znaczy lepiej". To zdecydowanie prawda, że w przypadku prawidłowo prowadzonego sondażu większa próba oznacza mniejszy błąd losowy – przy niezmienionych innych warunkach. Ale inne czynniki mogą być ważniejsze w ocenie jakości sondażu.

4. W jaki sposób dobrano badanych? Najważniejszym powodem, dla którego wyniki jednych sondaży dają wierny obraz opinii publicznej, a wyniki innych nadają się tylko do kosza na śmieci, są różnice w sposobie dobierania ludzi do badań.

W sondażach naukowych stosuje się specjalne techniki doboru respondentów. W nienaukowych respondenci dobierają się sami.

Metody stosowane przez profesjonalistów opierają się na mocnej matematycznej podstawie: wyniki sondażu przeprowadzonego na próbie odzwierciedlają opinie panujące w całej populacji wtedy – i tylko wtedy – jeśli dla każdej osoby należącej do tej populacji znamy prawdopodobieństwo bycia wylosowanym do próby. Próba taka nosi wtedy nazwę próby losowej lub probabilistycznej. Dzięki niej wywiady przeprowadzone na próbie 1000 dorosłych Amerykanów mogą trafnie odzwier-

ciedlać opinie ponad 185 milionów dorosłych Amerykanów.

W większości sondaży naukowych stosuje się metody doboru próby pozwalające na przeprowadzenie badań bez ponoszenia nadmiernych kosztów. Na przykład niektóre metody doboru próby w sondażach telefonicznych nie polegają po prostu na wydzwanianiu pod losowo wygenerowane numery – aby zredukować liczbę zmarnowanych połączeń, wybierane są wyłącznie centrale telefoniczne, o których wiadomo, że obsługują numery ze stref osiedli mieszkalnych. Ale w praktyce nawet próba losowa nigdy nie jest do końca losowa – niektórzy nie mają telefonów, niektórzy odmawiają udziału w badaniach, innych nie ma w domu.

5. Z jakiego obszaru: całego kraju, stanu, regionu, lub spośród jakiej grupy: nauczycieli, prawników, wyborców demokratów itp. dobrano respondentów? Wyniki uzyskane na próbie losowej mogą zostać uogólnione na szerszą populację, z której próba ta została dobrana – ale populacja ta musi być precyzyjnie określona. Musimy na przykład wiedzieć, czy próba jest reprezentatywna dla wszystkich mieszkańców USA, czy tylko dla mieszkańców jednego stanu lub miasta. W przypadku sondaży telefonicznych populacja ta ograniczona jest do osób, których mieszkania wyposażone są w telefony. Na ogół gospodarstwa domowe wyposażone w telefony możemy uznać za podobne do ogółu społeczeństwa, ale w przypadku, gdy omawiany sondaż dotyczy kwestii związanych z ubóstwem lub bezdomnością, próba telefoniczna nie byłaby właściwa. Pamiętaj – zastosowanie naukowej techniki doboru próby nie gwarantuje, że wywiady przeprowadzono z przedstawicielami właściwej populacji.

Wiedza o tym, z jakiej populacji dobrano respondentów, ma dla oceny badań znaczenie zasadnicze.

Na przykład, sondaż wśród biznesmenów może odzwierciedlać opinie biznesmenów – ale nie wszystkich dorosłych. Tylko wówczas, gdy respondenci zostali dobrani spośród wszystkich dorosłych Amerykanów, sondaż może odzwierciedlać opinie wszystkich dorosłych Amerykanów.

Sondaże polityczne są szczególnie wrażliwe na ten problem.

W sondażach przed prawyborami i w sondażach przedwyborczych kwestia określenia badanej populacji ma podstawowe znaczenie. Sondaż preferencji wszystkich dorosłych nie jest najlepszym źródłem informacji przed prawyborami[1], w których uczestniczy zwykle tylko 25% zarejestrowanych wyborców. Szukaj więc raczej sondaży przeprowadzanych na zarejestrowanych wyborcach, „prawdopodobnych wyborcach", „uczestnikach poprzednich prawyborów" itd. Takie rozróżnienia są istotne i należy je uwzględniać w opisie wyników sondaży.

Jednym z najtrudniejszych aspektów badań przedwyborczych jest określenie, kto rzeczywiście będzie głosować.

6. Czy wyniki oparte zostały na odpowiedziach wszystkich badanych? Jednym z najprostszych sposobów na wypaczenie wyników sondażu jest powoływanie się na odpowiedzi tylko jednej z grup ludzi biorących udział w badaniach. Na przykład w kwestiach politycznych zachodzi z reguły zasadnicza różnica pomiędzy opiniami demokratów i republikanów. Powoływanie się wyłącznie na opinie demokratów, wyrażone w sondażu, określanym jako obejmujący wszystkich dorosłych, prowadzi bez wątpienia do wypaczenia jego wyników.

Wyniki sondażu opinii demokratów muszą być określone jako takie – i podając je należy wyraźnie wskazywać, że dotyczą one jedynie opinii demokratów.

Oczywiście w pewnych sytuacjach podawanie wyników odnoszących się tylko do jednej grupy może być w pełni uzasadnione. Jeśli sondaż dotyczy republikańskich prawyborów, istotne są poglądy republikanów, mogących w nich uczestniczyć – a nie demokratów, którzy nie mogą brać udziału w głosowaniu Partii Republikańskiej.

[1] Autorzy odnoszą się do amerykańskiego systemu wyborczego, w którym prawybory są instytucją polityczną, służącą wyłonieniu kandydatów danej partii (przyp. red. nauk.).

7. Kto spośród tych, którzy powinni brać udział w badaniach, nie brał w nich udziału? Powinieneś wiedzieć, jak wielu ludzi odmówiło odpowiedzi na pytania sondażu, czy też z iloma w ogóle nie udało się skontaktować.

Poziom braków odpowiedzi (*non-response rate*) informuje, jaki procent ludzi, z którymi się skontaktowano i którzy mieli zostać objęci badaniami, nie wziął udziału w badaniach. Ludzie ci mogli odmówić udzielenia wywiadu, a być może po prostu nie było ich w domu, gdy zgłaszał się do nich ankieter.

Zupełnie inaczej trzeba traktować wyniki sondażu, w którym 100 badanych zostało wylosowanych spośród 1000 delegatów na konwencję, niż sondażu, w którym przeprowadzono tylko 100 wywiadów na 1000 delegatów, gdyż tylko tylu zgodziło się wziąć w nich udział. Podobne ryzyko zakłócenia wyników pojawia się w przypadku, gdy z częścią delegatów, którzy zostali wylosowani do próby, w rzeczywistości nigdy się nie skontaktowano.

8. Kiedy przeprowadzono sondaż? Pojedyncze wydarzenie może mieć dramatyczny wpływ na wyniki sondażu. Ich interpretacja powinna zależeć od tego, czy był on przeprowadzony przed jakimś istotnym wydarzeniem, czy po nim. Nawet najświeższe wyniki mogą zostać szybko zdezaktualizowane przez późniejsze wydarzenia. Prezydent mógł wygłosić porywające przemówienie, giełda mogła się załamać, a w wyniku katastrofy tankowca kilometry plaż mogły zostać zanieczyszczone ropą.

Wyniki sondaży sprzed kilku tygodni czy miesięcy mogą być cennym źródłem wiedzy dla historyków, ale niekoniecznie czymś, co warto publikować.

9. Jaką metodą przeprowadzono wywiady? Sondaże przeprowadza się na trzy główne sposoby: ankieterzy mogą przeprowadzać wywiady w domach respondentów, wywiady mogą być przeprowadzane telefonicznie, można też rozsyłać kwestionariusze pocztą.

Większość sondaży realizowanych jest obecnie przez telefon[2]; wykorzystuje się do tego specjalne pracownie do wywiadów telefonicznych, z których ankieterzy dzwonią do respondentów. Jednak niektóre sondaże wciąż przeprowadza się, wysyłając ankieterów do domów respondentów.

Są także sondaże przeprowadzane za pośrednictwem poczty. W badaniach naukowych badacze dobierają ludzi, do których wysyłane są kwestionariusze; respondenci je wypełniają i odsyłają z powrotem.

Sondaże pocztowe mogą być doskonałym źródłem informacji, ale ich realizacja trwa tygodniami – co oznacza, że nie da się z nich uzyskać tak aktualnych wyników, jak w przypadku sondaży telefonicznych. Sondaże pocztowe narażone są ponadto szczególnie na pewne rodzaje błędów, zwłaszcza na niski poziom realizacji próby – w wielu badaniach pocztowych mniej niż połowa ludzi odsyła wypełnione kwestionariusze, co powoduje, że wartość wyników jest bardzo ograniczona.

Badania realizowane w centrach handlowych, w restauracjach czy wśród przechodniów mogą przynieść jakieś korzyści ich zleceniodawcom, ale ich wyniki nie nadają się do publikowania w prasie. Robiąc wywiady z „ludźmi z ulicy", można usłyszeć wiele interesujących historii – ale nie należy tego mylić z badaniami opinii publicznej.

10. Czy mamy do czynienia z sondażem typu audiotele, opiniami przysyłanymi pocztą lub sondażem wśród prenumeratorów? Nie ma sensu powoływać się na wyniki sondaży typu audiotele, opinii nadsyłanych pocztą lub opierających się na ankietach załączonych do czasopism, ponieważ w ich przypadku uczestnicy sami zgłaszają się do udziału w badaniach. Tego rodzaju pseudosondaże nie mają żadnej wartości. Pamiętaj, że celem sondażu jest wnioskowanie o całej populacji, a nie tylko o próbie. W przypadku wzmiankowanych wyżej „sondaży" nie ma żadnej metody, która pozwoliłaby uogólnić ich wyniki na jakąś

[2] Dotyczy to oczywiście USA – już w znacznie mniejszym stopniu Europy Zachodniej; w Polsce wywiady telefoniczne stosowane są na stosunkowo niewielką skalę i tylko do specyficznych celów (przyp. tłum.).

większą grupę. Na ogół sondaże naukowe wykazują inne wyniki niż pseudosondaże.

Nie ma przeciwwskazań, żeby na podstawie audiotele zdecydować, czy Homar Larry ma być ugotowany w czasie najbliższej „Sobotniej Nocy Na Żywo", czy nawet pozwolić zagorzałym fanom futbolu wyrazić w ten sposób swoje opinie na temat tego, kto był najlepszym obrońcą w National Football League, ale są to zastosowania o charakterze jedynie rozrywkowym. Nie wiadomo, kto naprawdę dzwoni, ile ma lat, albo ile takich telefonów już wykonał.

Nigdy nie daj się nabrać, gdy podadzą ci liczbę przyjętych zgłoszeń. Bywa tak, że kilka osób dzwoni tysiące razy. Nawet jeśli przyjęto 500 tysięcy połączeń, nikt nie wie, co to naprawdę oznacza. Jeśli imponują ci duże liczby, pamiętaj, że nienaukowy sondaż „Literary Digest" na próbie 12 milionów ludzi przewidywał, że Landon wygra z Rooseveltem.

Wartość sondaży przeprowadzanych wśród prenumeratorów, polegających na tym, że czasopisma zamieszczają ankietę, którą po wypełnieniu należy odesłać do redakcji, jest nie lepsza. Znowu nie wiemy, kto ani ile razy odesłał ankietę. Takich wyników nie da się uogólnić nawet na populację prenumeratorów danego czasopisma.

11. Jakim błędem losowym obciążone są wyniki sondażu? Wywiady przeprowadzone na naukowo dobranej próbie 1000 respondentów mogą dać wierny obraz opinii ponad 185 milionów dorosłych Amerykanów. Inaczej mówiąc, przeprowadzenie wywiadów ze wszystkimi 185 milionami dorosłych – zakładając, że byłoby to możliwe – dałoby mniej więcej takie same wyniki jak poprawnie przeprowadzony sondaż.

Co jednak, jeśli dwa starannie przeprowadzone sondaże wśród 1000 dorosłych dadzą nieco różne wyniki? Nie wynika z tego, że któryś z nich był źle zrobiony. To niewielkie zróżnicowanie wyników jest określane jako błąd, wynikający z doboru próby, a potocznie często nazywane marginesem błędu.

Błąd losowy nie jest „błędem" w znaczeniu pomyłki; jest to raczej miara możliwego przybliżenia wyników, wynikająca z zastosowania próby losowej.

Badacze wyrażają rozmiar tej niepewności spowodowanej przez zastosowanie próby za pomocą określenia „poziomu ufności". Oznacza on prawdopodobieństwo, że próba znajduje się w granicach podanego zakresu błędu w stosunku do wyników, jakie uzyskano by przeprowadzając badania na całej populacji. Z reguły przyjmuje się poziom ufności 95%.

Tak więc, jeśli wyniki ogólnokrajowego sondażu podawane są z „3-procentowym marginesem błędu", należy przez to rozumieć, że gdyby przeprowadzono, w tym samym czasie i zadając w taki sam sposób takie same pytania, badania z udziałem wszystkich dorosłych w całym kraju, to w 95% przypadków wyniki sondażu przeprowadzonego na próbie nie różniłyby się od wyników badania wyczerpującego o więcej niż plus minus 3%.

Zwróć uwagę, że nie ma to nic wspólnego z tym, czy ludzie dobrze rozumieli pytania i szczerze na nie odpowiadali; nie ma to także nic wspólnego z innymi zagadnieniami metodologicznymi. „Błąd losowy" (inaczej: błąd z próby) jest tą składową potencjalnej niedokładności badań, która jest wynikiem prowadzenia ich nie na całej populacji, lecz na próbie. Wielkość błędu losowego nic nam nie mówi o liczbie osób, które odmówiły udziału w badaniach lub były nieosiągalne, ani też o błędach spowodowanych przez niewłaściwe sformułowanie jakiegoś pytania lub o obciążeniach wprowadzanych do sytuacji wywiadu przez poszczególnych ankieterów.

Należy pamiętać, że wielkość błędu losowego odnosi się do każdej z liczb podawanych w raporcie z badań – te plus minus 3% należy uwzględnić przy każdej z nich. Tak więc, jeśli w sondażu porównywani są dwaj kandydaci na prezydenta, błąd losowy należy uwzględnić przy obu wynikach.

Wiąże się z tym jeden z najpoważniejszych problemów prezentacji wyników sondaży. Kiedy – w przypadku badań przedwyborczych – można powiedzieć, że jeden z kandydatów rzeczywiście wyprzedza drugiego?

Na pewno wtedy, gdy różnica między nimi jest co najmniej dwa razy większa niż błąd z próby – a więc, w naszym przykładzie, wynosi co najmniej 6%.

Jeśli różnica między nimi jest mniejsza niż błąd z próby, również na pewno nie powinniśmy pisać, że jeden z kandydatów wyprzedza drugiego – a raczej, że „idą łeb w łeb", że „wyniki są prawie takie same", albo że „różnica między kandydatami jest niewielka".

Trzeba także pamiętać, że jeśli podaje się wyniki dla jakiejś podgrupy badanych – na przykład kobiet, Murzynów, młodzieży – marginesy błędu dla tych wyników jest większy niż w przypadku wyników dla wszystkich badanych łącznie.

12. Jakie inne rodzaje błędów mogły wypaczyć wyniki sondażu? Margines błędu jest tylko jednym ze źródeł niedokładności sondaży – niekoniecznie największym; jego wartość podawana jest tylko dlatego, że tylko ją można określić liczbowo. Sformułowanie pytań i ich kolejność również są potencjalnymi źródłami błędów, dlatego oceniając sondaż, należy przeanalizować także jego pytania pod kątem ewentualnych obciążeń.

Należy zawsze zapytać, czy wyniki sondażu były „ważone". Jest to procedura stosowana zwykle ze względu na nierówne prawdopodobieństwo dostania się do próby poszczególnych jednostek, a także po to, by skorygować strukturę demograficzną próby. Istnieje jednak zawsze ryzyko, że nadmiernie manipulowano ważeniem, aby uzyskać w sondażu pożądane wyniki.

Istnieje wiele innych potencjalnych źródeł błędów, jak słabe przeszkolenie i kontrola ankieterów, błędy w przetwarzaniu danych czy inne problemy operacyjne. Sondaże prowadzone przez zawodowców są znacznie mniej narażone na tego rodzaju błędy, niż sondaże robione przez amatorów, które są na ogół mniej wiarygodne.

13. Jak brzmiały zadawane pytania? Musisz poznać dokładne brzmienie zadawanych w sondażu pytań. Dlaczego? Ponieważ samo sformułowanie pytań może spowodować duże różnice w wynikach.

Chyba najlepszym sposobem oceny pytania jest sprawdzenie, jak się na nie samemu reaguje. Czy na pierwszy rzut oka pytanie wydaje się uczciwe i bezstronne? Czy możliwe odpowiedzi zapewniają zrównoważony wybór? Czy ludzie, których znasz, byliby w stanie na takie pytanie odpowiedzieć?

W przypadku pytań drażliwych – takich jak na przykład pytania dotyczące aborcji – w swojej relacji powinieneś przytoczyć pełne brzmienie pytania. A przynajmniej, przygotowując relację, musisz znać dokładnie to sformułowanie.

Warto porównać kilka różnych sondaży, przeprowadzanych przez różne instytucje, dotyczących tych drażliwych kwestii. W takim przypadku należy starannie porównać zarówno same wyniki, jak i dokładne sformułowanie pytań.

14. W jakiej kolejności zadawano pytania? W pewnych przypadkach na wyniki sondażu może wpływać sama kolejność pytań. Często wpływ ten jest zamierzony, ale czasami nie jest; efekty związane z kolejnością pytań mogą mieć bardzo złożony charakter.

Na przykład, w okresie kryzysu gospodarczego, jeśli najpierw zapytamy, jak respondent ocenia stan gospodarki, a następnie zapytamy o ocenę prezydenta, poziom poparcia dla prezydenta okaże się prawdopodobnie niższy, niż gdybyśmy zadawali pytania w innej kolejności – w okresie prosperity zaś sytuacja byłaby odwrotna.

W sondażach politycznych często zadawanych jest wiele pytań dotyczących stanowiska kandydatów w różnych istotnych kwestiach bądź też dotyczących oceny różnych cech kandydatów. Po tych pytaniach respondent pytany jest o to, na kogo będzie głosował, zwykle już po raz drugi w tym samym wywiadzie. To drugie pytanie o preferencje wyborcze jest później analizowane pod kątem ewentualnego wpływu wywieranego przez poprzednie pytania – może to być całkiem dobra metoda ustalenia istotności pewnych kwestii dla badanych, ale nie można w ten sposób dobrze ocenić rzeczywistego poparcia dla poszczególnych kandydatów.

Istotna jest kwestia, czy na odpowiedzi na pytanie nas interesujące mogła wpłynąć obecność jakiegoś innego wcześniejszego pytania. Na przykład zadawanie pytań dotyczących aborcji przed pytaniem o to, jak respondent zagłosuje w referendum w sprawie aborcji,

może zafałszować odpowiedzi udzielane na drugie pytanie.

15. Jakie inne sondaże przeprowadzano na ten sam temat? Czy ich wyniki były takie same? A jeśli były różne – to dlaczego? Wyniki innych sondaży – zamówionych przez konkurencyjnego kandydata, media, instytucje publiczne czy kogokolwiek innego – powinny zostać skonfrontowane z wynikami sondażu, które właśnie masz w ręku.

Jeśli wyniki się różnią, zacznij od sprawdzenia, kiedy każdy z nich był robiony. Różnica może wynikać ze zmienności opinii publicznej w czasie.

Jeśli sondaże były robione mniej więcej w tym samym czasie i nie znajdujesz żadnych innych czynników uzasadniających różnice, zwróć się o ich wyjaśnienie do autorów wszystkich tych sondaży: sprzeczne wyniki sondaży same w sobie stanowią dobry temat.

16. Z sondażu wynika, że wyniki wyborów są przesądzone – i co teraz? Nieważne, jak dobry był sondaż, nieważne, jak duży margines błędu, jak duża była próba – sondaże przedwyborcze nie przesądzają, że jeden z kandydatów już wygrał. Sytuacja – zwłaszcza w polityce – potrafi się zmieniać często i dramatycznie.

17. Czy sondaż jest elementem akcji zbierania funduszy? Mamy tutaj do czynienia z jeszcze innym przypadkiem pseudosondażu. Jakaś organizacja wysyła kwestionariusz do dużej liczby ludzi. Na ogół w ostatnim pytaniu zawarta jest prośba do respondenta o wpłacenie jakiejś kwoty – jako darowizny na wsparcie tej organizacji albo w charakterze opłaty za przetwarzanie wyników sondażu.

Ludzie, którzy decydują się na udział w takich sondażach, to na ogół ci, którzy organizację tę popierają. Poza tym pytania w takich sondażach są na ogół tendencyjne, a ich wyniki – bezwartościowe.

Praktyki takie stosowane są przez bardzo różnorodne organizacje, od partii politycznych i grup interesów po stowarzyszenia charytatywne. Jeśli sondaż jest elementem akcji zbierania funduszy – wyrzuć go do kosza.

18. No i zadałem wszystkie te pytania. Odpowiedzi brzmią dobrze, więc sondaż chyba jest OK? Na ogół tak. Trzeba jednak pamiętać, że z samego tylko rachunku prawdopodobieństwa wynika, że wyniki jednego na dwadzieścia sondaży mogą być błędne wyłącznie ze względu na błąd z próby.

19. Czy biorąc pod uwagę wszystkie te niebezpieczeństwa, w ogóle powinniśmy publikować wyniki sondaży? Tak. Ponieważ renomowane instytucje badawcze wykonują swoją robotę konsekwentnie dobrze. Mimo trudności, poprawnie prowadzone sondaże pozostają najlepszym obiektywnym miernikiem stanu opinii publicznej.

20. Czy ten sondaż wart jest omówienia? Jeśli sondaż prowadzony był poprawnie i jeśli udało ci się uzyskać odpowiedzi na powyższe pytania, ty i twoi wydawcy możecie to ocenić, opierając się na swoim doświadczeniu dziennikarskim, stosując te same kryteria, co w przypadku każdej innej wiadomości.

Słowniczek

analiza czynnikowa (*factor analysis*) złożona metoda algebraiczna, pozwalająca na wyróżnienie w całym zbiorze obserwacji struktury organizujących je ogólnych czynników (wymiarów). Patrz rozdział 17.

analiza dwuzmiennowa (*bivariate analysis*) jednoczesna analiza dwóch zmiennych, w celu określenia empirycznej relacji między nimi. Może ona polegać na przykład na stworzeniu tabeli krzyżowej albo na obliczeniu współczynnika korelacji. Patrz rozdział 15.

analiza ilościowa (*quantitative analysis*) (1) numeryczne przedstawianie i przetwarzanie danych w celu opisania i wyjaśnienia zjawisk, których dane te dotyczą. Patrz w szczególności rozdział 14, a także inne fragmenty w części 4. (2) DUŻA analiza.

analiza jakościowa (*qualitative analysis*) (1) analiza i interpretacja obserwacji, nie operująca liczbami, nastawiona na odkrywanie podstawowych znaczeń i struktur. Analiza jakościowa znajduje zastosowanie na przykład w badaniach terenowych i badaniach historycznych. Patrz rozdział 10. (2) uzasadnienie odmowy uiszczenia rachunku w restauracji.

analiza jednozmiennowa (*univariate analysis*) analiza jednej zmiennej w celach opisowych. Może polegać na wyznaczeniu rozkładu, średniej, miar rozproszenia. W przypadku analizy większej liczby zmiennych mamy do czynienia z *analizą dwuzmiennową* lub *wielozmiennową*. Patrz rozdział 15.

analiza konwersacyjna (*conversation analysis*) szczegółowa analiza przebiegu rozmowy, oparta na jej dokładnym zapisie uwzględniającym pauzy i inne komunikaty niewerbalne, takie jak „hm" czy wahanie. Patrz rozdział 13.

analiza pytań (*item analysis*) określenie, czy poszczególne pozycje wchodzące w skład złożonej miary wnoszą do niej własny wkład, czy jedynie powielają wkład pozostałych pytań.

analiza regresji (*regression analysis*) metoda analizy danych, w której zależności między zmiennymi przedstawiane są w postaci równania, zwanego równaniem regresji. W rozdziale 17 omówione są różne warianty analizy regresji.

analiza regresji cząstkowej (*partial regression analysis*) typ analizy regresji, w której analizowany jest wpływ jednej zmiennej przy kontroli (utrzymywaniu na stałym poziomie) innych zmiennych, odwołujący się do podobnej logiki jak model analizy rozbudowanej. Patrz rozdział 17.

analiza regresji krzywoliniowej (*curvilinear regression analysis*) typ analizy regresji, w którym związek między zmiennymi może być geometrycznie reprezentowany nie tylko w postaci linii prostej, ale także w postaci krzywej. Patrz rozdział 17.

analiza regresji liniowej (*linear regression analysis*) rodzaj analizy statystycznej, w której wyznacza się równanie prostej najlepiej opisującej relację łączącą dwie zmienne ilorazowe. Patrz rozdział 17.

analiza regresji wielokrotnej (*multiple regression analysis*) typ analizy statystycznej, w którym wyznacza się równanie reprezentujące wpływ dwóch lub więcej zmiennych niezależnych na zmienną zależną. Patrz rozdział 17.

analiza szeregów czasowych (*time-series analysis*) analiza zmian wartości zmiennej (na przykład *poziomu przestępczości*) w czasie. Patrz rozdział 17.

analiza ścieżkowa (*path analysis*) (1) typ analizy wielozmiennowej, którego wynikiem jest graficzna prezentacja związków przyczynowych między zmiennymi. Patrz rozdział 17. (2) to, z czego słyną Indianie i traperzy.

analiza treści (*content analysis*) badanie zapisów międzyludzkiej komunikacji, takich jak książki, strony internetowe, obrazy, przepisy prawa. Patrz rozdział 11.

analiza wielozmiennowa (*multivariate analysis*) jednoczesna analiza relacji między wieloma zmiennymi. Może być to na przykład badanie jednoczesnego wpływu *wieku*, *płci* i *przynależności klasowej* na *religijność*. Patrz rozdziały 15, 16 i 18.

analiza wtórna (*secondary analysis*) typ badań, w których dane zebrane i opracowane przez jednego badacza są analizowane ponownie – często

w innym celu – przez innego. Do wtórnej analizy nadają się zwłaszcza dane sondażowe; istnieją specjalne archiwa i biblioteki, udostępniające dane do analizy wtórnej.

analiza zorientowana na przypadek (*case-oriented analysis*) (1) analiza, której celem jest zrozumienie pojedynczego przypadku lub kilku przypadków, poprzez dokładne zbadanie wszystkich związanych z nimi szczegółów. Patrz rozdział 13. (2) metodyka pracy prywatnych detektywów.

analiza zorientowana na zmienną (*variable-oriented analysis*) analiza opisująca i/lub wyjaśniająca konkretną zmienną. Patrz rozdział 13.

anonimowość (*anonymity*) anonimowość badań zapewniona jest wtedy, gdy ani badacze, ani odbiorcy wyników badania nie mogą powiązać konkretnych odpowiedzi z konkretnym respondentem. Patrz rozdział 18.

arkusz kodowy (*transfer sheet*) duży arkusz papieru z wyznaczonymi kolumnami i wierszami, na który wpisuje się dane w celu późniejszego wprowadzenia ich do komputera i analizy. Patrz rozdział 14.

badania dynamiczne (*longitudinal study*) schemat badań, polegający na zbieraniu danych w różnych punktach czasowych, inaczej niż w przypadku *badań przekrojowych*. Patrz rozdział 4 oraz *badania kohort*, *badania panelowe*, *badania trendów*.

badania ewaluacyjne (*evaluation research*) badania, których celem jest ocena oddziaływania danej interwencji społecznej (na przykład programu, mającego rozwiązać jakiś problem społeczny). Patrz rozdział 12.

badania historyczno-porównawcze (*historical/comparative research*) badania społeczeństw (lub innych jednostek społecznych) na przestrzeni czasu i porównywanie ich między sobą. Patrz rozdział 11.

badania kohort (demograficznych) (*cohort study*) badania, w których pewna specyficzna podpopulacja (kohorta demograficzna) badana jest przez dłuższy czas (z tym że w poszczególnych punktach czasowych badaniu mogą być poddawani różni jej członkowie). Przykładem może być tutaj badanie historii kariery zawodowej rocznika 1970, w którym kwestionariusze rozsyłane były co 5 lat. Więcej na ten temat (jeśli *naprawdę* chcesz wiedzieć więcej) znajdziesz w rozdziale 4. Patrz także *badania dynamiczne*, *badania panelowe*, *badania trendów*.

badania panelowe (*panel study*) typ badań dynamicznych, w których dane zbierane są od tej samej grupy ludzi (próby czy inaczej panelu) w kilku punktach czasowych. Patrz rozdział 4 oraz *badania kohort*, *badania dynamiczne*, *badania trendów*.

badania przekrojowe (*cross-sectional study*) badania oparte na obserwacjach poczynionych w jednym punkcie czasowym. Porównaj z *badaniami dynamicznymi*.

badania trendów (*trend study*) (1) typ badań dynamicznych, w których dana cecha populacji monitorowana jest w dłuższym czasie, na przykład seria sondaży przedwyborczych Gallupa, ukazujących poparcie wyborców dla poszczególnych kandydatów, zmieniające się w trakcie kampanii wyborczej (nie jest przy tym konieczne, aby za każdym razem sondaż odbywał się na tej samej próbie). Patrz rozdział 4 oraz *badania kohort*, *badania panelowe*, *badania dynamiczne*. (2) stałe badania, zlecane przez producentów odzieży.

badania uczestniczące (*participatory action research*) podejście do badań społecznych, w którym osoby badane są dopuszczone do kontroli nad celami i środkami badania, pomyślane jako przeciwwaga dla przyjmowanego *implicite* poglądu, że badacz stoi ponad badanymi. Patrz rozdział 10.

błąd ekologiczny rozumowania (*ecological fallacy*) błędne wyciąganie wniosków na temat jednostek jedynie na podstawie obserwacji grup. Patrz rozdział 4.

błąd systematyczny (*nonsampling error*) (1) niedoskonałość danych wynikająca z przyczyn innych niż błąd z próby – na przykład konsekwencji niezrozumienia przez respondentów pytań kwestionariusza, błędnych zapisów, dokonanych przez ankieterów lub koderów, błędów przy wprowadzaniu danych. Patrz rozdział 17. (2) to, co robisz, jeśli zawsze piszesz „drzwi" przez „ż".

błąd z próby (błąd losowy) (*sampling error*) wielkość błędu, oczekiwana w doborze losowym. Wzór, na podstawie którego określa się błąd z próby, uwzględnia trzy czynniki: parametr, wielkość próby oraz błąd standardowy. Patrz rozdział 7.

czyszczenie danych (*data cleaning*) sprawdzanie skwantyfikowanych danych pod kątem wszelkich możliwych błędów. Patrz rozdział 14, *czyszczenie dotyczące dopuszczalnych kodów*, *czyszczenie logiczne*.

czyszczenie dotyczące dopuszczalnych kodów (kontrola zakresów kodowania) (*possible-code cleaning*) sprawdzanie skwantyfikowanych danych pod kątem tego, czy kody przypisane wartościom poszczególnych zmiennych mieszczą się w dopuszczalnym zakresie kodów, stworzonym dla każdej z tych zmiennych. Na przykład jeśli stwierdzimy, że wartość zmiennej, mającej tylko dwa atrybuty, dla jakiegoś przypadku określona jest jako „7", mamy prawdopodobnie do czynienia z błędem. Patrz rozdział 14.

czyszczenie logiczne (kontrola spójności danych) (*contingency cleaning*) sprawdzanie skwantyfikowanych danych pod kątem ewentualnych niespójności logicznych – na przykład sprawdzanie, czy w zbiorze danych nie pojawiają się mężczyźni w ciąży. Patrz rozdział 14.

debriefing (sesja poeksperymentalna) (*debriefing*) rozmowa z badanymi dotycząca ich osobistych doświadczeń, związanych z udziałem w badaniach; szczególnie ważna w przypadkach, gdy uczestnictwo w badaniach przyniosło im jakieś szkody. Patrz rozdział 18.

dedukcja (*deduction*) model logicznego rozumowania, w którym oczekiwania i hipotezy wywodzone są z ogólnych twierdzeń. Wychodząc od twierdzenia, że wszyscy dziekani to świnie, możesz sformułować przewidywanie, że twój nie zgodzi się, żebyś wprowadził zmiany do swojego programu studiów – przewidywanie to zostanie sformułowane na drodze dedukcji. Patrz *indukcja* oraz rozdziały 2 i 3.

definicja operacyjna (*operational definition*) konkretna i szczegółowa definicja jakiegoś pojęcia, sformułowana w kategoriach działań (operacji), pozwalających na zaklasyfikowanie danej obserwacji. Definicją operacyjną „uzyskania z egzaminu oceny bardzo dobrej" może być „udzielenie prawidłowych odpowiedzi na co najmniej 90% pytań". Patrz rozdział 5.

dobór kwotowy (*quota sampling*) typ próby nielosowej, w której jednostki dobierane są na podstawie wcześniej ustalonych cech, w ten sposób, aby rozkład tych cech w całej próbie był taki, jak w badanej populacji. Patrz rozdział 7.

dobór losowy prosty (*simple random sampling*) (1) typ losowego doboru próby, w którym elementom populacji przyporządkowuje się liczby, następnie generuje się zestaw liczb losowych i włącza się do próby te elementy populacji, które mają wylosowane numery. Chociaż rachunek prawdopodobieństwa i stosowane przezeń obliczenia zakładają użycie takiej właśnie próby, jest ona ze względów praktycznych rzadko stosowana. Takie same właściwości jak prosta próba losowa ma próba systematyczna (z losowo dobranym punktem startowym). Patrz rozdział 7. (2) przeciwieństwo doboru losowego krzywego.

dobór nieprobabilistyczny (nielosowy) (*nonprobability sampling*) metody doboru próby za pomocą technik nie odwołujących się do zasad rachunku prawdopodobieństwa, na przykład dobór na podstawie dostępności badanych, dobór celowy (arbitralny), kwotowy, metoda kuli śnieżnej. Patrz rozdział 7.

dobór probabilistyczny (losowy) (*probability sampling*) ogólne określenie dla prób dobranych zgodnie z zasadami rachunku prawdopodobieństwa, z reguły przy użyciu jakiejś procedury losowania, na przykład dobór losowy prosty, EPSEM, PPS, dobór systematyczny. Patrz rozdział 7.

dobór proporcjonalny PPS (*probability proportional to size, PPS*) rodzaj próby grupowej wielostopniowej, w której grupy nie są dobierane z równym prawdopodobieństwem (jak w EPSEM, czyli metodzie równego prawdopodobieństwa wyboru), ale z prawdopodobieństwem proporcjonalnym do ich wielkości, mierzonej liczbą jednostek, jakie mają być dobrane w ich obrębie. Patrz rozdział 7.

dopytywanie (*probe*) zwane także *sondowaniem* – technika stosowana w wywiadach, służąca uzyskaniu pełniejszej odpowiedzi na jakieś pytanie, wykorzystująca pośrednie pytania i zwroty, mające zachęcić respondenta do rozbudowania swojej wypowiedzi, jak „czy jeszcze coś?", „w jaki sposób?". Patrz omówienie wywiadu w rozdziale 9.

dyferencjał semantyczny (*semantic differential*) forma pytania kwestionariuszowego, w którym respondent proszony jest o dokonanie oceny czegoś na skali kończącej się dwoma przeciwstawnymi przymiotnikami (na przykład w przypadku oceny podręczników: „nudny" i „ciekawy"), której pośrednie stopnie wyrażone są sformułowaniami, takimi jak: „bardzo", „raczej", „ani taki, ani taki", „raczej", „bardzo", mającymi objąć całą przestrzeń między dwoma przeciwieństwami. Patrz rozdział 6.

eksperyment podwójnie ślepy (*double-blind experiment*) schemat eksperymentalny, w którym ani badani, ani eksperymentatorzy nie wiedzą, która grupa jest grupą eksperymentalną, a która kontrolną. Patrz rozdział 8.

eksplanacja (*explanation*) wynik modelu analizy rozbudowanej, stwierdzający, że początkowo zaobserwowana zależność między dwiema zmiennymi ma charakter pozorny, ponieważ znika ona po wprowadzeniu zmiennej kontrolnej.

element (*element*) pojedynczy obiekt należący do populacji, będący przedmiotem losowania w procedurze doboru próby. Pojęcie to należy odróżnić od pojęcia *jednostki analizy*, stosowanego w analizie danych. Patrz rozdział 7.

etnografia (*ethnography*) badania życia społecznego koncentrujące się raczej na szczegółowym i precyzyjnym opisie, niż na wyjaśnianiu obserwowanych zjawisk. Patrz rozdział 10.

etnografia instytucjonalna (*institutional ethnography*) technika badawcza, w której osobiste doświadczenia jednostek służą do odkrycia relacji

władzy oraz innych cech instytucji, w których jednostki działają. Patrz rozdział 10.

etnometodologia (*ethnomethodology*) podejście badawcze w naukach społecznych, koncentrujące się na poszukiwaniu ukrytych, z reguły nie formułowanych bezpośrednio założeń i konwencji; często stosowaną metodą jest celowe łamanie takich konwencji, aby wykazać ich istnienie. Patrz rozdział 10.

grupa eksperymentalna (*experimental group*) grupa, której w toku eksperymentu aplikowany jest bodziec eksperymentalny. Porównaj z *grupą kontrolną*. Patrz rozdział 8.

grupa kontrolna (*control group*) (1) w eksperymentach grupa badanych, którym nie został zaaplikowany bodziec eksperymentalny, nie różniąca się jednak pod żadnym innym względem od grupy eksperymentalnej. Porównanie pod koniec eksperymentu grupy eksperymentalnej z grupą kontrolną pozwala określić wpływ bodźca. Patrz rozdział 8. (2) Amerykańskie Stowarzyszenie Menedżerów.

hipoteza (*hypothesis*) określone, weryfikowalne oczekiwanie co do rzeczywistości, sformułowane na bazie ogólniejszego twierdzenia; szerzej – przewidywanie dotyczące natury rzeczy wyprowadzone z teorii. Stwierdzenie, że jeśli teoria jest prawdziwa, to powinno być zaobserwowane określone zjawisko. Patrz *dedukcja* oraz rozdziały 2 i 4.

hipoteza formułowana *post factum* (*ex post facto hypothesis*) hipoteza sformułowana po zebraniu obserwacji, które ją potwierdzają. Hipotezy takie jako niefalsyfikowalne są bezwartościowe. Patrz rozdział 16.

hipoteza zerowa (*null hypothesis*) w kontekście weryfikacji hipotez statystycznych i testów istotności, hipoteza, według której między analizowanymi zmiennymi nie zachodzi żaden związek. Istnienie takiego związku można uznać za potwierdzone, jeśli hipotezę zerową uda się odrzucić. Patrz rozdział 17.

idiograficzny (*idiographic*) tryb wyjaśniania, w którym dąży się do wyczerpującego przedstawienia specyficznych przyczyn konkretnego stanu lub zdarzenia. Wyobraź sobie, że starasz się wymienić wszystkie powody, dla których zdecydowałeś się studiować na tej, a nie innej uczelni. Wziąwszy pod uwagę wszystkie te powody, trudno wyobrazić sobie, że miałbyś podjąć inną decyzję. Patrz także: *nomotetyczny*.

indeks (*index*) rodzaj złożonego miernika, który syntetyzuje i porządkuje wyniki pewnej liczby szczegółowych obserwacji, jakiś ogólniejszy wymiar; porównaj ze *skalą*. Patrz rozdział 6.

indukcja (*induction*) logiczny model rozumowania, w którym ogólne twierdzenia wywodzone są ze szczegółowych obserwacji. Na przykład z obserwacji, że żydzi i katolicy są skłonni do głosowania na demokratów bardziej niż protestanci, w drodze indukcji możemy wysnuć wniosek, że w USA mniejszości religijne identyfikują się raczej z partią demokratyczną oraz wyjaśnić, dlaczego. Patrz *dedukcja* oraz rozdziały 2 i 3.

informator (*informant*) osoba dobrze obeznana ze zjawiskiem społecznym, które chcemy badać, i gotowa podzielić się z nami swoją wiedzą. Na przykład jeśli chcemy prowadzić obserwację uczestniczącą w sekcie religijnej, dobrze jest zaprzyjaźnić się z kimś, kto ją dobrze zna – na przykład z jej członkiem – i kto mógłby udzielić nam podstawowych informacji. Informatora nie należy mylić z *respondentem*.

interpretacja (*interpretation*) techniczne pojęcie, stosowane w związku z modelem analizy rozbudowanej. Oznacza sytuację, w której okazuje się, że zmienna niezależna wpływa na zależną za pośrednictwem zmiennej kontrolnej.

interwał (odstęp) losowania (*sampling interval*) w doborze systematycznym stała odległość między poszczególnymi elementami dobieranymi z populacji do próby. Patrz rozdział 7.

istotność statystyczna (*statistical significance*) (1) ogólne pojęcie, odnoszące się do prawdopodobieństwa, że zależności obserwowane w próbie są wyłącznie wynikiem błędu z próby. Patrz testy *istotności statystycznej* oraz rozdział 17. (2) na ile rzeczywiście jest ważne, czy zdasz egzamin ze statystyki – w końcu zawsze możesz zostać poetą.

jednostka losowania (*sampling unit*) jednostka lub grupa jednostek, która może być wylosowana na jakimś etapie procedury losowania próby.

jednostki analizy (*units of analysis*) to, co (kto) będzie badane(y). W naukach społecznych najbardziej typową jednostką analizy są pojedynczy ludzie. Patrz rozdział 4.

kodowanie (*coding*) (1) proces nadawania danym surowym standaryzowanej formy, pozwalającej na komputerowe przetwarzanie i analizę. (2) to, do czego szpiedzy używają książki kodowej.

kodowanie brzegowe (*edge-coding*) w przypadku kodowania danych z kwestionariuszy lub podobnych narzędzi – wpisywanie kodów na marginesach oryginalnych dokumentów, a nie na osobnym arkuszu. Patrz rozdział 14.

kodowanie otwarte (*open coding*) wstępna klasyfikacja i nazwanie pojęć w jakościowej analizie danych, przy czym badacze przypisują kody na

podstawie bezpośredniego wglądu w dane oraz ich dokładnej analizy. Patrz rozdział 13.

konceptualizacja (*conceptualization*) (1) proces myślowy, w którym nieostre, nieścisłe pojęcia (koncepty) są precyzowane i ściśle definiowane. No więc chciałbyś badać uprzedzenia – ale co to jest „uprzedzenie"? Czy są różne rodzaje uprzedzeń? Jakie? Rozdział 5 poświęcony jest w całości konceptualizacji i jej koleżance, operacjonalizacji. (2) sposób rozmnażania płciowego intelektualistów.

korelacja (*correlation*) (1) empiryczny związek, zaobserwowany między dwiema zmiennymi, polegający na tym, że: (a) zmiany wartości jednej zmiennej są powiązane ze zmianami drugiej lub (b) przyjmowanie określonych wartości przez jedną zmienną jest związane z przyjmowaniem określonych wartości przez drugą. Przykładowo, mówimy, że *wykształcenie* jest skorelowane z *dochodem*, ponieważ wyższy poziom dochodu powiązany jest z wyższym poziomem wykształcenia. Korelacja sama w sobie nie przesądza o istnieniu związku przyczynowego między dwiema zmiennymi, ale jest jednym z jego warunków. (2) ktoś, do kogo jednocześnie ty i twój przyjaciel jesteście w relacji.

książka kodowa (*codebook*) (1) dokument stosowany w czasie przetwarzania i analizy danych, zawierający informacje o lokalizacji poszczególnych danych w zbiorze. Z reguły książka kodowa określa położenie konkretnych danych oraz znaczenie kodów, użytych dla oznaczenia różnych wartości zmiennych. Więcej szczegółów i przykładów w rozdziale 14. (2) ściśle tajny dokument umożliwiający kontakty szpiegów z centralą; nie wiadomo, czemu dostępny na każdej poczcie.

kwestionariusz (*questionnaire*) dokument zawierający pytania i inne narzędzia służące do zebrania informacji, które mają zostać przeanalizowane. Kwestionariusze są stosowane przede wszystkim w badaniach sondażowych, ale także w eksperymentach, badaniach terenowych i innych typach obserwacji. Patrz rozdział 9.

losowanie (*random selection*) metoda doboru próby, w której każda jednostka ma takie same szanse na znalezienie się w próbie, niezależnie od innych zdarzeń występujących w procesie doboru. Patrz rozdział 7.

losowanie systematyczne (*systematic sampling*) typ doboru losowego, w którym do próby dobiera się co *k*-tą pozycję z listy – na przykład co dwudziestego piątego studenta z listy studentów danego wydziału. Wartość *k* (interwał losowania) wyznacza się, dzieląc wielkość populacji przez docelową wielkość próby. Z pewnymi zastrzeżeniami, próba systematyczna jest funkcjonalnie równoważna próbie losowej prostej, a często łatwiej ją dobrać. Pierwszy element próby wyznaczany jest z reguły losowo. Patrz rozdział 7.

makroteoria (*macrotheory*) teoria, której celem jest przedstawienie całościowego obrazu instytucji, społeczeństw czy relacji między zbiorowościami. Teoria „walki klas" Marksa jest przykładem makroteorii. Porównaj z *mikroteorią*.

mapa pojęciowa (*concept mapping*) ujęcie graficzne pojęć i ich wzajemnych relacji, pomocne przy formułowaniu teorii. Patrz rozdział 13.

mediana (*median*) przeciętna, równa wartości „środkowego" przypadku z porangowanego zbioru obserwacji. Jeśli mamy pięciu ludzi w wieku odpowiednio 16, 17, 20, 54 i 88 lat, to mediana ich wieku wynosi 20 (a średnia 39).

metoda kuli śnieżnej (*snowball sampling*) nielosowa metoda doboru próby stosowana często w badaniach terenowych, w której każdy badany może zostać poproszony o wskazanie kolejnych osób, które mogłyby wziąć udział w badaniu. Patrz rozdziały 7 i 10.

metoda rozszerzonych przypadków (*extended case method*) technika wykorzystująca obserwacje, uzyskane w studium przypadku, do odkrywania słabych punktów i udoskonalania istniejących teorii społecznych, rozwinięta przez Michaela Burawoya. Patrz rozdział 10.

metoda równego prawdopodobieństwa wyboru EPSEM (*EPSEM – equal probability of selection method*) schemat doboru próby, w którym każdy element populacji ma jednakowe prawdopodobieństwo bycia wylosowanym. Patrz rozdział 7.

metoda stałego porównywania (*constant comparative method*) (1) element metody teorii ugruntowanej – porównywanie poszczególnych obserwacji między sobą oraz z budowaną indukcyjnie teorią. Patrz rozdział 13. (2) technika stosowana przez uczestników randek w ciemno.

metoda teorii ugruntowanej (*grounded theory method, GTM*) indukcyjne podejście do badań, wprowadzone przez Barneya Glasera i Anselma Straussa, w którym teorie są tworzone raczej w wyniku samej analizy danych niż wyprowadzane dedukcyjnie. Patrz rozdział 13.

mikroteoria (*microtheory*) teoria mająca wyjaśnić zjawiska społeczne, zachodzące na poziomie jednostek i relacji między jednostkami. Teoria, dotycząca różnic w zachowaniach zabawowych chłopców i dziewczynek, byłaby mikroteorią. Porównaj z *makroteorią*.

modalna (*mode*) (1) przeciętna, odpowiadająca najczęściej obserwowanej wartości zmiennej. Jeśli w próbie jest 1000 protestantów, 275 katolików

i 33 żydów, wartością modalną jest „protestant". W rozdziale 15 znajdziesz więcej szokujących informacji o miarach tendencji centralnej. (2) pracownica odpowiedzialna za krój strojów roboczych.

model analizy rozbudowanej (*elaboration model*) technika badania zależności między dwiema zmiennymi poprzez kontrolowanie wpływu trzeciej, opracowana przez Paula Lazarsfelda. Efektem analizy rozbudowanej może być stwierdzenie *replikacji*, *specyfikacja*, *eksplanacja* lub *interpretacja*.

naturalizm (*naturalism*) podejście do badań terenowych, oparte na założeniu, że obiektywna rzeczywistość społeczna istnieje i daje się obserwować oraz trafnie opisywać. Patrz rozdział 10.

nierównoważna grupa kontrolna (*nonequivalent control group*) grupa kontrolna podobna do grupy eksperymentalnej, ale wyłoniona w sposób inny niż losowy. Taki typ grupy kontrolnej różni się istotnie od grupy eksperymentalnej ze względu na zmienną zależną lub zmienne z nią powiązane. Patrz rozdział 12.

nietrafność wewnętrzna (*internal invalidity*) niebezpieczeństwo, że wnioski, wyciągnięte z wyników eksperymentu, nie są odzwierciedleniem samego eksperymentu. Patrz rozdział 8 oraz *nietrafność zewnętrzna*.

nietrafność zewnętrzna (*external invalidity*) niebezpieczeństwo, że wnioski wyciągnięte z wyników eksperymentu mogą nie nadawać się do uogólniania w „realnym świecie". Patrz rozdział 8 oraz *nietrafność wewnętrzna*.

nomotetyczny (*nomothetic*) tryb wyjaśniania, zmierzający do ustalenia kilku czynników o charakterze przyczynowym, które na ogół wpływają na pewną klasę zdarzeń lub stanów. Pomyśl o dwóch lub trzech kluczowych czynnikach determinujących wybór uczelni przez studentów, takich jak odległość od miejsca zamieszkania, reputacja itd. Patrz także: *idiograficzny*.

notatki, robienie notatek (*memos, memoing*) zapisywanie notatek, traktowanych jako część danych do późniejszej analizy w badaniach jakościowych, takich jak teoria ugruntowana. Zapiski mogą opisywać lub definiować pojęcia, omawiać kwestie metodologiczne bądź też proponować wstępne sformułowanie elementów teorii.

obciążenie (*bias*) (1) właściwość narzędzia pomiarowego, powodująca przekłamywanie w określonym kierunku wyników pomiaru. Na przykład pytanie „Czy zgodzisz się z opinią, że prezydent dobrze sprawuje swój urząd?" będzie obciążone, ponieważ skłania respondentów do wyrażenia pozytywnej opinii. Patrz rozdział 9. (2) coś wewnątrz ciebie, co powoduje, że innych ludzi lub grupy konsekwentnie postrzegasz jako lepszych lub gorszych, niż są w rzeczywistości. (3) coś, co poczujesz, gdy ktoś ci postawi na głowie kowadło (nie zadawaj się z idiotami).

odchylenie standardowe (*standard deviation*) jedna z miar dyspersji (rozproszenia wyników wokół średniej). Dla zmiennej o rozkładzie normalnym około 68% przypadków charakteryzuje się wartością zmiennej, różniącą się od średniej o nie więcej niż jedno odchylenie standardowe, 95% o nie więcej niż dwa, a 99,9% o nie więcej niż trzy odchylenia standardowe. Tak więc w grupie, w której wiek ma rozkład normalny o średniej 30 i odchyleniu standardowym 10 lat, 68% osób ma od 20 do 40 lat. Im mniejsze jest odchylenie standardowe, tym bardziej wartości zmiennej są skupione wokół średniej; jeśli odchylenie standardowe jest duże, wartości rozrzucone są szerzej. Patrz rozdział 15.

operacjonalizacja (*operationalization*) (1) krok dalej w stosunku do konceptualizacji – proces tworzenia definicji operacyjnych czy inaczej określania dokładnych procedur pomiaru wartości zmiennej. (2) zabieg chirurgiczny na intelektualiście.

operat losowania (*sampling frame*) lista, lub quasi-lista jednostek tworzących populację, z której losowana jest próba. Jeśli chcemy wylosować próbę reprezentatywną dla całej populacji, musimy zadbać o to, by operat losowania zawierał wszystkie (lub prawie wszystkie) jej elementy. Patrz rozdział 7.

paradygmat (*paradigm*) model czy schemat służący obserwacji i interpretacji. Kształtuje on zarówno nasze spostrzeżenia, jak i to, jak je rozumiemy. Paradygmat konfliktu sprawia, że postrzegamy zachowania społeczne w jeden sposób, a paradygmat interakcjonistyczny – że w inny sposób.

parametr (*parameter*) syntetyczny opis wartości zmiennej w populacji. Patrz rozdział 7.

pomiar ilorazowy (*ratio measure*) poziom pomiaru zmiennej, mający wszystkie właściwości pomiaru nominalnego, porządkowego i przedziałowego, w którym dodatkowo na skali pomiarowej można wyróżnić obiektywny punkt zerowy. Przykładem zmiennej mierzonej na skali ilorazowej może być wiek. Patrz rozdział 5 oraz *pomiar nominalny*, *pomiar porządkowy*, *pomiar interwałowy* (*przedziałowy*).

pomiar interwałowy (przedziałowy) (*interval measure*) poziom pomiaru zmiennej, przy którym wartości zmiennej można uszeregować, a odległości pomiędzy sąsiadującymi wartościami zmiennej są równe. Pomiar temperatury na skali Fahrenheita jest pomiarem interwałowym, ponieważ

odległość między 17 i 18 stopniami jest taka sama jak między 89 a 90. Patrz rozdział 5 oraz *pomiar nominalny, pomiar porządkowy, pomiar ilorazowy*.

pomiar nominalny (*nominal measure*) poziom pomiaru, którego jedyną własnością jest to, że wartości zmiennej są wyczerpujące i wzajemnie rozłączne. Innymi słowy, jeśli zmienna mierzona jest na skali nominalnej, to o jej wartościach możemy orzec jedynie, że są różne – inaczej niż w przypadku *pomiaru porządkowego, interwałowego* czy *ilorazowego*. Przykładem zmiennej nominalnej może być *płeć*. Patrz rozdział 5.

pomiar porządkowy (rangowy) (*ordinal measure*) poziom pomiaru zmiennej, który umożliwia uszeregowanie wartości zmiennej według jakiegoś kryterium. Przykładem zmiennej mierzonej na skali porządkowej może być status społeczno-ekonomiczny z możliwymi wartościami: *niski, średni, wysoki*. Patrz rozdział 5 oraz *pomiar nominalny, pomiar interwałowy, pomiar ilorazowy*.

populacja (*population*) określony teoretycznie zbiór obiektów będących przedmiotem badania. Patrz rozdział 7.

populacja badana (*study population*) zbiór elementów, z którego faktycznie dobierana jest próba. Patrz rozdział 7.

posttest (*posttesting*) ponowny pomiar wartości zmiennej zależnej po tym, jak badani zostali poddani działaniu zmiennej niezależnej. Patrz rozdział 8.

poufność (*confidentiality*) projekt badawczy zapewnia badanym poufność, jeśli badacz może wprawdzie zidentyfikować, który z badanych udzielił danej odpowiedzi, ale zobowiązuje się, że tej informacji nie ujawni. Patrz rozdział 18.

poziom istotności (*level of significance*) (1) w kontekście testów istotności statystycznej prawdopodobieństwo, że obserwowana zależność empiryczna jest wynikiem jedynie błędu z próby. Zależność jest istotna na poziomie 0,05, jeżeli prawdopodobieństwo, że jest ona skutkiem jedynie błędu z próby, jest nie większe niż 5 na 100. Patrz rozdział 17. (2) maksymalna wysokość, na której warto wieszać plakaty.

poziom realizacji próby (poziom zwrotów) (*response rate, completion rate*) wyrażona procentowo liczba osób, które wzięły udział w badaniu sondażowym, w stosunku do liczby osób, które zostały wylosowane do próby. W sondażach realizowanych za pomocą ankiet wypełnianych samodzielnie przez respondentów używa się także określenia „poziom zwrotów" (*return rate*). Patrz rozdział 9.

poziom ufności (*confidence level*) (1) miara prawdopodobieństwa, że parametr w populacji należy do wyznaczonego przedziału ufności. Tak więc możemy być na 95% pewni, że procent wyborców popierających kandydata A należy do przedziału 35–45%. Patrz rozdziały 7 i 17. (2) stopień, w jakim ufasz, że w kupionym od faceta na ulicy za 10 dolarów pierścionku rzeczywiście jest trzykaratowy diament.

pretest (*pretesting*) pomiar wartości zmiennej zależnej. Patrz rozdział 8.

proporcja losowania (*sampling ratio*) stosunek wielkości próby do wielkości populacji. Patrz rozdział 7.

proporcjonalna redukcja błędu (*proportional reduction of error, PRE*) logiczny model oceny siły związku, oparty na badaniu, o ile można zredukować błąd przewidywania wartości jednej zmiennej dzięki znajomości wartości drugiej zmiennej. Na przykład znajomość wykształcenia badanych ułatwia przewidywanie, ile wynoszą ich zarobki – co pokazuje, że zmienne te są ze sobą związane. Patrz rozdział 17.

próba celowa (arbitralna) (*purposive sampling, judgmental sampling*) rodzaj próby nielosowej, do której dobiera się jednostki, które według subiektywnej oceny badacza mają być najbardziej użyteczne lub reprezentatywne. Patrz rozdział 7.

próba grupowa (gronowa, zespołowa) (*cluster sampling*) wielostopniowa próba, w której w pierwszym etapie losowane są pewne naturalne grupy („grona"), w których następnie losuje się poszczególne obiekty. Na przykład najpierw losuje się próbę wyższych uczelni, a następnie z listy studentów każdej z wylosowanych szkół losuje się konkretne osoby. Patrz rozdział 7.

przeciętna wartość (*average*) nieprecyzyjny termin, sugerujący wartość typową lub normalną; miara tendencji centralnej. Przykładami sprecyzowanych matematycznie wartości przeciętnych są średnia, mediana i modalna. Patrz rozdział 15.

przedział ufności (*confidence interval*) (1) przedział wartości, w którym leży szacowany parametr populacji. Przykładowo – wyniki sondażu wskazują, że na kandydata A chce głosować 40% próby. Chociaż najlepszą estymacją wielkości poparcia wśród wszystkich wyborców również jest 40%, trudno się spodziewać, żeby było to dokładnie tyle. Możemy jednak wyznaczyć przedział ufności (na przykład 35–45%), w którym odsetek ten prawdopodobnie się mieści. Zauważ, że dla każdego przedziału ufności zawsze musimy podać także poziom ufności. Patrz rozdziały 7 i 17. (2) odległość, na jaką odważymy się zbliżyć do aligatora.

przekrojowa analiza przypadków (*cross-case analysis*) analiza wiążąca się z badaniem więcej niż jednego przypadku. Może to być analiza *zorientowana na zmienną* lub *zorientowana na przypadek*. Patrz rozdział 13.

pytania otwarte (*open-ended questions*) pytania, na które respondent może odpowiedzieć własnymi słowami. Pogłębione wywiady jakościowe opierają się prawie wyłącznie na pytaniach otwartych. Patrz rozdziały 9 i 10.

pytania warunkowe (*contingency questions*) pytania zadawane w badaniu sondażowym jedynie tym respondentom, którzy udzielili określonych odpowiedzi na inne pytania. Na przykład badani mogliby najpierw zostać zapytani, czy należą do mafii, po czym tylko tych, którzy odpowiedzieliby twierdząco, zapytano by, jak często biorą udział w spotkaniach mafijnych. To drugie pytanie byłoby właśnie pytaniem warunkowym. Więcej przykładów w rozdziale 9.

pytania zamknięte (*closed-ended questions*) pytania zadawane w sondażu, na które respondent odpowiada wybierając jeden z wariantów odpowiedzi proponowanych przez badacza. Często stosowane, ze względu na większą jednorodność odpowiedzi i łatwość obróbki wyników. Patrz rozdział 9.

quasi-eksperyment (*quasi-experiment*) metoda badawcza przypominająca nieco eksperyment kontrolowany, ale znacznie słabiej sformalizowana, nie uwzględniająca jego kluczowych elementów, takich jak pretest i posttest czy grupa kontrolna. Patrz rozdział 12.

randomizacja (*randomization*) technika losowego przypisywania badanych do grup eksperymentalnej i kontrolnej. Patrz rozdział 8.

redukcjonizm (*reductionism*) (1) błąd niektórych badaczy polegający na ścisłym ograniczaniu (redukcji) rodzajów pojęć uznawanych za adekwatne do danego zjawiska. (2) kłopoty z dykcją powracające przy każdej kolejnej sesji podczas egzaminów ustnych, ograniczające wypowiedź do tzw. dukania.

replikacja (*replication*) ogólnie, powtórzenie eksperymentu, aby wychwycić lub zmniejszyć błąd. W kontekście modelu analizy rozbudowanej jest to techniczny termin używany do określenia sytuacji, gdy zależność początkowa między dwiema zmiennymi utrzymuje się także, gdy jest kontrolowana ze względu na trzecią zmienną – co stanowi argument za rzeczywistym charakterem tej zależności. Patrz rozdział 16.

reprezentatywność (*representativeness*) (1) właściwość próby, polegająca na tym, że uzyskane w niej rozkłady pewnych cech są takie same jak w całej populacji. Wnioskuje się z tego, że opisy i wyjaśnienia, konstruowane na bazie analizy danych z próby, można odnieść także do populacji. Reprezentatywność można osiągnąć poprzez stosowanie prób losowych; uprawnia ona do uogólniania wyników oraz stosowania metod wnioskowania statystycznego. Patrz rozdział 7. (2) to, czym we własnym mniemaniu charakteryzują się niektórzy członkowie Kongresu.

respondent (*respondent*) osoba, której odpowiedzi na pytania w badaniach sondażowych poddawane są analizie. Patrz rozdział 9.

rozkład częstości (*frequency distribution*) (1) opis, w jakiej części obserwacji z próby zmienna przyjmuje określone wartości. Stwierdzenie, że w próbie 53% stanowią mężczyźni, a 47% kobiety, jest przykładem rozkładu częstości dla zmiennej *płeć*. Innym przykładem jest stwierdzenie, że spośród badanych miast 15 miało mniej niż 10 tysięcy mieszkańców, 23 między 10 a 25 tysięcy itd. Patrz rozdział 15. (2) panel radioodbiornika.

rozproszenie (*dispersion*), inaczej *dyspersja*. Rozkład wartości wokół jakiejś wartości centralnej, na przykład średniej. Najprostszą miarą rozproszenia jest zakres – możemy stwierdzić, że w badanej grupie średnia wieku wynosi 37,9, przy zakresie od 12 do 89. Patrz rozdział 15.

rzetelność (*reliability*) właściwość metody pomiaru, polegająca na tym, że przy każdym powtórzeniu obserwacji tego samego zjawiska uzyskalibyśmy taki sam wynik. Na przykład w przypadku sondażu możemy się spodziewać, że pytanie „Czy był(a) Pan(i) w kościele w ostatnim tygodniu?" charakteryzowałoby się większą rzetelnością niż pytanie „Ile razy w życiu był(a) Pan(i) w kościele?". Rzetelności nie należy mylić z *trafnością*. Patrz rozdział 5.

schemat z uwzględnieniem szeregów czasowych (*time-series design*) plan badawczy zakładający dokonywanie pomiarów w dłuższym okresie (na przykład badanie liczby wypadków drogowych przed wprowadzeniem ograniczenia prędkości i po jego wprowadzeniu). Patrz rozdział 12.

schemat z uwzględnieniem wielu szeregów czasowych (*multiple time series design*) wykorzystanie w celach porównawczych większej liczby zbiorów danych, które były gromadzone w czasie, na przykład poziomu wypadków w różnych miastach w różnych latach. Patrz rozdział 12.

semiotyka (*semiotics*) nauka o znakach i ich znaczeniach, powiązana z analizą treści. Patrz rozdział 13.

skala (*scale*) typ złożonej miary, której elementy tworzą logiczną lub empiryczną strukturę. Przykładami

skal są skala dystansu społecznego Bogardusa, skale Guttmana, Likerta i Thurstone'a. Skale przeciwstawia się *indeksom*. Patrz rozdział 6.

skala dystansu społecznego Bogardusa (*Bogardus social distance scale*) technika pomiaru, określająca skłonność do wchodzenia w różne – bliższe lub dalsze – stosunki społeczne z innymi kategoriami ludzi. W tej wyjątkowo skutecznej technice stosuje się kilka pytań zamkniętych, które można później zagregować, nie tracąc przy tym szczegółowych danych. Patrz rozdział 6.

skala Guttmana (*Guttman scale*) typ złożonego narzędzia pomiarowego, stosowanego do syntetycznego ujęcia wyników kilku odrębnych obserwacji i przedstawienia pewnej ogólniejszej zmiennej. Patrz rozdział 6.

skala Likerta (*Likert scale*) typ złożonego narzędzia pomiarowego, opracowanego przez Rensisa Likerta w celu poprawienia poziomu pomiaru w naukach społecznych poprzez zastosowanie zestandaryzowanych kategorii odpowiedzi w pytaniach kwestionariusza. Ma ono na celu określenie względnego natężenia różnych postaw i poglądów. Na pytania tego rodzaju mogą być udzielane odpowiedzi w rodzaju: „zdecydowanie się zgadzam", „raczej się zgadzam", „raczej się nie zgadzam", „zdecydowanie się nie zgadzam". Pytania takie mogą być stosowane zarówno do budowy właściwych skal Likerta, jak też w innego rodzaju złożonych narzędziach pomiarowych. Patrz rozdział 6.

skala Thurstone'a (*Thurstone scale*) typ złożonego narzędzia pomiarowego, skonstruowanego na podstawie określonych przez „kompetentnych sędziów" wag przypisanych poszczególnym wskaźnikom zmiennych. Patrz rozdział 6.

specyfikacja (*specification*) proces, w wyniku którego uściślane są pojęcia. W kontekście modelu analizy rozbudowanej jest to techniczny termin określający sytuację, gdy pierwotnie obserwowana zależność między dwiema zmiennymi replikuje się w niektórych (ale nie we wszystkich) podgrupach wyznaczonych przez zmienną kontrolną. W takiej sytuacji formułuje (specyfikuje) się warunek zachodzenia pierwotnej zależności – na przykład zachodzi ona w grupie mężczyzn, ale nie wśród kobiet.

statystyka (*statistic*) także (oprócz ogólnego znaczenia „dział matematyki"): wartość syntetycznie charakteryzująca rozkład zmiennej w próbie, służąca do estymacji parametru w populacji. Patrz rozdział 7.

statystyka opisowa (*descriptive statistics*) część statystyki, zajmująca się opisem cech próby lub zależności między zmiennymi w próbie – streszczaniem do syntetycznej postaci wyników dużej liczby obserwacji dokonanych na elementach próby. Aby jednak przejść z poziomu opisu konkretnych obserwacji z próby do formułowania wniosków dotyczących całej populacji, z której pochodzi próba, należy odwołać się do *wnioskowania statystycznego*.

studium przypadku (*case study*) pogłębiona analiza pojedynczego przykładu pewnego ogólniejszego zjawiska społecznego (na przykład wsi, rodziny, gangu młodzieżowego). Patrz rozdział 10.

średnia (*mean*) wartość przeciętna obliczona przez zsumowanie wartości wielu obserwacji, po czym podzielenie sumy przez ich liczbę. Jeśli do tej pory zdałeś 10 egzaminów i twoja średnia ocen wynosi 4, a z egzaminu z metodologii dostaniesz dwóję, to twoja nowa średnia wyniesie 3,81.

tabela krzyżowa (tabela kontyngencji) (*contingency table*) format przedstawienia relacji między zmiennymi w postaci procentowych rozkładów ich łącznych częstości. Przykłady i wskazówki, jak takie tabele przygotowywać, znajdują się w rozdziale 15.

teoria (*theory*) systematyczne wyjaśnienie obserwacji wiążących się z jakimś konkretnym zjawiskiem społecznym (na przykład przestępczością nieletnich, stratyfikacją społeczną, rewolucją polityczną).

teoria ugruntowana (*grounded theory*) (1) podejście do badania zjawisk społecznych, zorientowane na indukcyjne budowanie teorii poprzez stałe porównywanie kolejnych pojawiających się obserwacji, zasadniczo różne od podejścia zorientowanego na testowanie hipotez, w którym to teoria ma służyć kreowaniu hipotez poddawanych testom empirycznym. Patrz rozdział 10. (2) teoria, która ma zakaz fruwania.

testy istotności statystycznej (*tests of statistical significance*) (1) procedury obliczeń statystycznych, przeprowadzanych w celu ustalenia, jakie jest prawdopodobieństwo, że zależność obserwowana w próbie może być powodowana jedynie przez błąd z próby. Patrz *wnioskowanie statystyczne* oraz rozdział 17. (2) badanie, jak istotne jest zdanie egzaminu ze statystyki dla ukończenia studiów.

trafność (*validity*) cecha pomiaru, stwierdzająca, że narzędzie pomiarowe mierzy rzeczywiście to, co ma mierzyć. Na przykład wynik testu IQ może być trafniejszą miarą twojej inteligencji niż liczba godzin spędzonych w bibliotece. Ponieważ rzeczywistej trafności pomiaru nigdy nie da się ustalić, pozostaje nam umówić się na określanie względnej trafności na podstawie *trafności fasadowej, trafności kryterialnej, trafności treściowej, trafności teoretycznej* oraz *walidacji ze-*

wnętrznej i *walidacji wewnętrznej*. Trafności nie należy mylić z *rzetelnością*. Patrz rozdział 5.

trafność fasadowa (*face validity*) (1) własność wskaźnika (narzędzia pomiarowego), powodująca, że jest on uznawany za odpowiedni do pomiaru jakiejś zmiennej. Częstotliwość udziału w nabożeństwach jest bez wątpienia jakimś wskaźnikiem religijności – a więc jako miara religijności charakteryzuje się trafnością fasadową. Patrz rozdział 5. (2) właściwie dobrany kolor elewacji frontowej ściany budynku.

trafność kryterialna (trafność predykcyjna) (*criterion-related validity, predictive validity*) stopień, w jakim dany miernik powiązany jest z pewnymi kryteriami zewnętrznymi. Na przykład trafność pomiaru dokonywanego przez komisję egzaminacyjną jest wykazywana przez późniejsze osiągnięcia przyjętych kandydatów. Patrz rozdział 5.

trafność teoretyczna (*construct validity*) stopień, w jakim dany miernik powiązany jest z innymi zmiennymi, z którymi zgodnie z założeniami teorii powinien być powiązany. Patrz rozdział 5.

trafność treściowa (*content validity*) stopień, w jakim dany miernik uwzględnia wszystkie znaczenia zawarte w pojęciu określającym mierzone zjawisko. Patrz rozdział 5.

treść jawna (*manifest content*) w analizie treści – zwroty bezpośrednio zawarte w analizowanym przekazie, w odróżnieniu od *treści ukrytej*. Patrz rozdział 11.

treść ukryta (*latent content*) (1) w analizie treści – ukryty sens przekazu, odróżniany od jego *treści jawnej*. Patrz rozdział 11. (2) to, czego nie widać na białej kartce papieru.

typologia (*typology*) klasyfikacja (z reguły nominalna) obserwacji na podstawie wartości dwóch lub więcej zmiennych. Gazety można skategoryzować na przykład tak, jak to przedstawiono w tabeli 6.4 na s. 197. Patrz rozdział 6.

walidacja zewnętrzna (*external validation*) proces testowania trafności narzędzia pomiarowego (na przykład indeksu lub skali) poprzez badanie jego związku z innymi domniemanymi wskaźnikami mierzonej zmiennej. Jeśli, powiedzmy, dany indeks rzeczywiście mierzy poziom uprzedzeń, powinien być skorelowany z innymi ich wskaźnikami. W rozdziale 6 znajdziesz dłuższe omówienie tego problemu wraz z przykładami.

warstwowanie (stratyfikacja) (*stratification*) zestawianie jednostek populacji w homogeniczne grupy (warstwy) przed losowaniem. Procedura ta zwiększa reprezentatywność próby, przynajmniej ze względu na zmienne, użyte do warstwowania. Może być stosowana łącznie z doborem prostym, systematycznym lub grupowym.

wartości zmiennej (atrybuty) (*attributes*) cechy osób lub rzeczy. Patrz *zmienna* i rozdział 1.

ważenie (*weighting*) przypisywanie różnych wag przypadkom, które miały różne szanse na dostanie się do próby. W najprostszym przykładzie każdemu przypadkowi przypisuje się wagę, będącą odwrotnością jego prawdopodobieństwa wylosowania. Jeśli wszystkie przypadki miały takie same szanse na dostanie się do próby, nie ma potrzeby przeprowadzania ważenia. Patrz rozdział 7.

wnioskowanie statystyczne (*inferential statistics*) dział statystyki, zajmujący się formułowaniem wniosków, dotyczących szerszej populacji, na bazie obserwacji dokonanych na próbie. Patrz *statystyka opisowa* oraz rozdział 17.

wskaźnik (*indicator*) wybrana obserwacja, którą traktujemy jako odzwierciedlenie zmiennej, którą chcemy badać. Na przykład udział w nabożeństwach możemy traktować jako wskaźnik religijności. Patrz rozdział 15.

wskaźniki społeczne (*social indicators*) miary opisujące warunki lub charakter życia społecznego, takie jak poziom przestępczości, śmiertelność niemowląt, liczba lekarzy na 100 tysięcy mieszkańców itd. Wskaźniki społeczne monitoruje się często w celu określenia charakteru zmian zachodzących w społeczeństwie. Patrz rozdział 12.

wymiar (*dimension*) możliwy do wyszczególnienia aspekt pojęcia. Patrz rozdział 5.

wytwory społeczne (*social artifacts*) wszelkie wytwory istot społecznych lub ich zachowania; mogą być jednostką analizy. Patrz rozdział 4.

wywiad (*interview*) spotkanie, którego celem jest zebranie danych, w czasie którego jedna osoba (ankieter, badacz, moderator) zadaje pytania drugiej (respondentowi). Może być przeprowadzany bezpośrednio (*face-to-face*) albo telefonicznie. W rozdziale 9 znajdziesz więcej informacji o wywiadzie w badaniach sondażowych.

wywiad jakościowy (*qualitative interview*) w przeciwieństwie do wywiadu standaryzowanego, stosowanego w badaniach sondażowych, wywiad jakościowy prowadzony jest raczej na podstawie listy tematów do pogłębionego omówienia niż zestandaryzowanego zestawu pytań. Patrz rozdział 10.

zależność cząstkowa (*partial relationship*) w modelu analizy rozbudowanej – zależność między dwiema zmiennymi, rozpatrywana w podzbiorze przypadków, zdefiniowanym przez trzecią zmienną. Wychodząc od zależności początkowej, na przykład między *identyfikacją partyjną* a *postawami wobec aborcji* możemy badać, czy zależność ta zachodzi zarówno dla kobiet, jak

i dla mężczyzn (tzn. kontrolujemy wartość zmiennej *płeć*). Związek między dwiema pierwszymi zmiennymi, obserwowany wśród mężczyzn lub kobiet, będzie *zależnością cząstkową*.

zależność początkowa (*zero-order relationship*) w modelu analizy rozbudowanej pierwotna zależność między dwiema zmiennymi, niekontrolowana ze względu na wartości innych zmiennych.

zależność pozorna (*spurious relationship*) przypadkowa korelacja pomiędzy dwiema zmiennymi, którą można wyjaśnić, odwołując się do wpływu trzeciej zmiennej. Na przykład istnieje dodatnia zależność między liczbą wozów strażackich, biorących udział w gaszeniu pożaru, a wielkością szkód – im więcej wozów strażackich, tym zniszczenia są większe. Trzecią zmienną jest w tym przypadku wielkość pożaru: do dużego pożaru, powodującego duże szkody, wysyła się więcej załóg, do małego pożaru zaś wysyła się mało wozów, a ze względu na niewielki rozmiar pożaru szkody są niewielkie. Wysłanie większej liczby wozów nie powoduje zwiększenia szkód. Właściwie wysłanie do pożaru o określonej sile większej liczby wozów powinno ograniczyć straty wywołane przez ogień.

zasada świadomej zgody (*informed consent*) zasada, według której badani mogą wyrazić zgodę na udział w badaniach tylko wtedy, gdy są świadomi wszelkich związanych z badaniami elementów ryzyka. Patrz rozdział 18.

zmienna ciągła (*continuous variable*) zmienna, której wartości zmieniają się płynnie, jak *wiek* czy *dochód*. Na przykład w grupie ludzi mogą znaleźć się osoby w wieku 21, 22, 23, 24 lata; można też wiek wyznaczyć jeszcze dokładniej, w ułamkach lat. Przeciwieństwem zmiennych ciągłych są *zmienne dyskretne*, takie jak *płeć* lub *afiliacja religijna*, których wartości tworzą odrębne elementy. Patrz rozdział 15.

zmienna dyskretna (skokowa) (*discrete variable*) zmienna, na przykład *płeć*, *afiliacja religijna*, której wartości są wyraźnie od siebie oddzielone (nieciągłe), w przeciwieństwie do wartości *zmiennych ciągłych*, płynnie przechodzących jedna w drugą. Tak więc w przypadku ciągłej zmiennej *wiek* jej wartości zmieniają się, na przykład od 21 poprzez 22 do 23 itd., nie można natomiast wskazać takiego przejścia między wartościami zmiennej *płeć* – „mężczyzna" i „kobieta". Patrz rozdział 15.

zmienna kontrolna (testowa) (*control variable, test variable*) zmienna, której wartość kontroluje się (utrzymuje się na ustalonym poziomie), aby wyjaśnić charakter zależności między dwiema innymi zmiennymi. Jeśli na przykład stwierdzimy istnienie zależności między wykształceniem a poziomem uprzedzeń, możemy zbadać tę zależność ze względu na zmienną kontrolną „płeć", badając zależność między wykształceniem a uprzedzeniami osobno w grupach kobiet i mężczyzn. Patrz rozdział 16, w którym dowiesz się więcej o tym, jak ważne jest właściwe stosowanie zmiennych kontrolnych w analizie.

zmienna niezależna (*independent variable*) (1) zmienna, której wartości w analizie traktuje się jako dane i nie próbuje się tłumaczyć. Zakłada się, że zmienne niezależne determinują wartości zmiennych zależnych lub wpływają na nie. Jeśli stwierdzimy, że *religijność* jest częściowo funkcją *płci* – ponieważ kobiety są bardziej religijne od mężczyzn – to zmienną niezależną jest *płeć*, a zmienną zależną *religijność*. Zauważmy, że zmienna traktowana jako niezależna w innej części analizy może stać się zmienną zależną – na przykład *religijność* może być zmienną niezależną tam, gdzie analizujemy *poziom przestępczości*. Patrz *zmienna zależna*. (2) zmienna, której nikogo się nie słucha.

zmienna zniekształcająca (*distorter variable*) w modelu analizy rozbudowanej zmienna kontrolna której włączenie do modelu odwraca kierunek zależności początkowej.

zmienna tłumiąca (*suppressor variable*) w modelu analizy rozbudowanej zmienna, która sprawia, że rzeczywiście istniejąca zależność między dwiema zmiennymi nie ujawnia się jako zależność początkowa.

zmienna zależna (*dependent variable*) (1) zmienna, co do której zakłada się, że jest powodowana przyczynowo lub uzależniona od innej zmiennej (zmiennej niezależnej). Jeśli stwierdzisz, że *dochód* jest częściowo funkcją *osiągniętego formalnie wykształcenia*, *dochód* traktujesz jako zmienną zależną. (2) zmienna nieudacznica, zmienna na posyłki, zmienna z kompleksem niższości.

zmienne (*variables*) logicznie powiązane zestawy wartości (atrybutów), na przykład zmienna *płeć* składa się z wartości *mężczyzna* i *kobieta*.

zogniskowany wywiad grupowy (grupa fokusowa) (*focus group*) technika polegająca na zebraniu grupy badanych i zainicjowaniu ich dyskusji. Stosowana często w badaniach rynku, na przykład dla oceny produktów lub usług. Patrz rozdział 10.

Bibliografia

A Guide to Writing Sociology Papers, Judith Richlin-Klonsky, Ellen Strenski (red.), St. Martin's Press, New York 1998.

Abdulhadi Rabab, *The Palestinian Women's Autonomous Movement: Emergence, Dynamics, and Challenges*, „Gender and Society" 1998, 12(6), s. 649.

Anderson Andy B., Basilevsky Alexander, Hum Derek P. J., *Measurement: Theory and Techniques*, w: *Handbook of Survey Research*, Peter H. Rossi, James D. Wright, Andy B. Anderson (red.), Academic Press, New York 1983, s. 231–287.

Anderson Walt, *Reality Isn't What It Used to Be: Theatrical Politics, Ready-to-Wear Religion, Global Myths, Primitive Chic, and Other Wonders of the Postmodern World*, Harper & Row, San Francisco 1990.

Andorka Rudolf, *The Importance and the Role of the Second Economy for the Hungarian Economy and Society*, „Quarterly Journal of Budapest University of Economic Sciences" 1990, 12(2), s. 95–113.

Aneshensel Carol S., Becerra Rosina, Fielder Eve, Schuler Roberleigh, *Participation of Mexican American Female Adolescents in a Longitudinal Panel Survey*, „Public Opinion Quarterly" 1989, 53, s. 548–562.

Asch Solomon, *Nacisk grupy na motywacje i wypaczenia sądów, tłum.* Danuta Jedlicka, w: *Zagadnienia psychologii społecznej*, A. Malewski (red.), PWN, Warszawa 1962.

Asher Ramona M., Fine Gary Alan, *Fragile Ties: Sharing Research Relationships with Women Married to Alcoholics*, w: *Experiencing Fieldwork: An Inside View of Qualitative Research*, William B. Shaffir, Robert A. Stebbins (red.), Sage, Newbury Park, Cal. 1991, s. 196–205.

Auster Carol J., *Manuals for Socialization: Examples from Girl Scout Handbooks 1913–1984*, „Qualitative Sociology" 1985, 8(4), s. 359–367.

Babbie Earl, *The Third Civilization*, „Review of Religious Research", zima 1966, s. 101–121.

– *A Religious Profile of Episcopal Churchwomen*, „Pacific Churchman" styczeń 1967, 12, s. 6–8.

– *Science and Morality in Medicine*, University of California Press, Berkeley 1970.

– *Social Research for Consumers*, Wadsworth Belmont, Cal. 1982.

– *You* Can *Make a Difference*, St. Martin's Press, New York 1985.

– *Survey Research Methods*, Wadsworth, Belmont, Cal. 1990.

– *The Sociological Spirit*, Wadsworth, Belmont, Cal. 1994.

– *Observing Ourselves: Essays in Social Research*, Waveland Press, Prospect Heights, Ill. 1998.

Babbie Earl, Halley Fred, Zaino Jeanne, *Adventures in Social Research*, Pine Forge Press, Newbury Park, Cal. 2000.

Bailey William C., *Murder and Capital Punishment*, w: *Criminal Law in Action*, William J. Chambliss (red.), Wiley, New York 1975.

Bailey William C., Peterson. Ruth D., *Murder, Capital Punishment, and Deterrence: A Review of the Evidence and an Examination on Police Kilings*, „Journal of Social Issues" 1994, 50, s. 53–74.

Baker Vern, Lambert Charles, *The National Collegiate Athletic Association and the Governance of Higher Education*, „Sociological Quarterly" 1990, 31(3), s. 403–421.

Ball-Rokeach Sandra J., Grube Joel W., Rokeach Milton, *Roots: The Next Generation – Who Watched and with What Effect*, „Public Opinion Quarterly" 1981, 45, s. 58–68.

Bart Pauline, Frankel Linda, *The Student Sociologist's Handbook*, General Learning Press, Morristown, N.J. 1986.

Bart Pauline, O'Brien Patricia, *Stopping Rape: Successful Survival Strategies*, Pergamon, New York 1985.

Beck E. M., Tolnay Stewart E., *The Killing Fields of the Deep South: The Market for Cotton and the Lynching of Blacks, 1882–1930*, „American Sociological Review" 1990, 55, s. 526–539.

Becker Howard S., *Tricks of the Trade: How to Think about Your Research While You're Doing It*, University of Chicago, Chicago 1997.

Bednarz Marlene, *Push Polls Statement*, raport dla sieci AAPOR 5 kwietnia 1996 (dostępne online: http://www.aapor.org/ethics/ pushpoll.html).

Belenky Mary Field, McVicker Blythe Clinchy, Goldberger Nancy Rule, Tarule Jill Mattuck, *Women's Ways of Knowing: The Development of Self, Voice, and Mind*, Basic Books, New York 1986.

Bellah Robert N., *Tokugawa Religion*, Free Press, Glencoe, Ill. 1957.

– *Research Chronicle: Tokugawa Religion*, w: *Sociologists at Work*, Phillip E. Hammond (red.), Anchor Books, Garden City, NY 1967, s. 164–185.

– *Christianity and Symbolic Realism*, „Journal for the Scientific Study of Religion" 1970, 9, s. 89–96.
– *Comment on the Limits of Symbolic Realism*, „Journal for the Scientific Study of Religion" 1974, 13, s. 487–489.
Benton Edwin J., Daly John L., *A Question Order Effect in a Local Government Survey*, „Public Opinion Quarterly" 1991, 55, s. 640–642.
Berbrier Mitch, *„Half the Battle": Cultural Resonance, Framing Processes, and Ethnic Affectations in Contemporary White Separatist Rhetoric*, „Social Problems" 1998, 45(4), s. 431–450.
Berg Bruce L., *Qualitative Research Methods for the Social Sciences*, Allyn and Bacon, Boston 1989 (wyd. 3, Allyn and Bacon, Boston 1998).
Berg Richard, Rossi Peter H., *Thinking about Program Evaluation*, Sage, Thousand Oaks, Cal. 1998.
Bernstein Ira H., Havig Paul, *Computer Literacy: Getting the Most from Your PC*, Sage, Thousand Oaks, Cal. 1999.
Bian Yanjie, *Work and Inequality in Urban China*, State University of New York Press, Albany 1994.
Birchfeld R. W., *The New Fowler's Modern English Usage*, wyd. 3, Oxford University Press, New York 1998.
Black Donald, *Production of Crime Rates*, „American Sociological Review" 1970, 35, s. 733–748.
Blair Johnny, Zhao Shanyang, Bickart Barbara, Kuhn Ralph, *Sample Design for Household Telephone Surveys: A Bibliography 1949–1995*, Survey Research Center, University of Maryland, College Park 1995.
Blalock Hubert M. Jr., *Statystyka dla socjologów*, tłum. Marek Tabin, Irena Topińska, Krzysztof Starzec, PWN, Warszawa 1975.
Bolstein Richard, *Comparison of the Likelihood to Vote among Preelection Poll Respondents and Nonrespondents*, „Public Opinion Quarterly" 1991, 55, s. 648–650.
Bradburn Norman M., Sudman Seymour, *Polls and Surveys: Understanding What They Tell Us*, Jossey-Bass, San Francisco 1988.
Brownlee K. A., *A Note on the Effects of Nonresponse on Surveys*, „Journal of the American Statistical Association" 1975, 52(227), s. 29–32.
Campbell Donald, Stanley Julian, *Experimental and Quasi-Experimental Designs for Research*, Rand McNally, Chicago 1963.
Campbell M. L., *Institutional Ethnography and Experience as Data*, „Qualitative Sociology" 1998, 21(1), s. 55–73.
Carpini Michael X. Delli, Keeter Scott, *Stability and Change in the U.S. Public's Knowledge of Politics*, „Public Opinion Quarterly" 1991, 55, s. 583–612.
Carr C. Lynn, *Tomboy Resistance and Conformity: Agency in Social Psychological Gender Theory*, „Gender and Society" 1998, 12(5), s. 528–553.
Cassell James W., *Internet Data Archives for Sociologists*, raport zaprezentowany przed American Sociological Association, New York, 19 sierpnia 1996.
Centers for Disease Control and Prevention, National Center for Health Statistics, „Monthly Vital Statistics Report" 1998, 46 (11, Suppl.), s. 29.
Chafetz Janet, *A Primer on the Construction and Testing of Theories in Sociology*, Peacock, Itasca, Ill. 1978.
Chen Huey-Tsyh, *Theory-Driven Evaluations*, Sage, Newbury Park, Cal. 1990.
Chossudovsky Michel, *The Globalization of Poverty: Impacts of IMF and World Bank Reforms*, Zed Books, London 1997.
Civil Rights and the Black American, Albert Blaunstein, Robert Zangrando (red.), Washington Square Press, New York 1970.
Coffey Amanda, Atkinson Paul, *Making Sense of Qualitative Data: Complementary Research Strategies*, Sage, Thousand Oaks, Cal. 1996.
Cole Stephen, *Making Science: Between Nature and Society*, Cambridge, Harvard University Press, Mass. 1992.
Collins G. C., Blodgett Timothy B., *Sexual Harassment ... Some See It ... Some Won't*, „Harvard Business Review", marzec–kwiecień 1981, s. 76–95.
Comparative Methodology: Theory and Practice in International Social Research, Else Øyen (red.), Sage, Newbury Park, Cal. 1990.
Computer-Aided Qualitative Data Analysis: Theory, Methods, and Practice, Udo Kelle (red.), Sage, Thousand Oaks, Cal. 1995.
Comstock Donald, *Dimensions of Influence in Organizations*, „Pacific Sociological Review", styczeń, 1980, s. 67–84.
Conrad Clifton F., *A Grounded Theory of Academic Change*, „Sociology of Education" 1978, 51, s. 101––112.
Contemporary Field Research, Robert M. Emerson (red.), Little Brown, Boston 1988.
Continuities in Social Research: Studies in the Scope and Method of „The American Soldier", Robert K. Merton, Paul F. Lazarsfeld (red.), Free Press, New York 1950.
Continuities in the Language of Social Research, Lazarsfeld Paul F., Pasanella Ann, Rosenberg Morris (red.), Free Press, New York 1972.
Cook Elizabeth, *Communication to the METHODS listserv*, komunikacja mailowa między Michelem de Seve (T120 @music.ulaval.ca) a Elizabeth Cook (EC1645A@american.edu) [online], 25 kwietnia 1995.

Cook Thomas D., Campbell Donald T., *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Rand McNally, Chicago 1979.

Cooper-Stephenson Cynthia, Theologides Athanasios, *Nutrition in Cancer: Physicians' Knowledge, Opinions, and Educational Needs*, „Journal of the American Dietetic Association" maj 1981, s. 472–476.

Craig R. Stephen, *The Effect of Television Day Part on Gender Portrayals in Television Commercials: A Content Analysis*, „Sex Roles" 1992, 26(5/6), s. 197–211.

Crawford Kent S., Thomas Edmund D., Fink Jeffrey J., *Pygmalion at Sea: Improving the Work Effectiveness of Low Performers*, „Journal of Applied Behavioral Science" październik–grudzień 1980, s. 482–505.

Cunningham J. Barton, *Action Research and Organizational Development*, Praeger, Westport, Conn. 1993.

Davis Fred, *The Martian and the Convert: Ontological Polarities in Social Research*, „Urban Life" 1973, 2(3) s. 333–343.

Davis James A., *The Logic of Causal Order*, Sage, Beverly Hills, Cal. 1985.

– *Changeable Weather in a Cooling Climate atop the Liberal Plateau: Conversion and Replacement in Forty-Two General Social Survey Items, 1972–1989*, „Public Opinion Quarterly" 1992, 56, s. 261–306.

DeFleur Lois, *Biasing Influences on Drug Arrest Records: Implications for Deviance Research*, „American Sociological Review" 1975, 40, s. 88–103.

Denzin Norman K., Lincoln Yvonna S., *Handbook of Qualitative Research*, Sage, Newbury Park, Cal. 1994.

DeVault Marjorie L., *Liberating Method: Feminism and Social Research*, Temple University Press, Philadelphia 1999.

Dillman Don A., *Mail and Telephone Surveys: The Total Design Method*, Wiley, New York 1978.

Donald Marjorie N., *Implications of Nonresponse for the Interpretation of Mail Questionnaire Data*, „Public Opinion Quarterly" 1960, 24, s. 99–114.

Doyle Sir Arthur Conan, *Skandal w Czechach*, tłum. St. Zawirska-Michalska, Nakładem Wacława Pawłowskiego, Warszawa 1922.

Durkheim Émile, *O podziale pracy społecznej*, tłum. Krzysztof Wakar, PWN, Warszawa 1999.

– *Suicide*, Free Press, Glencoe, Ill. [1897] 1951.

Ekins R., *Male Femaling: A Grounded Theory Approach to Cross-Dressing and Sex-Changing*, Routledge, New York 1997.

Elder Glen H. Jr., Pavalko Eliza K., Clipp Elizabeth C., *Working with Archival Data: Studying Lives*, Sage, Newbury Park, Cal. 1993.

Ellison Christopher G., Sherkat Darren E., *Patterns of Religious Mobility among Black Americans*, „Sociological Quarterly" 1990, 31(4), s. 551–568.

Emerson Robert M., Ferris Kerry O., Brooks Gardner Carol, *On Being Stalked*, „Social Problems" 1998, 45(3) s. 289–314.

Ethnography Unbound: Power and Resistance in the Modern Metropolis, M. Burawoy i in. (red.), University of California Press, Berkeley 1991.

Evans William, *Computer-Supported Content Analysis: Trends, Tools, and Techniques*, „Social Science Computer Review" 1996, 14(3), s. 269–279.

Experiencing Fieldwork: An Inside View of Qualitative Research, William B. Shaffer, Robert A. Stebbins (red.), Sage, Newbury Park, Cal. 1991.

Fairness and Accuracy in Reporting, 1999. *Action Alert: ABC News Gives up on Accuracy?*, 28 września. Dostępne na stronie: http://www.fair.org/activism/stossel-america.html.

Feick Lawrence F., *Latent Class Analysis of Survey Questions That Include Don't Know Responses*, „Public Opinion Quarterly" 53, 1989, s. 525–547.

Ferrante Joan, Vaughn Angela, *Let's Go Sociology: Travels on the Internet*, Wadsworth, Belmont, Cal. 1999.

Festinger L., Reicker H. W., Schachter S., *When Prophecy Fails*, University of Minnesota Press, Minneapolis 1956.

Ford David A., *Preventing and Provoking Wife Battery through Criminal Sanctioning: A Look at the Risks*, 1989, rękopis nie publikowany.

Ford David A., Regoli Mary Jean, *The Preventive Impacts of Policies for Prosecuting Wife Batterers*, w: *Domestic Violence: The Changing Criminal Justice Response*, E. S. Buzawa, C. G. Buzawa (red.), Auburn, New York 1992, s. 181–208.

Forslund Morris A., *Patterns of Delinquency Involvement: An Empirical Typology*, raport zaprezentowany podczas Annual Meeting of the Western Association of Sociologists and Anthropologists, Lethbridge, Alberta, 8 lutego 1980.

Foschi Martha, Warriner Keith G., Hart Stephen D., *Standards, Expectations, and Interpersonal Influence*, „Social Psychology Quarterly" 1985, 48(2), s. 108–117.

Fowler Floyd J. Jr., *Improving Survey Questions: Design and Evaluation*, Sage, Thousand Oaks, Cal. 1995.

Fox Katherine J., *The Politics of Prevention: Ethnographers Combat AIDS among Drug Users*, w: *Ethnography Unbound: Power and Resistance in the Modern Metropolis*, M. Burawoy i in. (red.), University of California Press, Berkeley 1991, s. 227–249.

Frankfort-Nachmias Chava, *Social Statistics for a Diverse Society*, Pine Forge Press, Newbury Park, Cal. 1997.

Gall John, *Systemantics: How Systems Work and Especially How They Fail*, Quadrangle, New York 1975.

Gallup George, *Where Parents Go Wrong*, „San Francisco Chronicle" 13 grudnia 1984, s. 7.

Gallup George Jr., Roper Burns, Yankelovich Daniel i in., *Polls that Made a Difference*, „Public Perspective" maj–czerwiec 1990, s. 17–21.

Gamson William A., *Talking Politics*, Cambridge University Press, New York 1992.

Garant Carol, *Stalls in the Therapeutic Process*, „ American Journal of Nursing" grudzień 1980, s. 2166–2167.

Garfinkel H., *Studies in Ethnomethodology*, Prentice-Hall, Englewood Cliffs, N.J. 1967.

Gaventa J., *Towards a Knowledge Democracy: Viewpoints on Participatory Research in North America*, w: *Action and Knowledge: Breaking the Monopoly with Participatory Action-Research*, O. Fals-Borda, M. A. Rahman (red.), Apex Press, New York 1991, s. 121–131.

Gee Wilson, *Social Science Research Methods*, Appleton-Century-Crofts, New York 1950.

Glaser Barney, Strauss Anselm, *The Discovery of Grounded Theory*, Aldine, Chicago 1967.

Glock Charles Y., Ringer Benjamin B., Babbie Earl R., *To Comfort and to Challenge*, University of California Press, Berkeley 1967.

Glock Charles Y., Stark Rodney, *Christian Beliefs and Anti-Semitism*, Harper & Row, New York 1967.

Goffman Erving, *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*, Aldine, Chicago 1961.

– *Stigma: Notes on the Management of a Spoiled Identity*, Prentice-Hall, Englewood Cliffs, N.J. 1963.

– *Frame Analysis*, Harvard University Press, Cambridge, Mass. 1974.

– *Gender Advertisements*, Harper & Row, New York 1979.

Gottlieb Bruce, *Cooking the School Books: How U.S. News Cheats in Picking Its „Best American Colleges"*, 31 sierpnia 1999 (dostępne online: pod http://www.slate.com/crapshoot/99–08–31/crapshoot.asp).

Graham Laurie, Hogan Richard, *Social Class and Tactics: Neighborhood Opposition to Group Homes*, „Sociological Quarterly" 1990, 31(4), s. 513–529.

Greenwood Peter W., Rydell Peter C., Model Karyn, *Diverting Children from a Life of Crime: Measuring Costs and Benefits*, Rand Corporation, Santa Monica, Cal. 1996.

Greenwood Peter W. i in., *Three Strikes and You're Out: Estimated Benefits and Costs of California's New Mandatory-Sentencing Law*, Rand Corporation, Santa Monica, Cal. 1994.

Griffith Alison I., *Mothering, Schooling, and Children's Development*, w: *Knowledge, Experience, and Ruling Relations: Studies in the Social Organization of Knowledge*, M. Campbell, A. Manicom (red.), University of Toronto Press, Toronto 1995, s. 108–121.

Groves Robert M., *Theories and Methods of Telephone Surveys*, w: „Annual Review of Sociology" (t. 16), W. Richard Scott, Judith Blake (red.), Annual Reviews, Palo Alto, Cal. 1990, s. 221–240.

Gubrium Jaber F., Holstein James A., *The New Language of Qualitative Method*, Oxford University Press. New York 1997.

Hamnett Michael P. i in., *Ethics, Politics, and International Social Science Research*, University of Hawaii Press, Honolulu 1984.

Handbook of Applied Social Research Methods, Leonard Bickman, Debra J. Rog (red.), Sage, Thousand Oaks, Cal. 1998.

Hannon Lance, Defronzo James, *The Truly Disadvantaged, Public Assistance, and Crime*, „Social Problems" 1998, 45(3), s. 383–392.

Havemann Ernest, West Patricia Salter, *They Went to College*, Harcourt, Brace, New York 1952.

Healey Joseph F., *Statistics: A Tool for Social Research*, Wadsworth, Belmont, Cal. 1999.

Hedrick Terry E., Bickman Leonard, Rog Debra J., *Applied Research Design: A Practical Guide*, Sage, Newbury Park, Cal. 1993.

Hein Jeremy, *Refugees, Immigrants, and the State*, „Annual Review of Sociology" 1993, 19, s. 43–59.

Hempel Carl G., *Fundamentals of Concept Formation in Empirical Science*, „International Encyclopedia of United Science II", no. 7, 1952.

Heritage J., *Garfinkel and Ethnomethodology*, Polity Press, Cambridge 1984.

Heritage Johen, Greatbatch David, *On the Institutional Character of Institutional Talk*, w: *Talk at Work*, P. Drew, J. Heritage (red.), Cambridge University Press, Cambridge 1992.

Higginbotham A. Leon Jr., *In the Matter of Color: Race and the American Legal Process*, Oxford University Press, New York 1978.

Hilts Philip J., *Values of Driving Classes Disputed*, „San Francisco Chronicle" 25 czerwca 1981, s. 4.

Hirschi Travis, Selvin Hanan, *Principles of Survey Analysis*, Free Press, New York 1973.

Hogan Richard, Perrucci Carolyn C., *Producing and Reproducing Class and Status Differences: Racial and Gender Gaps in U.S. Employment and Retirement Income*, „Social Problems" 1998, 45(4), s. 528–549.

Homan Roger, *The Ethics of Social Research*, Longman, London 1991.

Hoover Kenneth R., *The Elements of Social Scientific Thinking*, St. Martin's Press, New York 1992.

Horowitz Irving Louis, *The Rise and Fall of Project Camelot*, MIT Press, Cambridge, Mass. 1967.

How the Poll Was Conducted, „New York Times" 1 października 1995, s. 15.

Howard Edward N., Norman Darlene M., *Measuring Public Library Performance*, „Library Journal" luty 1981, s. 305–308.

Howell Joseph T., *Hard Living on Clay Street*, Doubleday Anchor, Garden City, NY 1973.

Huberman Michael A., Miles Matthew B., *Data Management and Analysis Methods*, w: *Handbook of Qualitative Research*, Norman K. Denzin, Yvonna S. Lincoln (red.), Sage, Thousand Oaks, Cal. 1994, s. 428–444

Hughes Michael, *The Fruits of Cultivation Analysis: A Reexamination of Some Effects of Television Watching*, „Public Opinion Quarterly" 1980, 44, s. 287–302.

Humphreys Laud, *Tearoom Trade: Impersonal Sex in Public Places*, Aldine, Chicago 1970.

Hurst Leslie, *Mr. Henry Makes a Deal*, w: *Ethnography Unbound: Power and Resistance in the Modern Metropolis*, M. Burawoy i in. (red.), University of California Press, Berkeley 1991, s. 183–202.

Indrayan A. i in., *Three-Decade Trend in Human Development Index in India and Its Major States*, „Social Indicators Research" 1999, 46(1), s. 91–120.

Irwin John, Austin James, *It's About Time: America's Imprisonment Binge*, Wadsworth, Belmont, Cal. 1997.

Isaac Larry W., Griffin Larry J., *A Historicism in Time-Series Analyses of Historical Process: Critique, Redirection, and Illustrations from U.S. Labor History*, „American Sociological Association" 1989, 54, s. 873–890.

Jackman Mary R., Senter Mary Scheuer, *Images of Social Groups: Categorical or Qualified?*, „Public Opinion Quarterly" 1980, 44, s. 340–361.

Jacobs Bruce A., Miller Jody, *Crack Dealing, Gender, and Arrest Avoidance*, „Social Problems" 1998, 45(4), s. 550–569.

Jasso Guillermina, *Principles of Theoretical Analysis*, „Sociological Theory" 1988, 6, s. 1–20.

Jensen Arthur, *How Much Can We Boost IQ and Scholastic Achievement?*, „Harvard Educational Review" 1969, 39, s. 273–274.

Jobes Patrick C. i in., *Shopping as a Social Problem: A Grounded Theoretical Analysis of Experiences among Romanian Shoppers*, „Journal of Applied Sociology" 1997, 14(1), s. 124–146.

Johnson Jeffrey C., *Selecting Ethnographic Informants*, Sage, Newbury Park, Cal. 1990.

Johnston Hank, *The Marketed Social Movement: A Case Study of the Rapid Growth of TM*, „Pacific Sociological Review" lipiec 1980, s. 333–354.

Johnston Hank, Snow David A., *Subcultures and the Emergence of the Estonian Nationalist Opposition 1945–1990*, „Sociological Perspectives" 1998, 41(3), s. 473–497.

Jones James H., *Bad Blood: The Tuskegee Syphilis Experiments*, Free Press, New York 1981.

Jones Stephen R. G., *Worker Independence and Output: The Hawthorne Studies Reevaluated*, „American Sociological Review" 1990, 55, s. 176–190.

Kahane Howard, *Logic and Contemporary Rhetoric*, Wadsworth, Belmont, Cal. 1992.

Kaplan Abraham, *The Conduct of Inquiry*, Chandler, San Francisco 1964.

Kardas Edward P., Milford Tommy M., *Using the Internet for Social Science Research and Practice*, Wadsworth, Belmont, Cal. 1996.

Karl Marx: Selected Writings in Sociology and Social Philosophy, T. B. Bottomore, Maximilien Rubel (red.), McGraw-Hill, New York [1843] 1956.

Kasl Stanislav V., Chisolm Rupert F., Eskenazi Brenda, *The Impact of the Accident at Three Mile Island on the Behavior and Well-Being of Nuclear Workers*, „American Journal of Public Health" maj 1981, s. 472–495.

Kasof Joseph, *Sex Bias in the Naming of Stimulus Persons*, „Psychological Bulletin" 1993, 113(1), s. 140–163.

Kebede Alemseghed, Knottnerus J. David, *Beyond the Pales of Babylon: The Ideational Components and Social Psychological Foundations of Rastafari*, „Sociological Perspectives" 1998, 42(3), s. 499–517.

Kendall Patricia L., Lazarsfeld Paul F., *Problems of Survey Analysis*, w: *Continuities in Social Research: Studies in the Scope and Method of „The American Soldier"*, Robert K. Merton, Paul F. Lazarsfeld, (red.), Free Press, New York 1950, s. 133–196.

Khayatt Didi, *Compulsory Heterosexuality: Schools and Lesbians Students*, w: *Knowledge, Experience, and Ruling Relations: Studies in the Social Organization of Knowledge*, M. Campbell, A. Manicom (red.), University of Toronto Press, Toronto 1995, s. 149–163.

Kilburn John C. Jr., *It's a Matter of Definition: Dilemmas of Service Delivery and Organizational Structure in a Growing Voluntary Organization*, „Journal of Applied Sociology" 1998, 15(1), s. 89–103.

Kinnell Ann Marie, Maynard Douglas W., *The Delivery and Receipt of Safer Sex Advice in Pretest Counseling Sessions for HIV and AIDS*, „Journal of Contemporary Ethnography" 1946, 24, s. 405–437.

Krueger Richard A., *Focus Groups*, Sage, Newbury Park, Cal. 1988.

Kuhn Thomas, *Struktura rewolucji naukowych*, tłum. J. Nowotniak, H. Ostromęcka, Aletheia, Warszawa [1970] 2001.

Kurland Daniel J., *The Net, the Web, and You*, Wadsworth, Belmont, Cal. 1996.

Kvale Steinar, *InterViews: An Introduction to Qualitative Research Interviewing*, Sage, Thousand Oaks, Cal. 1996.

Landon, 1 293 669: Roosevelt, 972 897, „Literary Digest" 31 października, 1936, s. 5–6.

Lazarsfeld Paul F., *Foreword*, w: Herbert Hyman, *Survey Design and Analysis*, Free Press, New York 1955.
– *Problems in Methodology*, w: *Sociology Today*, Robert K. Merton (red.), Basic Books, New York 1959.
– *The Varied Sociology of Paul F. Lazarsfeld*, Patricia L. Kendall (red.), Columbia University Press, New York 1982.
Lee Raymond, *Doing Research on Sensitive Topics*, Sage, Newbury Park, Cal. 1993.
Lever Janet, *Sex Differences in the Complexity of Children's Play and Games*, w: *Structure and Process*, Richard J. Peterson, Charlotte A. Vaughan (red.), Wadsworth, Belmont, Cal. 1986, s. 74–89.
Li Tze-chung, *Social Science Reference Sources: A Practical Guide*, Greenwood Press, Westport, Conn. 1990.
Linton Ralph, *The Study of Man*, D. Appleton-Century. New York 1937.
Lofland John, Lofland Lyn H., *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*, wyd. 3 Wadsworth, Belmont, Cal. 1995.
Lopata-Znaniecki Helena, *Widowhood and Husband Sanctifcation*, „Journal of Marriage and the Family" maj 1981, s. 439–450.
Lynd Robert S., Lynd Helen M., *Middletown*, Harcourt Brace, New York 1929.
– *Middletown in Transition*, Harcourt, Brace, New York 1937.
– *Middletown: A Study in American Culture*, Harcourt, Brace, Jovanovich, New York 1956.
Madison Anna-Marie, *Primary Inclusion of Culturally Diverse Minority Program Participants in the Evaluation Process*, „New Directions for Program Evaluation" 1992, 53, s. 35–43.
Manning Peter K., Cullum-Swan Betsy, *Narrative, Content, and Semiotic Analysis*, w: *Handbook of Qualitative Research*, Norman K. Denzin, Yvonna S. Lincoln (red.), Sage, Thousand Oaks, Cal. 1994, s. 463–477.
Marshall Catherine, Rossman Gretchen B., *Designing Qualitative Research*, Sage, Thousand Oaks, Cal. 1995.
Martin David W., *Doing Psychology Experiments*, wyd. 4 Brooks/Cole, Monterey, Cal. 1996.
Marx Karl, *Revue Socialist*, July 5. Reprinted in *Karl Marx: Selected Writings in Sociology and Social Philosophy*, B. Bottomore, Maximilien Rubel (red.), McGraw-Hill, New York [1880] 1956.
– *Kapitał*, t. 1–3, KiW, Warszawa 1951–1959.
Maxwell Joseph A., *Qualitative Research Design*, Sage, Thousand Oaks, Cal. 1996.
McAlister Alfred i in., *Pilot Study of Smoking, Alcohol, and Drug Abuse Prevention*, „American Journal of Public Health" lipiec 1980, s. 719–721.
McGrane Bernard, *The Un-TV and the 10 mph Car: Experiments in Personal Freedom and Everyday Life*, The Small Press, Fort Bragg, Cal. 1994.
McIver John P., Carmines Edward G., *Unidimensional Scaling*, Sage, Newbury Park, Cal. 1981.
Meadows Donella H., Meadows Dennis L., Randers Jørgen, *Beyond the Limits: Confronting Global Collapse, Envisioning a Sustainable Future*, Chelsea Green, Post Mills, Vt. 1992.
Meadows Donella, Meadows Dennis L., Randers Jørgen, Behrens William W. III. *The Limits to Growth*, Universe Books, New York 1972.
Merton Robert K., *Struktura społeczna i anomia*, w: *Teoria socjologiczna i struktura społeczna*, tłum. Ewa Morawska, Jerzy Wertenstein-Żuławski, PWN, Warszawa 2002.
Milgram Stanley, *Some Conditions of Obedience and Disobedience to Authority*, „Human Relations" 1965, 18, s. 57–76.
– *Behawioralne badanie posłuszeństwa*, tłum. Józef Radzicki, w: A. Aronson (red.), *Człowiek istota społeczna, Wybór tekstów*, PWN, Warszawa 2001.
Miller Camilo, E-mail prywatny, 8 września 1999.
Miller Delbert, *Handbook of Research Design and Social Measurement*, Sage, Newbury Park, Cal. 1991.
Mitchell Richard G. Jr., *Secrecy and Disclosure in Field Work*, w: *Experiencing Fieldwork: An Inside View of Qualitative Research*, William B. Shaffer, Robert A. Stebbins (red.), Sage, Newbury Park, Cal. 1991, s. 97–108.
Mitofsky Warren J., *Miscalls Likely in 2000*, „Public Perspective" 1999, 10(5), s. 42–43.
Mohr Lawrence B., *Understanding Significance Testing*, Sage, Newbury Park, Cal. 1990.
Morgan Lewis H., *Systems of Consanguinity and Affinity*, Smithsonian Institution, Washington, DC 1870.
Moskowitz Milt, *The Drugs That Doctors Order*, „San Francisco Chronicle" 23 maja 1981, s. 33.
Moynihan Daniel, *The Negro Family: The Case for National Action*, U.S. Government Printing Office, Washington, DC 1965.
Myrdal Gunnar, *An American Dilemma*, Harper & Row, New York 1944.
Naisbitt John, *Megatrendy: dziesięć nowych kierunków zmieniających nasze życie*, tłum. Paweł Kwiatkowski, Zysk, Poznań 1997.
Neuman W. Lawrence, *Negotiated Meanings and State Transformation: The Trust Issue in the Progressive Era*, „Social Problems" 1998, 45(3), s. 315–335.
Nicholls William L. II, Baker Reginald P., Martin Jean, *The Effect of New Data Collection Technology on Survey Data Quality*, w: *Survey Measurement and Process Quality*, L. Lyberg i in. (red.), In press, Wiley, New York (w druku)

O'Neill Harry W., *They Can't Subpoena What You Ain't Got*, „AAPOR News" 1992, 19(2), s. 4, 7.

Our Sociological Eye: Personal Essays on Society and Culture, Arthur Shostak (red.), Alfred, Port Washington, NY 1977.

Perinelli Phillip J., *Nonsuspecting Public in TV Call-in Polls*, „New York Times" 14 lutego 1986, list do wydawcy.

Petersen Larry R., Maynard Judy L., *Income, Equity, and Wives' Housekeeping Role Expectations*, „Pacific Sociological Review" styczeń 1981, s. 87–105.

Picou J. Steven, *Compelled Disclosure of Scholarly Research: Some Comments on High Stakes Litigation*, „Law and Contemporary Problems" 1996, 59(3), s. 149–157.

– *Sociology and Compelled Disclosure: Protecting Respondent Confidentiality*, „Sociological Spectrum" 1996, 16(3), s. 207–238.

Plutzer Eric, Zipp John F., *Identity Politics, Partisanship, and Voting for Women Candidates*, „Public Opinion Quarterly" 1996, 60, s. 30–57.

Polivka Anne E., Rothgeb Jennifer M., *Redesigning the CPS Questionnaire*, „Monthly Labor Review" 1993, 116(9), s. 10–28.

Population Communications International, *International Dateline*, New York, Population Communications International, luty 1996.

Powell Elwin H., *Occupation, Status, and Suicide: Toward a Redefinition of Anomie*, „American Sociological Review" 1958, 23, s. 131–139.

Presser Stanley, Blair Johnny, *Survey Pretesting: Do Different Methods Produce Different Results?*, w: *Sociological Methodology 1994*, Peter Marsden (red.), Jossey-Bass, San Francisco, Cal. 1994, s. 73–104.

Qualitative and Quantitative Social Research, Robert K. Merton, James S. Coleman, Peter H. Rossi (red.), Free Press, New York 1979.

Ragin Charles C., Becker Howard S., *What Is a Case?: Exploring the Foundations of Social Inquiry*, Cambridge University Press, Cambridge 1992.

Rasinski Kenneth A., *The Effect of Question Wording on Public Support for Government Spending*, „Public Opinion Quarterly" 1989, 53, s. 388–394.

Ray William J., *Methods toward a Science of Behavior and Experience*, wyd. 6 Wadsworth, Belmont, Cal. 2000.

Redfield Robert, *The Folk Culture of Yucatan*, University of Chicago Press, Chicago 1941.

Reinharz Shulamit, *Feminist Methods in Social Research*, Oxford University Press, New York 1992.

Riecken Henry W., Boruch Robert F., *Social Experimentation: A Method for Planning and Evaluating Social Intervention*, Academic Press, New York 1974.

Ritzer George, *Sociological Theory*, Knopf, New York 1988.

Roethlisberger F. J., Dickson W. J., *Management and the Worker*, Harvard University Press, Cambridge, Mass. 1939.

Rogers Everett M. i in., *Effects of an Entertainment-Education Radio Soap Opera on Family Planning and HIV/AIDS Prevention Behavior in Tanzania*. Raport przedstawiony na spotkaniu Tanzania Entertainment-Education Project, Rockefeller Foundation, New York, 27 marca 1996.

Roper Burns, *...But Will They Give the Poll Its Due?*, „AAPOR News" 1992, 19(2), s. 5–6.

Rosenberg Morris, *The Logic of Survey Analysis*, Basic Books, New York 1968.

– *Logika analizy sondażowej* (tłum. skrócone), w: A. Sułek (red.), *Logika analizy sondażowej*, Wyd. UW, Warszawa 1979.

Rosenthal Robert, Jacobson Lenore, *Pygmalion in the Classroom*, Holt, Rinehart & Winston, New York 1968.

Rossi Peter H., Freeman Howard E., *Evaluation: A Systematic Approach*, Sage, Newbury Park, Cal. 1996.

Rothman Ellen K., *The Written Record*, „Journal of Family History" wiosna 1981, s. 47–56.

Rubin Herbert J., Rubin Riene S., *Qualitative Interviewing: The Art of Hearing Data*, Sage, Thousand Oaks, Cal. 1995.

Sacks Jeffrey J., Krushat Mark W., Newman Jeffrey, *Reliability of the Health Hazard Appraisal*, „American Journal of Public Health" lipiec 1980, s. 730–732.

Sanders William B., *Gangbangs and Drive-bys: Grounded Culture and Juvenile Gang Violence*, Aldine De Gruyter. New York 1994.

Schiflett Kathy L., Zey Mary, *Comparison of Characteristics of Private Product Producing Organizations and Public Service Organizations*, „Sociological Quarterly" 1990, 31(4), s. 569–583.

Schmitt Frederika E., Martin Patricia Yancey, *Unobtrusive Mobilization by an Institutionalized Rape Crisis Center: „All We Do Comes from Victims"*, „Gender and Society" 1999, 13(3), s. 364–384.

Schutz Alfred, *The Phenomenology of the Social World*, Northwestern University Press, Evanston, Ill. 1967.

– *On Phenomenology and Social Relations*, University of Chicago Press, Chicago 1970.

Shaver Kelly G., *The Attribution of Blame: Causality, Responsibility, and Blameworthiness*, Springer-Verlag, New York 1985.

Sheatsley Paul F., *Questionnaire Construction and Item Writing*, w: *Handbook of Survey Research*, Peter H. Rossi, James D. Wright, Andy B. Anderson (red.), Academic Press, New York 1983, s. 195–230.

Silverman David, *Interpreting Qualitative Data*, Sage, Newbury Park, Cal. 1993.

– *Doing Qualitative Research: A Practical Handbook*, Sage, Thousand Oaks, Cal. 1999.

Smith Andrew E., Bishop G. F., *The Gallup Secret Ballot Experiments: 1944–1988*, raport przedstawiony na corocznej konferencji American Association for Public Opinion Research, St. Petersburg, Fl. maj 1992.

Smith Dorothy E., *The Everyday World as Problematic*, Northeastern University Press, Boston 1978.

Smith Eric R. A. N., Squire Peverill, *The Effects of Prestige Names in Question Wording*, „Public Opinion Quarterly" 1990, 54, s. 97–116.

Smith Vicki, *The Fractured World of the Temporary Worker: Power, Participation, and Fragmentation in the Contemporary Workplace*, „Social Problems" 1998, 45(4), s. 411–430.

Snow David A., Leon Anderson, *Identity Work among the Homeless: The Verbal Construction and Avowal of Personal Identities*, „American Journal of Sociology" 1987, 92, s. 1336–1371.

Sorokin Pitirim A., *Social and Cultural Dynamics*, Bedminster Press, Englewood Cliffs, N.J. 1937–1940, 4 t.

Spohn Cassie, Horney Julie, *A Case of Unrealistic Expectations: The Impact of Rape Reform Legislation in Illinois*, „The Criminal Justice Policy Review" 1990, 4(1), s. 1–18.

Srole Leo, *Social Integration and Certain Corollaries: An Exploratory Study*, „American Sociological Review" 1956, 21, s. 709–716.

State Saved $21.7 Billion with Five-Year-Old „Three Strikes" Law, „BayInsider" 1 marca 1999 (dostępne pod adresem: http://www.bayinsider.com/news/1999/03/01/three_strikes.html).

Stearns Cindy A., *Breastfeeding and the Good Maternal Body*, „Gender and Society" 1999, 13(3), s. 308.

Sterling T. D., Rosenbaum W. L., Weinkam J. J., *Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa*, „American Statistician" 1995, 49(1), s. 108–112.

Stouffer Samuel i in., *The American Soldier*, Princeton University Press, Princeton, NY 1949–1950, 3 t.

Stouffer Samuel [1949–1950], *Studia z psychologii społecznej w czasie II wojny światowej*, Zarząd Propagandy i Agitacji GZP WP, Biuletyn Informacyjny. Seria psychologii i pedagogiki wojskowej, nr 16, 17, 20, 21, Warszawa 1960.

Strang David, Baron James N., *Categorical Imperatives: The Structure of Job Titles in California State Agencies*, „American Sociological Review" 1990, 55, s. 479–495.

Strauss Anselm, Corbin Juliet, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage, Newbury Park, Cal. 1990.

– *Grounded Theory Methodology: An Overview*, w: *Handbook of Qualitative Research*, Norman K. Denzin, Yvonna S. Lincoln (red.), Sage, Thousand Oaks, Cal. 1994. s. 273–285

Strunk William Jr., White E. B., *The Elements of Style*, 4 wyd. Allyn and Bacon, Boston 1999.

Successful Focus Groups: Advancing the State of the Art, David L. Morgan (red.), Sage, Newbury Park, Cal. 1993.

Survey Research in the Social Sciences, Charles Glock (red.), Russell Sage Foundation, New York 1967.

Swafford Michael, *Soviet Survey Research: The 1970's vs. the 1990's*, „AAPOR News" 1992, 19(3), s. 3–4.

Swalehe Ramadhan i in., *A Content Analysis of the Entertainment-Education Radio Soap Opera „Twende na Wakati" (Let's Go with the Times) in Tanzania*, Arusha, Tanzania: Population Family Life and Education Programme (POFLEP), Ministry of Community Development, Women Affairs, and Children, 15 listopada 1995.

Sweet Stephen, *Using a Mock Institutional Review Board to Teach Ethics in Sociological Research*, „Teaching Sociology" 27 stycznia 1999, s. 55–59.

Tacq Jacques, *Multivariate Analysis Techniques in Social Science Research: From Problem to Analysis*, Sage, Thousand Oaks, Cal. 1997.

Takeuchi David, *Grass in Hawaii: A Structural Constraints Approach*, M.A. thesis, University of Hawaii, 1974.

Tan Alexis S., *Mass Media Use, Issue Knowledge and Political Involvement*, „Public Opinion Quarterly" 1980, 44, s. 241–248.

Tandon Rajesh, Brown L. Dave, *Organization-Building for Rural Development: An Experiment in India*, „Journal of Applied Behavioral Science" kwiecień–czerwiec 1981, s. 172–189.

Taylor Humphrey, Terhanian George, *Heady Days Are Here Again: Online Polling Is Rapidly Coming of Age*, „Public Perspective" 1999, 10(4), s. 20–23.

The Politics of Field Research: Sociology beyond Enlightenment, Jaber F. Gubrium, David Silverman (red.), Sage, Newbury Park, Cal. 1989.

Theory Building in Sociology: Assessing Theoretical Cumulation, Jonathan H. Turner (red.), Sage, Newbury Park, Cal. 1989.

Thomas William I., Znaniecki Florian, *Chłop polski w Europie i Ameryce*, LSW, Warszawa 1976.

Tiano Susan, *Patriarchy on the Line: Labor, Gender, and Ideology in the Mexican Maquila Industry*, Temple University Press, Philadelphia, PA 1994.

Tourangeau Roger i in., *Carryover Effects in Attitude Surveys*, „Public Opinion Quarterly" 1989, 53, s. 495–524.

Tuckel Peter S., Feinberg Barry M., *The Answering Machine Poses Many Questions for Telephone Survey Re-*

searchers, „Public Opinion Quarterly" 1991, 55, s. 200–217.

Turk Guminski Theresa, *Hospital Support: Urban Correlates of Allocation Based on Organizational Prestige*, „Pacific Sociological Review" lipiec 1980, s. 315–332.

Turner Stephen Park, Turner Jonathan H., *The Impossible Science: An Institutional Analysis of American Sociology*, Sage, Newbury Park, Cal. 1990.

United Nations, 1995, *Human Development Report 1995, New York: United Nations Development Program*, [streszczenie w Population Communication International, *International Dateline*, luty 1996, s. 1–4.]

U.S. Bureau of the Census, *Male-Female Differences in Work Experience, Occupation, and Earning*, 1984, „Current Population Reports", Series P-70, No. 10, U.S. Government Printing Office, Washington, DC 1987.

– *Statistical Abstract of the United States, 1995*, CD-ROM CD-SA-95, kwiecień 1996a.

– *Statistical Abstract of the United States, 1996, National Data Book and Guide to Sources*, Washington, DC: U.S. Government Printing Office, 1996b.

– *Statistical Abstract of the United States*, U.S. Government Printing Office, Washington, DC 1999.

U.S. News and World Report, „America's Best Colleges" 1999, 30.

Uwe Flick, *An Introduction to Qualitative Research*, Sage, Thousand Oaks, Cal. 1998.

Veroff Joseph, Hatchett Shirley, Douvan Elizabeth, *Consequences of Participating in a Longitudinal Study of Marriage*, „Public Opinion Quarterly" 1992, 56, s. 325–327.

Votaw Carmen Delgado, *Women's Rights in the United States*, United States Commission on Civil Rights, Inter-American Commission on Women, Clearinghouse Publications, Washington, DC 1979.

Wagner-Pacifici Robin, *Discourse and Destruction: The City of Philadelphia versus MOVE*, University of Chicago Press, Chicago 1995.

Walker Jeffery T., *Fax Machines and Social Surveys: Teaching an Old Dog New Tricks*, „Journal of Quantitative Criminology" 1994, 10(2), s. 181–188.

Walker Research, *Industry Image Study*, wyd. 8 Walker Research, Indianapolis, IN 1988.

Wallace William A., *Causality and Scientific Explanation*, University of Michigan Press, Ann Arbor 1972.

Warner W. Lloyd, *Democracy in Jonesville*, Harper, New York 1949.

Weber Eugene J. i in., *Nonreactive Measures in the Social Sciences*, Houghton Miffin, Boston 1981.

Weber Max, *The Religion of China*, tłum. Hans H. Gerth, Free Press, New York 1951.

– *Ancient Judaism*, tłum. Hans H. Gerth, Don Martindale. Free Press, New York [1934] 1952.

– *The Religion of India*, tłum. Hans H. Gerth, Don Martindale, Free Press, New York [1934] 1958.

– *Etyka protestancka i duch kapitalizmu*, tłum. Jan Miziński, Test, Lublin 1994.

– *Nauka jako powołanie*, w: *Polityka jako zawód i powołanie*, tłum. Andrzej Kopacki, Paweł Dybel, Kraków–Warszawa 1998.

Weber Robert Philip, *Basic Content Analysis*, Sage, Newbury Park, Cal. 1990.

Weiss Carol, *Evaluation Research*, Prentice-Hall, Englewood Cliffs, N.J. 1972.

Wellman David, *The Union Makes Us Stronger: Radical Unionism on the San Francisco Waterfront*, Cambridge University Press, Cambridge, Mass. 1995.

Wharton Amy S., Baron James N., *So Happy Together? The Impact of Gender Segregation on Men at Work*, „American Sociological Review" 1987, 52, s. 574–587.

What Went Wrong with the Polls?, „Literary Digest" 14 listopada 1936, s. 7–8.

Whyte W. F., *Street Corner Society*, University of Chicago Press, Chicago 1943.

Whyte W. F., Greenwood D. J., Lazes P., *Participatory Action Research: Through Practice to Science in Social Research*, w: *Participatory Action Research*, W. F. Whyte (red.), Sage, New York 1991, s. 19–55.

Wieder D. L., *Language and Social Reality: The Case of Telling the Convict Code*, University Press of America, Landman, MD 1988.

Williams Robin M. Jr., *The American Soldier: An Assessment, Several Wars Later*, „Public Opinion Quarterly" 1989, 53, s. 155–174.

Wolcott Harry F., *The Art of Fieldwork*, AltaMira Press, Walnut Creek, Cal. 1995.

World Opinion Update, październik 1985, maj 1986.

Yammarino Francis J., Skinner Steven J., Childers Terry L. Terry L., *Understanding Mail Survey Response Behavior: A Meta-Analysis*, „Public Opinion Quarterly" 1991, 55, s. 613–639.

Yerg Beverly J., *Reflections on the Use of the RTE Model in Physical Education*, „Research Quarterly for Exercise and Sport" marzec 1981, s. 38–47.

Yinger J. Milton i in., *Middle Start: An Experiment in the Educational Enrichment of Young Adolescents*, Cambridge University Press, London 1977.

Bibliografia uzupełniająca

Zestawił Antoni Sułek

Bibliografia obejmuje ważniejsze prace z zakresu metodologii badań społecznych, które nawiązując do treści podręcznika, wyraźnie poza te treści wykraczają. Bibliografia obejmuje wyłącznie teksty w języku polskim, zarówno oryginalne, jak i przekłady z języków obcych. Prace w bibliografii ułożono według rozdziałów podręcznika. Jeśli praca miała więcej niż jedno wydanie, wskazywano wydanie najłatwiej dostępne.

Dodatkowe lektury znaleźć można w przygotowanej przez K. Lutyńską bibliografii *Metodologia badań socjologicznych*, cz. 1–2, IFiS PAN, Warszawa 2000–2002; bibliografia ta obejmuje lata 1945–1999 i jest adnotowana. Można je znaleźć także w metodologicznych hasłach *Encyklopedii Socjologii* (Oficyna Naukowa, Warszawa 1998–2002) oraz w innych podręcznikach metodologii, zwłaszcza we *Wprowadzeniu do metod socjologii empirycznej* R. Mayntz, K. Holma i P. Hübnera (PWN, Warszawa 1985) i *Metodach badawczych w naukach społecznych* C. i D. Nachmiasów (Zysk, Poznań 2001).

1. Działalność badawcza człowieka a nauka

Ideę nauki przedstawiają czasem sami uczeni, na przykład socjolog R. K. Merton (*Teoria socjologiczna i struktura społeczna*, PWN, Warszawa 2002, rozdział *Nauka i demokratyczny ład społeczny*) i biolog R. Dunkbar (*Kłopoty z nauką*, Marabut, Gdańsk 1995, rozdział *Co to takiego nauka*). Ideę nauki łatwiej zrozumieć, gdy naukę porówna się z inni rodzajami wiedzy. Filozof E. Nagel (*Struktura nauki*, PWN, Warszawa 1970, rozdział 1) porównuje naukę z wiedzą potoczną, natomiast M. Przełęcki (*Poza granicami nauki. Z semantyki poznania pozanaukowego* (Pol. Tow. Semiotyczne, Warszawa 1996) – z poznaniem wartości, poznaniem filozoficznym, religijnym i artystycznym. *Logika odkrycia naukowego* (1935) K. R. Poppera (PWN, Warszawa 2002) i *Struktura rewolucji naukowych* T. S. Kuhna (Aletheia, Warszawa 2001) są być może dwiema najważniejszymi książkami o nauce, jakie napisano w XX wieku.

Wykład spraw podstawowych nauki zawierają podręczniki filozofii nauki: A. F. Chalmers, *Czym jest to, co zwiemy nauką* (Siedmioróg, Wrocław 1993), C. G. Hempel, *Filozofia nauk przyrodniczych* (Aletheia, Warszawa 2001) i J. Losee, *Wprowadzenie do filozofi nauki* (Pruszyński, Warszawa 2001). Podstawowe pro blemy nauk społecznych przedstawione są gdzie in dziej. P. Sztompka wydał wybór przekładów *Metodo logiczne podstawy socjologii* (UJ, Kraków 1975) a S. Nowak w *Metodologii badań społecznych* (PWN Warszawa 1985) wyłożył metodologiczne podstaw socjologii w tradycji i języku filozofii nauki. *O osobli wościach nauk społecznych* (1962) S. Ossowskiego jes ważną książką o socjologii jako nauce humanistycz nej. Swoistości socjologii jako nauki bronią także książki E. Mokrzyckiego *Filozofia nauki a socjologi* (PWN, Warszawa 1980) i Z. Krasnodębskiego, *Rozu mienie ludzkiego zachowania* (PIW, Warszawa 1986).

Metodologia jako dziedzina zainteresowa zwięźle przedstawiona jest w napisanym prze A. Sułka haśle *Metodologia socjologiczna* w *Encyklope dii Socjologii*, t. 2, 1999 – tam dalsza literatura; arty kuł K. Szaniawskiego *Metoda i twórczość w nauce* w jego tomie *O nauce, rozumowaniu i wartościac* (PWN, Warszawa 1994) jest szczególnie wart lektur

2. Paradygmaty, teoria i badania społeczne

Pojęcie paradygmatu wprowadził T. S. Kuhn w *Stru turze rewolucji naukowych* (Aletheia, Warszawa 2001 Większość paradygmatów opisanych w niniejszy podręczniku jest dokładniej scharakteryzowan w *Strukturze teorii socjologicznej* J. H. Turnera (PW Warszawa 2003) i wydaniu nowym *Historii myśli s cjologicznej* J. Szackiego (PWN, Warszawa 2002). W dany pod red. S. Nowaka zbiór *Metodologiczne probl my teorii socjologicznych* (PWN, Warszawa 197 przedstawia założenia różnych podejść teoretyc nych w naukach społecznych, m.in. historyzm (J. Szacki), funkcjonalizmu (P. Sztompka) i strateg idealizacyjnej (L. Nowak).

S. Nowak w *Metodologii badań społecznych* (PW Warszawa 1985), omawia różne strategie budowy t orii, a P. Sztompka w *Teorii i wyjaśnianiu* (PWN, Wa szawa 1973) przedstawił obszernie strategię budow teorii jako systemu wyjaśnień. Książka A. Malewski go *O zastosowaniach teorii zachowania* (1964), *O nowy kształt nauk społecznych*, PWN, Warszaw 1975, (zwł. cz. 2) jest pięknym przykładem budow

nia teorii wyjaśniających. R. K. Merton w *Teorii socjologicznej i strukturze społecznej* (PWN, Warszawa 2002) pokazuje związki między teorią a badaniami: pisze o znaczeniu teorii socjologicznej dla badań empirycznych (rozdział 4) i znaczeniu badań empirycznych dla teorii socjologicznej (rozdział 5).

3. Pojęcie przyczynowości w badaniach społecznych

Rozumienie przyczynowości pogłębi lektura *Struktury nauki* E. Nagla (PWN, Warszawa 1970); pisze on o zasadzie przyczynowości (10.V), o prawach przyczynowych i nieprzyczynowych (14.I), o wyjaśnieniach przyczynowych i kryteriach ważności przyczynowej w historii, dyscyplinie bliskiej socjologii (15.III), a także o determinizmie w odniesieniu do nauk społecznych. Podstawową książką dla socjologii jest *Przyczynowość w badaniach socjologicznych* J. Karpińskiego (PWN, Warszawa 1985).

Ujęcie przyczynowości wypracowane przez P. F. Lazarsfelda przedstawił sam jego autor w pracy *Przyczynowa interpretacja zależności statystycznych przy zastosowaniu zmiennej kontrolnej*, w: S. Nowak (red.), *Metody badań socjologicznych* (PWN, Warszawa 1965), a rozbudował S. Nowak w pracy *Przyczynowa interpretacja zależności statystycznych w nieeksperymentalnych badaniach socjologicznych*, w: K. Szaniawski (red.), *Metody statystyczne w socjologii*, PWN, Warszawa 1968. Skrócony przekład książki M. Rosenberga, *The Logic of Survey Analysis* ukazał się po polsku w skrypcie A. Sułka *Logika analizy socjologicznej* (UW, Warszawa 1979), a inny przekład – w wydaniu książkowym jako *Zmienne kontrolne (faktory testujące) w badaniach socjologicznych* (KUL, Lublin 1981).

4. Plan badań

Dobrym uzupełnieniem tego rozdziału jest rozdział 1, *Problematyka i metody badań* w pracy S. Nowaka *Metodologia badań społecznych* (PWN, Warszawa 1985).

Problematyka jednostek i poziomu analizy została wyłożona przez J. Karpińskiego w książce *Przyczynowość w badaniach socjologicznych* (PWN, Warszawa 1985), w rozdziale *Ludzie i zbiorowości. Zagadnienia „poziomu analizy" w socjologii*. Typologia własności jednostek i własności zbiorowości, sformułowana przez P. F. Lazarsfelda i H. Menzla została opisana w ich pracy *O stosunku między własnościami jednostek a własnościami zbiorowości*, przetłumaczonej w zbiorze A. Sułka *Logika analizy socjologicznej* (UW, Warszawa 1979). Przekład klasycznego artykułu W. S. Robinsona *Korelacje ekologiczne a zachowanie jednostek* oraz ważnej pracy G. Firebaugha, *Pewna reguła wnioskowania o zależnościach indywiduowych z danych agregatowych* znajduje się w zbiorze A. Sułka *Metody analizy socjologicznej* (UW, Warszawa 1986).

Wykład metodologii badań uwzględniających wymiar czasowy zawarty jest w dwóch pracach: J. Karpiński, *Przyczynowość w badaniach socjologicznych*, rozdział *Weryfikacja tez o zależnościach przyczynowych w badaniach diachronicznych* oraz A. Sułek i K. Koseła, *Metody badania procesów społecznych*, w zbiorze *Metody analizy socjologicznej*.

5. Konceptualizacja, operacjonalizacja i pomiar

Podstawowy problem, „teoria a doświadczenie", przedstawiony jest w książce A. Chalmersa, *Czym jest to, co zwiemy nauką* (Siedmioróg, Wrocław 1993), głównie w rozdziale 3 *Zależność obserwacji od teorii*, a w odniesieniu do poznania humanistycznego przez Z. Krasnodębskiego, *Rozumienie ludzkiego zachowania* (PIW, Warszawa 1986), głównie w rozdziale 2 zatytułowanym *Czy rozumienie jest empatią?* Metody wprowadzania do nauki pojęć teoretycznych przedstawia M. Przełęcki w studium *Pojęcia teoretyczne i doświadczenie*, w zbiorze T. Pawłowskiego *Logiczna teoria nauki* (PWN, Warszawa 1966).

Jeśli chodzi o definicje, najlepiej czytać teksty polskich logików. Część I zbioru *Logiczna teoria nauki* traktuje o definicjach i zawiera podstawowe prace K. Ajdukiewicza (*Trzy pojęcia definicji*), J. Kotarbińskiej (*Definicja, Tak zwana definicja dejktyczna*) i M. Przełęckiego (*Operacjonizm*); zob. też pracę T. Pawłowskiego *Definicja i argumentacja perswazyjna* w jego *Metodologicznych zagadnieniach humanistyki* (PWN, Warszawa 1969). Spośród najnowszych podręczników godny polecenia jest: E. Żarnecka-Biały, *Mała logika* (UJ, Kraków 1999), głównie rozdział 9 *Poprawianie języka naturalnego: precyzja logiczna, definiowanie*.

O operacjonalizacji zmiennych w socjologii, zob. S. Nowak, *Studia z metodologii nauk społecznych* (PWN, Warszawa 1965), rozdział *Pojęcia i wskaźniki*. Doskonałe przykłady operacyjnego definiowania pojęć znajdują się w pracy M. L. Kohna i C. Schoolera, *Praca a osobowość* (PWN, Warszawa 1986), zwłaszcza rozdział 2 *Klasa społeczna, zawód i orientacje*, w którym zoperacjonalizowana jest zawodowa samosterowność oraz rozdział 4 *Struktura zawodowa a alienacja*, w którym zoperacjonalizowane jest (za Seemanem) pojęcie tak abstrakcyjne i nieuchwytne jak alienacja; zob. też K. M. Słomczyński i M. L. Kohn, *Sytuacja pracy i jej psychologiczne konsekwencje* (Ossolineum, Wrocław 1988).

O pomiarze, kryteriach jego jakości, trafności i rzetelności najlepiej czytać u psychologów: u J. Brzezińskiego *Metodologia badań psychologicznych* (PWN, Warszawa 1996), rozdziały 12, 17 i 20 – ostatni traktuje o adaptacji kulturowej testów, kwestii ważnej dla badań porównawczych, lub w książce E. Hornowskiej *Testy psychologiczne. Teoria i praktyka* (Scholar, Warszawa 2001).

6. Indeksy, skale i typologie

O naturze pomiaru i typach skal szczególnie jasno pisze T. Pawłowski w *Metodologicznych zagadnieniach humanistyki* (PWN, Warszawa 1969), rozdział 4 *Pomiar w naukach społecznych*.

O konstrukcji typologii i innych systemów opisowych zob. A. H. Barton, P. F. Lazarsfeld, *Niektóre funkcje analizy jakościowej w badaniach społecznych*, w zbiorze A. Sułka, *Metody analizy socjologicznej*, UW, Warszawa 1985; I. Lazari-Pawłowska, *O pojęciu typologicznym w humanistyce*, w zbiorze T. Pawłowskiego, *Logiczna teoria nauki*, PWN, Warszawa 1966.

O budowie skal więcej można przeczytać w innych podręcznikach metodologii: o skali Bogardusa i Thurstone'a w *Metodach badań socjologicznych*, pod red. S. Nowaka (PWN, Warszawa 1965), o skali Likerta w *Metodach badawczych w naukach społecznych* Ch. i D. Nachmiasów (Zysk, Poznań 2001), a o dyferencjale semantycznym, czyli skali Osgooda w artykule J. Czapińskiego *Dyferencjał semantyczny*, w *Materiałach do nauczania psychologii*, seria 3, t. 3 (PWN, Warszawa 1978).

7. Logika doboru próby

Podstawowe podręczniki badań opartych na próbach reprezentacyjnych to: Z. Pawłowski, *Wstęp do statystycznej metody reprezentacyjnej* (PWN, Warszawa 1972), R. Zasępa, *Metoda reprezentacyjna* (PWE, Warszawa 1972) i C. Bracha, *Teoretyczne podstawy metody reprezentacyjnej* (PWN, Warszawa 1996). Bardziej bezpośrednio do badań społecznych odnosi się książka C. Brachy *Metoda reprezentacyjna w badaniu opinii publicznej i marketingu* (Efekt, Warszawa 1998), a pośrednio – dwie ważne prace G. Lissowskiego: *Z zagadnień doboru próby*, w: K. Szaniawski (red.), *Metody statystyczne w socjologii* (PWN, Warszawa 1968) oraz *Problem jednostek niedostępnych w reprezentacyjnych badaniach socjologicznych*, w: K. Szaniawski (red.), *Metody matematyczne w socjologii* (Ossolineum, Wrocław 1971).

8. Eksperyment

Eksperyment w socjologii jest tematem monografii A. Sułka, *Eksperyment w badaniach społecznych* (PWN, Warszawa 1979). Zaawansowany wykład metodologii eksperymentalnej zawierają podręczniki J. Brzezińskiego: *Metodologia badań psychologicznych* (PWN, Warszawa 1996), zwłaszcza rozdziały 4 i 12 oraz *Badania eksperymentalne w psychologii i pedagogice* (Scholar, Warszawa 2000). Książka *Społeczny kontekst badań psychologicznych i pedagogicznych* wydana przez J. Brzezińskiego i J. Siutę (UAM, Poznań 1991) zawiera przekłady podstawowych, w tym klasycznych prac z zakresu „psychologii społecznej eksperymentu", m.in. M. T. Orne, M. J. Rosenberga i R. Rosenthala. Książka E. Aronsona i G. Wieczorkowskiej, *Kontrola naszych myśli i uczuć* (Santorski, Warszawa 2001) dobrze pokazuje metodę eksperymentalną w działaniu. *Psychologia jako nauka*, rozdział znanego podręcznika Aronsona *Człowiek – istota społeczna* (PWN, Warszawa 2001) stanowi doskonałą prezentację metody eksperymentalnej, napisaną przez jej praktyka, natomiast wydany przez Aronsona zbiór pod tym samym tytułem, uzupełnienie do podręcznika, podaje wiele przykładów ważnych eksperymentów w naukach społecznych.

9. Badania sondażowe

Polska ma bogatą tradycję badań sondażowych, ale nie ukazał się u nas żaden solidny podręcznik tej metody. Istnieje natomiast bogata literatura poświęcona różnym aspektom i etapom takich badań; jej wykaz zawiera bibliografia K. Lutyńskiej, *Metodologia badań socjologicznych*, cz. II (IFiS PAN, Warszawa 2002).

O ogólnych problemach sondaży napisano wiele, począwszy od pracy S. Nowaka *Funkcje wskaźnikowe odpowiedzi w badaniach ankietowych*, w jego *Studiach z metodologii nauk społecznych* (PWN, Warszawa 1965), a skończywszy na książce A. Sułka, *Sondaż polski* (IFiS PAN, Warszawa 2001).

Budowa kwestionariusza i układanie pytań jest przedmiotem następujących książek: K. Lutyńska *Wywiad kwestionariuszowy: przygotowanie i sprawdzanie narzędzia badawczego* (Ossolineum, Wrocław 1984) i L. Gruszczyński, *Kwestionariusze w socjologii. Budowa narzędzi do badań surveyowych* (UŚ, Katowice 1999). Praca A. Sułka, *Jak działa kwestionariusz?*, rozdział w *Sondażu polskim* zapoznaje z kognitywnym podejściem do pytań i odpowiedzi ankietowych.

Podręcznik ankietera, pod red. Z. Sawińskiego P. B. Sztabińskiego i F. Sztabińskiego (SMG KRC Poland i IFiS PAN, Warszawa 2000) przedstawia proce realizacji różnych rodzajów sondaży, a szczególn uwagę zwraca na wywiad osobisty i pracę ankietera Problemy ankiety pocztowej szerzej omawiaj książki: F. Sztabiński *Ankieta pocztowa i wywiad kwes tionariuszowy* (IFiS PAN, Warszawa 1997) ora

T. W. Mangione, *Ankietowanie pocztowe w badaniach marketingowych i socjologicznych* (PWN, Warszawa 1999). O sondażach telefonicznych traktują prace: Z. Sawiński, *Sondaże telefoniczne*, „Ask" 1996, nr 1; P. B. Sztabiński, *Wywiad telefoniczny ze wspomaganiem komputerowym (CATI): co zyskujemy, co tracimy?*, „Ask" 1999, nr 8.

Sondaże są od wielu lat przedmiotem badań metodologicznych w środowisku łódzkim. Ich główne wyniki zebrane są w książkach: *Wywiad kwestionariuszowy: analizy teoretyczne i badania empiryczne*, pod red. K. Lutyńskiej i A. P. Wejlanda (Ossolineum, Wrocław 1983); K. Lutyńska, *Surveye w Polsce: spojrzenie socjologiczno-antropologiczne* (IFiS PAN, Warszawa 1993). Od 1995 roku ukazuje się czasopismo metodologiczne „Ask. Społeczeństwo, Badania, Metody", które przynosi najnowsze wyniki badań i informacje z zakresu metodologii sondażowej.

10. Jakościowe badania terenowe

Podręcznik M. Hammersleya i P. Atkinsona *Metody badań terenowych* (Zysk, Poznań 2000) rozszerza i pogłębia wiedzę zawartą w tym rozdziale. Pod nawą „badania terenowe" kryje się cała klasa strategii i metod, jednak istotą takich badań jest obecność badanego w zbiorowości badanej i korzystanie ze wszystkich metod, które pozwolą ją poznać, nade wszystko z obserwacji, swobodnych wywiadów oraz analizy dokumentów wytwarzanych w zbiorowości lub regulujących jej życie.

Wartościową książką dotyczącą obserwacji jest praca M. Q. Pattona, *Obserwacja – metoda badań terenowych*, a wywiadu – praca H. J. i I. S. Rubinów, *Jak zmierzać do celu, nie wiążąc sobie rąk. Projektowanie wywiadów jakościowych* – obie opublikowane w tomie *Ewaluacja w edukacji* pod red. L. Korporowicza (Oficyna Naukowa, Warszawa 1997); o analizie dokumentów traktuje rozdział 6 *Dokumenty* w podręczniku Hammersleya i Atkinsona. Metodologię badań w instytucjach przedstawia książka M. Kostery, *Antropologia organizacji. Metodologia badań terenowych* (PWN, Warszawa 2002). Badania terenowe są często monografiami jednego przypadku – praca R. E. Stake'a *Studium przypadku*, w tomie *Ewaluacja w edukacji*, pokazuje i przybliża problemy z tym związane. Metodologii badań terenowych dobrze uczyć się z prac, których autorzy przedstawiają doświadczenia własne. Zob. np.: P. Moczydłowski, *O sposobach wglądu w sekrety stosunków międzyludzkich. Przypadek instytucji totalnej*, „Kultura i Społeczeństwo" 1990, nr 1, a z klasyki: K. Zawistowicz-Adamska, *Społeczność wiejska. Wspomnienia i materiały z badań terenowych* (LSW, Warszawa 1958).

11. Badania niereaktywne

Poszerzony wykład klasycznej analizy treści zawiera książka W. Pisarka, *Analiza zawartości prasy* (Ośrodek Badań Prasoznawczych, Kraków 1983). Analizę pól semantycznych zaproponował R. Robin, *Badanie pól semantycznych*, w: M. Głowiński (red.), *Język i społeczeństwo* (Czytelnik, Warszawa 1980), a jako jeden z pierwszych w Polsce zastosował ją J. Leoński do badania świadomości robotników na podstawie ich życiorysów konkursowych (*Drogi życiowe i świadomość społeczna robotników polskich*, PWN, Warszawa–Poznań 1987). Analizę treści wspomaganą komputerowo przedstawia i ilustruje własnym badaniem C. Trutkowski w artykule *Analiza treści wspomagana komputerowo. Badanie społecznych reprezentacji polityki*, „Ask" 1999, nr 8. Analiza „słów sztandarowych" jest podstawą książki W. Pisarka, *Polskie słowa sztandarowe i ich publiczność* (Universitas, Kraków 2002). O metodach analizy dyskursu społecznego, wykraczającej znacznie poza tradycyjną analizę treści, mówi gruntowna książka *Dyskurs jako struktura i proces* pod red. T. A. van Dijka (PWN, Warszawa 2001), natomiast książka *Rytualny chaos* pod red. M. Czyżewskiego, S. Kowalskiego i A. Piotrowskiego (Aureus, Kraków 1997) jest przykładem studium dyskursu publicznego.

Dane statystyczne są tylko jednym z rodzajów danych zastanych. Dane zastane tworzone są przez instytucje („urzędy") lub osoby. O tym, jak analizować dane urzędowe, pisze A. Sułek w pracy *Socjolog wobec danych urzędowych* w pracy *Ogród metodologii socjologicznej* (Scholar, Warszawa 2002), a pokazuje to J. Bartkowski w książce *Tradycja i polityka. Wpływ tradycji kulturowych polskich regionów na współczesne zachowania społeczne i polityczne* (Żak, Warszawa 2003). Metody analizy „dokumentów osobistych" są w socjologii lepiej opisane, zob. *Metoda biograficzna w socjologii* J. Włodarek, M. Ziółkowski (red.), (PWN, Warszawa–Poznań 1990). Badania biograficzne opierają się dziś częściej już na materiałach pozyskanych – wywiadach i narracjach biograficznych; ich obszerną prezentację i przykład ich wykorzystania zawiera książka A. Rokuszewskiej-Pawełek, *Chaos i przymus. Wojenne trajektorie Polaków – analiza biograficzna* (UŁ, Łódź 2002).

Metodologii analiz historyczno-porównawczych najlepiej szukać w dziełach wybitnych historyków, np. W. Kuli (*Problemy i metody historii gospodarczej*, rozdział 15) i Ch. Tilly'ego (zob. A. Kolasa-Nowak, *Socjolog w badaniu przeszłości. Koncepcja socjologii historycznej Charlesa Tilly'ego*, UMCS, Lublin 2001). *Trzecia fala demokratyzacji* S. P. Huntingtona (PWN, Warszawa 1995) jest świetnym przykładem analizy po-

równawczej zastosowanej do historii współczesnej. W socjologii prowadzi się międzykrajowe badania porównawcze. Do ich metodologii wprowadzają prace: K. Berting, *Teoretyczne i metodologiczne problemy międzynarodowych badań porównawczych*, „Ruch Prawniczy, Ekonomiczny i Socjologiczny" 1988, nr 2 oraz K. M. Słomczyński, *Typologia międzykrajowych badań porównawczych a kluczowe problemy metodologiczne*, „Studia Socjologiczne" 1995, nr 1; książka Słomczyńskiego i M. L. Kohna *Sytuacja pracy i jej psychologiczne konsekwencje. Polsko-amerykańskie analizy porównawcze* (Ossolineum, Wrocław 1988) stanowi pociągający przykład takich badań.

12. Badania ewaluacyjne

Ewaluacja w edukacji, zbiór przekładów wydany przez L. Korporowicza (Oficyna Naukowa, Warszawa 1997) zawiera wiele cennych tekstów dotyczących zarówno podstaw badań ewaluacyjnych, jak i metod tych badań; zbiór ten bynajmniej nie jest ograniczony do edukacji. *Podręcznik ewaluacji programów profilaktycznych* J. D. Hawkinsa i B. Nederhood (Inst. Psychiatrii i Neurologii – Polskie Tow. Psychologiczne, Warszawa–Olsztyn 1994) łączy wykład metodologii badań ewaluacyjnych z praktycznymi zaleceniami dla reformatorów społecznych. Wykład metodologii ewaluacji programów profilaktycznych zawiera artykuł J. i B. Grzelaków, *Ocena efektywności programów profilaktycznych. Perspektywa metodologiczna*, „Alkoholizm i Narkomania" 1993, nr 13. Eksperymentalne i quasi-eksperymentalne badania ewaluacyjne są omówione w dwóch tekstach A. Sułka: *O eksperymentalnych próbach zmian społecznych*, „Studia Socjologiczne" 1977, nr 4 oraz *Quasi-eksperymentalne oceny reform społecznych*", w jego *Ogrodzie metodologii socjologicznej* (Scholar, Warszawa 2002). Badania wskaźników społecznych omówił M. Tabin w pracy *Wskaźniki społeczne*, „Polska 2000" 1981, nr 1.

13. Analiza danych jakościowych

Dobrym wprowadzeniem do analizy danych jakościowych jest praca N. K. Denzina i Y. S. Lincoln, *Wkraczanie na pole badań jakościowych*, „Socjologia Wychowania", t. 13 (UMK, Toruń 1997) – jest to przekład wprowadzenia do wydanego przez nich podstawowego podręcznika *Handbook of Qualitative Research* (1994). Wprowadzenie do podręcznika M. B. Milesa i A. M. Hubermana *Analiza danych jakościowych* ma tę przewagę, że cały ten podręcznik ukazał się po polsku (Trans Humana, Białystok 2000). Książki A. Wyki, *Badacz społeczny wobec doświadczenia* (IFiS, PAN Warszawa 1993) i K. Koneckiego, *Studia z metodologii badań jakościowych. Teoria ugruntowana* (PWN, Warszawa 2000) pomogą zrozumieć teoretyczne podstawy takich badań, zawierają też ciekawe przykłady. Bibliografia K. Lutyńskiej *Metodologia badań socjologicznych* (IFiS PAN, Warszawa 2000–2002), głównie część II podaje wiele innych prac metodologicznych z zakresu analizy jakościowej i prac zawierających przykłady jej zastosowań.

14. Kwantyfikacja danych i 15. Podstawy analizy ilościowej

Książka J. Górniaka i J. Wachnickiego, *SPSS PL for Windows. Pierwsze kroki w analizie danych* (SPSS Polska, Kraków 2000) jest praktycznym podręcznikiem podstawowych analiz ilościowych przy użyciu SPSS.

16. Model analizy rozbudowanej

Oryginalne autorskie sformułowanie modelu elaboracji znajduje się w pracy P. F. Lazarsfelda, *Przyczynowa interpretacja zależności statystycznych przy zastosowaniu zmiennej kontrolnej*, w zbiorze S. Nowaka *Metody badań socjologicznych* (PWN, Warszawa 1965). Model ten rozbudował M. Rosenberg, skrót *The Logic of Survey Analysis* ukazał się po polsku w skrypcie A. Sułka *Logika analizy socjologicznej* (UW, Warszawa 1979). Teoretycznie model rozbudował S. Nowak w studium *Przyczynowa interpretacja zależności statystycznych w nieeksperymentalnych badaniach socjologicznych*, w: K. Szaniawski (red.), *Metody statystyczne w socjologii* (PWN, Warszawa 1968). Nowak był pionierem analizy wielozmiennowej, jako pierwszy w Polsce wykorzystał ją w analizie wyników badań nad studentami Warszawy (1958); zob. jego *Studenci Warszawy. Studium długofalowych przemian postaw i wartości* (UW, Warszawa 1991).

17. Statystyka w naukach społecznych

Na polski przełożono dwa amerykańskie podręczniki statystyki dla badaczy społecznych: H. M. Blalock, *Statystyka dla socjologów* (PWN, Warszawa 1975) oraz G. A. Ferguson i Y. Takane, *Analiza statystyczna w psychologii i pedagogice* (PWN, Warszawa 1997). W zbiorze A. Sułka *Metody analizy socjologicznej* (UW, Warszawa 1985) znajdują się prezentacje dwóch metod zaawansowanej analizy wielozmiennowej – analizy ścieżkowej i analizy modeli log-liniowych: D. R. Heise, *Problemy analizy ścieżkowej i wnioskowań przyczynowych* oraz G. Lissowski, *Zastosowanie modeli logarytmiczno-liniowych do analizy związków między wieloma zmiennymi jakościowymi*; tamże polskie ilustracje.

18. Etyka i polityka w badaniach społecznych

Badania społeczne mają swoje problemy etyczne, ale w Polsce nie są one jeszcze szerzej dyskutowane; częściej rozważano uwikłania polityczne socjologii. Do wyjątków należą prace: J. Lutyński, *Nauka i polskie problemy* (PIW, Warszawa 1990), zwłaszcza rozdział *Problematyka etyczna w socjologii a system społeczny i kulturowe tradycje*; W. A. Rostocki, *Socjologia a prywatność*, „Przegląd Socjologiczny" 1999, t. 48, nr 1. O etycznych problemach badań społecznych więcej piszą psychologowie, zob.: *Etyczne problemy działalności badawczej i praktycznej psychologów*, J. Brzeziński, W. Poznaniak (red.) (Fund. Humaniora, Poznań 1994); *Etyczne problemy psychologii*, J. Brzeziński, M. Toeplitz-Winiewska (red.) (Fund. Humaniora, Poznań 2000). Zob. też E. Aronson, *Człowiek – istota społeczna* (PWN, Warszawa 2001), głównie rozdział *Psychologia jako nauka*.

Na polski przełożono dwa najważniejsze w świecie kodeksy badań społecznych: opracowany przez Europejskie Stowarzyszenie Badaczy Opinii i Rynku (ESOMAR) *Międzynarodowy Kodeks Postępowania w Dziedzinie Badań Rynkowych i Społecznych* „Kultura i Społeczeństwo" 1994, nr 3 oraz *Kodeks Etyczny Amerykańskiego Stowarzyszenia Socjologicznego*, „Kultura i Społeczeństwo" 1997, nr 4.

O politycznych uwikłaniach socjologii w Polsce zob.: J. Lutyński, *Nauka i polskie problemy*, rozdział *Niektóre uwarunkowania rozwoju socjologii polskiej i ich konsekwencje*; J. Szacki, *W służbie społeczeństwa czy władzy? Socjologia polska w latach 1944–1989*, w: S. Kozyr-Kowalski i in. (red.), *Krytyka rozumu socjologicznego* (Zysk, Poznań 1997); A. Sułek, *Sondaż polski* (IFiS PAN, Warszawa 2001), rozdział *Sługa dwóch panów. Badania opinii publicznej w ostatniej dekadzie socjalizmu realnego*. Szerzej o problemach z praktycznym wykorzystaniem socjologii pisze K. W. Frieske w książce *Socjologia w działaniu. Nadzieje i rozczarowania* (IS UW, Warszawa 1990).

Indeks osób*

(Opracował Jakub Radziejowski)

Abdulhadi Rabab 127
Aburdene Patricia 342
al Haq Mahbub 189
Alicja (Alicja Pleasaunce Liddell), bohaterka książek *Alicja w krainie czarów* i *O tym, co Alicja odkryła po drugiej stronie lustra* autorstwa Lewisa Carrolla 68
Aminzade Ron 364, 365
Anderson Andy B. 170, 199, 241, 307
Anderson Leon 316, 317
Anderson Walt 33
Andorka Rudolf 372
Aneshensel Carol S. 124, 125
Asch Solomon E. 63, 64, 521
Asher Ramona 122
Auster Carol 347, 348
Austin James 387

Babbie Earl R. 52, 70, 72, 179, 231, 306, 310, 364, 457, 467, 470, 471, 509, 517, 531, 586
Babbie Sheila 10
Bailey William C. 392, 393
Baird Anne 9
Baker Vern 368
Ball-Rokeach Sandra 264
Banks Rae 9
Baron James N. 137, 184
Bart Pauline 137, 383, 553
Bartnicki Andrzej 423
Basilevsky Alexander 199
Beck E. M. 104
Becker Howard S. 52, 320
Bednarz Marlene 269
Belenky Mary Field 63
Bellah Robert 314, 362–364, 366
Benton J. Edwin 280
Berbrier Mitch 127
Berg Bruce L. 173, 342, 348, 350, 368
Berg Richard 395
Bergez John 9
Bernstein Ira H. 437
Berra Yogi 327
Beyer William H. 581
Bian Yanjie 121, 122, 273
Bickman Leonard 338, 374, 395
Birchfield Robert W. 563
Bishop George F. 273
Black Donald 357
Blair Johnny 282, 301
Blalock Hubert M. Jr. 509
Blaunstein Albert 529
Blodgett Timothy 236, 237
Bogardus Emory 190, 193
Bohrnstedt George W. 170
Bollen Kenneth 178
Bolstein Richard 288
Bonney Charles 95
Boruch Robert F. 374
Borucki W. C. 56
Bottomore Tom B. 361
Boys Don 98
Bradburn Norman N. 306, 307
Brown L. Dave 381, 382
Brownlee K. A. 287
Burawoy Michael 320, 321, 627
Burneson Ray 388

Campbell Donald T. 253, 255, 258, 259, 264, 265, 266, 368
Campbell Marie L. 322, 323
Carpini Michael X. Delli 122
Carr C. Lynn 126, 127
Carroll Lewis 68
Casley Dennis J. 137
Cassell James W. 562
Chafetz Janet 82
Chamblis William J. 392
Chawla Anil 199
Chen Huey-Tsych 395
Childers Terry L. 288
Chilton Roland 9
Chisolm Rupert F. 263
Chossudovsky Michel 59
Ciechomski Marcin 327
Clinton Bill 66, 89, 96, 201, 202
Cole Stephen 52
Coleman James S. 138, 476, 529, 53
Collins Eliza G. C. 236, 237
Comstock Donald 137
Comte Auguste 57, 58, 61, 63, 36
Conrad Clifton F. 319
Cook Elisabeth 473
Cook Thomas 255, 264, 266
Cooley Charles Horton 60, 260
Cooper Harris M. 137
Cooper-Stephenson Cynthia 205
Corbin Juliet 319, 338, 400, 40 408, 411, 423
Cox James 202
Craig R. Stephen 351, 352
Cramer Richard M. 9
Crawford Kent S. 384
Cullum-Swan Betsy 403
Cunningham J. Barton 395

Daly John 280
D'Andrade Roy 307
Dannemiller James E. 297
Darwin Karol 38, 58
Davis Fred 313
Davis James A. 104, 123, 457
Davis Miles 15
DeFleur Lois 356, 357
Defronzo James 62
DeNuzzo Rinaldo V. 221
Denzin Norman K. 82, 338, **423**
DeVault Marjorie L. 82
Dewey Thomas E. 203, 241
Dickson William John 249
Dillman Don 288, 296, 306
Dole Robert 66, 89, 93, 96, 201
Donald Marjorie N. 287
Doolittle Eliza, bohaterka szt *Pigmalion* autorstwa Georg Bernarda Shawa 260
Douvan Elizabeth 124
Doyle sir Arthur Conan 75
Dunlap Riley E. 71

* Indeks osób nie obejmuje *Bibliografii* i *Bibliografii uzupełniającej*.

Durkheim Émile 20, 61, 74, 138, 148–150, 341, 353–356, 361

Eagan Andrea Boroff 550
Eichman Adolf 150
Elder Glen H. Jr. 306
Ellison Christopher G. 137
Elvis zob. Presley Elvis
Emerson Robert M. 60, 338
Engels Fryderyk 268, 361
Eskenazi Brenda 263
Evans William 368

Feick Lawrance F. 306
Feinberg Barry 296
Ferrante Joan 437, 562
Ferris Kerry O. 60
Festinger Leon 122
Fine Gary Alan 122
Fink Jeffrey J. 384
Fischer sir Ronald A. 582
Fisher Patricia 152
Fletcher Joseph 9
Ford David A. 390, 391
Forslund Morris A. 495, 496
Foschi Martha 260, 261
Foster Bruce D. 492
Fowler Floyd J. Jr. 306
Fox Katherine 322
Frankel Linda 137, 553
Frankfort Ellen 550
Frankfort-Nachmias Chava 241, 509
Freeman Howard E. 395
Freud Sigmund 405
Fundi Mitindo, bohater programu radiowego *Twende na Wakati* 370

Gabbaya Shaul 9
Gall John 566, 567
Gallup George 52, 203, 204, 206, 220, 241, 268
Gamson William 330, 331
Garant Carol 335
Gardner Brooks Carol 60
Garfinkel Harold 60, 61, 317
Gates Bill 443
Gaventa John 324
Gawiser Sheldon R. 616
Gee Wilson 58, 59
Ghidina Marcia 9
Glaser Barney G. 78, 318, 350, 402, 403, 408, 412, 423, 627
Glock Charles Y. 9, 70, 72, 73, 102, 231, 449, 454, 470, 471, 474–476, 492
Goffman Erving 78, 404, 405, 412
Gonzalez Jose Luis 325
Gorbaczow Michaił Siergiejewicz 272
Gottlieb Bruce 185
Górniak Jarosław 586
Graham Laurie 137
Greatbatch David 61
Greenwood Davydd J. 324, 325
Greenwood Peter W. 386, 387
Griffin Larry J. 366, 367
Griffith Allison I. 323
Grimes Michael D. 170
Grove Janet Belew 368
Groves Robert M. 306
Grube Joel 264
Gubrium Jaber F. 316–318, 338
Guttman Louis 193

Hald Anders 583
Haley Alex 264
Halley Fred 458, 509, 586
Hamnett Michael P. 531, 536
Hannon Lance 62
Harding Warren 202
Hart Stephen D. 261
Hatchett Shirley 124
Haveman Ernest 469
Having Paul 437
Hawkes Roland 9
Healey Joseph F. 509
Hedrick Terry E. 374, 395
Hein Jeremy 89, 90
Heise David R.
Hempel Carl G. 146
Heritage Johen 61, 317
Higginbotham A. Leon Jr. 364
Higgins Henry, prof., bohater sztuki *Pigmalion* autorstwa George'a Bernarda Shawa 260
Hilts Philip 371, 372, 388
Hirchi Travis 104
Hite Shere 269
Hitler Adolf 150
Hogan Richard 127, 137
Hollerith Herman 425, 426
Holmes Sherlock, bohater cyklu opowiadań autorstwa Arthura Conan Doyle'a 75, 92
Holstein James A. 316–318, 338
Homan Roger 536
Hoover Kenneth R. 52
Horney Julie 388, 389
Horowitz Irving Louis 530
Howard Edward N. 373
Howard Ewa 9
Howell Joseph T. 335
Huberman A. Michael 401, 402
Hughes Michael 136
Hum Derek P. J. 199
Humphreys Laud 525, 526, 534
Humpty Dumpty, bohater książki *Alicja w krainie czarów* autorstwa Lewisa Carrolla 68
Hunt Morton 138
Hurst Leslie 321, 322
Hutchby Ian 423

Indrayan Abhaya 199
Irwin John 387
Isaac Larry W. 366, 367
Iversen Gudmund R. 138

Jackman Mary R. 221
Jacobs Bruce A. 127
Jacobson Lenore 260
Jacquard Joseph-Marie 425, 428, 432
Jacques Jeffrey 9
Jasso Guillermina 76, 77, 81
Jensen Arthur 529, 530
Jobes Patrick C. 320
Johnson Jeffrey C. 127, 207, 338
Johnston Hank 137
Jones Bill 386
Jones James H. 387
Jones Stephen R. G. 266

Kahane Howard 98, 99
Kalton Graham 241
Kalwin Jan 361, 362
Kaplan Abraham 66, 143
Kardas Edward P. 562
Kasl Stanislav V. 263
Kebede Alemseghed 80
Keeter Scott 122
Kelle Udo 338
Kendall Diana 366
Kendall Patricia L. 460–464, 468–470, 476
Kernion Jean de 10
Khayatt Didi 323
Kiereński Aleksander 361
Kilburn John C. 127
Kinloch Graham C. 552
Kinnel Ann Marie 406
Kish Leslie 241
Klenow Daniel J. 9
Kloskowska-Dudzińska Agnieszka 561

Knottnerus J. David 80
Kołakowski Leszek 361
Kopaliński Władysław 55
Korporowicz Leszek 327
Kosinski Wanda 9
Krushat W. Mark 164, 165
Kuhn Thomas S. 56, 57, 82
Kumar Rajeev 199
Kurland Daniel J. 562
Kvale Steinbar 147, 327, **330, 338**

Lambert Charles 368
Landon Alf 202, 203, 269, 620
Laslett Barbara 364, 365
Lawrence Bobo 199
Lazarsfeld Paul 9, 92, 104, 170, 199, 398, 460, 462–464, 471, 476, 628
Lazes Peter 324, 325
Leaon-Guerrero Anna 241
Lee Raymond 536
Leon Joseph J. 125
Lever Janet 62, 63
Lewis-Beck Michael 457
Li Tze-chung 553
Licari Frederick C. 199
Likert Rensis 192, 193, 270, 631
Lincoln Yvonna S. 82, 338, 423
Linton Ralph 62
Lofland John 82, 93, 119, 208, 310, 314, 329, 332, 336, 338, 400, 407, 411, 412
Lofland Lyn H. 82, 93, 119, 208, 310, 314, 329, 332, 336, 338, 400, 407, 411, 412
Lopata-Znaniecki Helena 485
Lury D. A. 137
Lynd Helen M. 320
Lynd Robert S. 320

Madison Anna-Marie 373
Mallon Mary, zwana Tyfusową Mary 479
Mama Waridi, bohaterka programu radiowego *Twende na Wakati* 370
Manning Peter K. 403
Marks Karol (właśc. Marx Karl) 15, 59, 60, 268, 361, 627
Marshall Catherine 147, 312, 313
Martin David W. 266
Martin Patricia Yancey 80, 81, 127
Marx Karl (zob. Marks Karol)
Maxwell Joseph A. 92, 138
Maynard Douglas W. 221, 406
McAlister Alfred 380
McGrane Bernard 61, 82
McIver John P. 199
McKinney Kathleen 522
Mead George Herbert 60, **62, 260**
Meadows Dennis L. 394
Meadows Donella H. 394
Menard Scott 138
Menjivar Cecilia 9
Merton Robert K. 67, 149, 463, **476**
Michałek Krzysztof 423
Miles Matthew B. 401, 402
Milford Tommy M. 562
Milgram Stanley 521, 526, 527, 534
Miller Delbert 138, 170, 199, 288, 306
Miller Jody 127
Mitchell Richard G. Jr. 273
Mitofsky Warren J. 299
Miziński Jan 361
Mkwaju, bohater programu radiowego *Twende na Wakati* 370
Model Karyn 387
Mohr Lawrence B. 509
Molen Nynke van der 401
Monroe Marilyn 403
Morgan David L. 331, 338
Morgan Lewis H. 360, 361
Morin Gaelle T. 403
Morris Joan 9
Moskowitz Milt 221
Moynihan Daniel 529, 530
Myrdal Gunnar 529

Nader Ralph 360
Naisbitt John 342
Nardi Peter 489, 490
Neuman W. Lawrence 80
Newman Jeffrey 164, 165
Newton Rae R. 457
Norman Darlene M. 373
Nowotniak Justyna 82

O'Brien Patricia 383
O'Neill Harry W. 520
Oppenheimer Robert **530**
Ostromęcka Helena 82
Øyen Else 368

Parker Charlie 15
Parkinson Northote C. **566**
Passanella Ann 199, **476**
Perinelli Philip J. 210
Perot Ross 201
Perrucci Carolyn C. 127
Peterson Ruth D. 221, 393
Peverill Squire 307
Picou J. Steven 519, 520
Piotrowski Kazimierz 260
Plutzer Eric 89
Polivka Anne E. 271
Porter Douglas J. 536
Potter Alisa 9
Powell Elwin H. 149
Presley Elvis 269
Presser Stanley 282

Quinley Harold E. 492

Radcliffe-Brown Alfred Reginald 15(
Ragin Charles C. 320
Rasinski Kenneth A. 275, 307
Rausch Mitchell 10
Ray William J. 266
Reagan Ronald 98
Regoli Mary Jean 390, 391
Reicker Henry W. 122
Reinharz Shulamit 83, 342
Reynolds Robert W. 9
Richlin-Klonsky Judith 553
Riecken Henry W. 374
Rimer Sara 232
Ringer Benjamin R. 70, 72, 231, 47(
471
Ritzer George 83
Roethlisberger Fritz J. 249
Rog Debra J. 338, 374, 395
Rogers Everett M. 370, 374
Rokeach Milton 264
Roosevelt Franklin Delano **20**
203, 269, 620
Roper Burns 520
Rosenbaum Wilf L. 507
Rosenberg Morris 119, 170, 1(
471–473, 476
Rosenthal Robert 260
Rossi Peter H. 170, 199, 241, 3(
395, 476
Rossman Gretchen B. 147, 312, 3
Roth Molly 9, 10
Rothgeb Jennifer M. 271
Rothman Ellen K. 363
Rubel Maximilien 361
Rubin Herbert J. 327, 329
Rubin Riene S. 327, 329
Rudestam Kjell Erik 457
Rusinowa Izabela 423
Rydell C. Peter 387

Sacks Jeffrey J. 164, 165
Sakumoto Tom, fikcyjne nazwisko polityka, wprowadzone do jednego z sondaży przez Earla Babbiego 274
Sanders William B. 113
Scarce Rik 519
Schachter Stanley 122
Scheuble Laurie K. 9
Schiflett Kathy L. 137
Schmitt Frederika E. 80, 81, 127
Schutz Alfred 317
Schwartz Richard D. 368
Sechrest Lee 368
Selvin Hanan 104, 213
Senter Mary Scheuer 221
Seve Michel de 473
Shaffir William B. 314, 338
Shapiro Lillian L. 550
Shaver Kelly G. 104
Shaw George Bernard 259, **260**
Sheatsley Paul F. 307
Shelton Beth Anne 9
Sherkat Darren E. 137
Shostak Arthur 338
Silverman David 100, **170, 338,** 405, 448
Singh Nihal 199
Simmel Georg 59, 60
Skinner Steven J. 288
Sloan Matthew 9
Słomczyński Maciej 68
Smith Andrew 273
Smith Dorothy F. 322, 323
Smith Eric R. A. N. 307
Smith Vicky 127
Snow David A. 127, 316, 317
Sorofman Bernard 9
Sorokin Pitirim A. 361
Spencer Herbert 58–61
Spock Benjamin 26
Spohn Cassie 388, 389
Srole Leo 149, 150
Stanley Julian 253, 255, 258, 259, 263, 265, 266
Stark Rodney 102, 492
Stearns Cindy A. 127
Stebbins Robert A. 314, 338
Steel Paul 383
Stein Gertrude 34
Sterling Theodore D. 507
Stewart Ron 9
Stoker Randy 9
Stossel John 357
Stouffer Samuel 9, 37, 38, 460–464, 466, 474, 476, 531
Strang David 137
Strauss Anselm L. 78, 318, 319, 338, 350, 400, 402, 403, 408, 411, 423, 627
Strenski Ellen 553
Strunk William Jr. 563, 569
Sudman Seymour 241, 306
Sułek Antoni 353
Summers Lawrence 357
Sumner William Graham 528
Swafford Michael 307
Swalehe Ramadhan 383
Sweet Stephen 536
Szawiel Tadeusz 353

Tacq Jacques 476
Takeuchi David 79, 81
Tan Alexis S. 221
Tandon Rajesh 381, 382
Terhanian George 299
Theologides Athanasios 221
Thomas Edmund D. 384
Thomas William I. 363
Thurstone Louis 191, 193
Tiano Susan 46
Tolnay Stewart R. 102
Tołstoj Lew 346
Tourangeau Roger 307
Truman Harry 203, 206, 241
Tuckel Peter S. 296
Turk Theresa Guminski 137
Turner Jonathan H. 67, 83
Turner Stephen Park 83
Tyfusowa Mary (pseud.) zob. Mallon Mary
Tylor Humphrey 299

Uwe Flick 339

Vaughn Angela 437, 562
Veroff Joseph 124
Votaw Carmen Delgado 137

Wachnicki Janusz 586
Wagenaar Theodore 9, 10
Wagner-Pacifici Robin 88, 118
Walker Jeffrey T. 298
Wallace Walter L. 47
Wallace William A. 104
Warner William Lloyd 150, 320
Warriner G. Keith 261
Watson John H., dr, bohater cyklu opowiadań o Sherlocku Holmesie autorstwa Arthura Conan Doyle'a 75
Webb Eugene 341, 368
Weber Max 20, 361, 362, 365, 366, 528
Weber Robert Philip 368
Weinkam James J. 507
Weiss Carol 380, 381
Wellman David 46
Wells Ida B. 342
West Patricia Salter 468, 469
Wharton Amy S. 184
White E. B. 563, 569
Whyte William Foote 316, 317, **324,** 325
Wieder D. Lawrence 318
Williams Robin M. Jr. 307
Wilson Camilo 299
Witt G. Evans 616
Wolcott Harry 99
Wolf Robert 9
Wolfe Jerome 9
Wollfitt Robin 423
Wolińska Zofia 60
Worcester Robert M. 201
Wright James D. 170, 199, **241, 307**
Wysocki M. J. 199

Yammarino Francis 288
Yates Frank 582
Yerg Beverly J. 488, 489
Yinger J. Milton 253

Zaino Jeanne 458, 509, **586**
Zangrando Robert 529
Zawirska-Michalska St. 75
Zerbib Sandrine 5, 10, **331, 417–421**
Zey Mary 137
Zhang Ray 49
Zhenchoa Qian 9
Ziesel Hans 458
Zipp John F. 89
Znaniecki Florian 363

Indeks rzeczowy*

(Opracowała Agnieszka Kloskowska-Dudzińska)

aksjomaty **67, 76, zob. także: założenia**
analiza
- a. czynnikowa **494–497**
- a. danych 13, 21–24, 132, 397–510, 545, zob. także: interpretacja
- a. dyskursu **88, 89**
- a. historyczno-porównawcze **360–367**
- – techniki a.h.-p. **364–367**
- – źródła danych w a.h.-p. **363, 364**
- a. ilościowa (statystyczna) 21–23, **438–510**, zob. także: dane ilościowe
- – dwuzmiennowa a.i. 22, **449–454**
- – jednozmiennowa a.i. 21, 22, **439–445**
- – wielozmiennowa a.i. 22, 439, **491–497**, zob. także: model analizy rozbudowanej
- a. istniejących *danych statystycznych* 20, **353–360**, 544
- – jednostki analizy w a.i.d.s. **355, 356**
- – rzetelność w a.i.d.s. **356–358**
- – trafność w a.i.d.s. **356**
- – źródła istniejących danych statystycznych **358–360**
- a. jakościowa 21, **399–423**, zob. także: dane jakościowe
- – wykorzystanie komputerów w a.j. **412–422**
- a. konwersacyjna **405, 406**
- a. kosztów i korzyści **376**
- a. pytań zob. indeks (walidacja)
- a. regresji zob. regresja (analiza r.)
- a. rozbudowana zob. model analizy rozbudowanej
- a. statystyczna zob. a. ilościowa
- a. szeregów czasowych **492–494**
- a. ścieżkowa **491, 492**
- a. treści 13, 20, **342–353**
- – dobór próby w a.t. **343–347**
- – jednostki analizy w a.t. **343–346**
- – *kodowanie* w a.t. **347–350**
- – mocne i słabe strony a.t. **352, 353**
- – tematy odpowiednie dla a.t. **342, 343**
- a. wtórna **268, 303, 304**
- a. zorientowana na przypadek **401**
- a. zorientowana na zmienną **401**

ankieta zob. badania sondażowe
- a. pocztowa **282, 285–290**, 300, zob. także: kwestionariusz do samodzielnego wypełniania
- – zwrotność a.p. zob. poziom zwrotów

ankieter **290–295**
- bezpieczeństwo a. **296**, 301
- kontrola pracy a. **293–295**
- ogólne zasady pracy a. **291–293**
- rola a. **290, 291**
- szkolenie a. **293–295**
- wygląd i zachowanie a. **291, 292**
- wypełnianie kwestionariusza przez a. zob. kwestionariusz wypełniany przez ankietera

anomia **148–151**, 355
anonimowość 50, **518**, 522
arkusz kodowy **432, 433**
artefakty 264, **302**
Asha eksperyment zob. eksperyment Asha
atrybut zmiennej zob. zmienna (wartość)
autorytet **31**

badani
- ochrona b. 14, 23, 50, 51, **516–520**, 522
- oszukiwanie b. 51, 312, **520**, 521
- stosunek do b. **314–316**

badania
- b. ankietowe zob. badania sondażowe
- b. dynamiczne **122–125**, zob. także: badania zbliżone do dynamicznych
- b. ewaluacyjne 20, 21, 36, 369–396, 544
- – etyka w b.e. **385–387**
- – ewaluacja jakościowa **383**
- – pomiar w b.e. **372–377**
- – problemy logistyczne w b. **383–385**
- – schematy badawcze w b. **377–383**
- – tematy odpowiednie dla b. **371, 372**
- – wykorzystanie wyników b. **387–391**
- b. historyczno-porównawcze z. analizy historyczno-porównawc
- b. ilościowe **49, 50**, zob. tak dane ilościowe
- b. jakościowe 19, 20, 49, 50, z. także: badania terenowe, d. jakościowe
- – opisy liczbowe w b.j. **448**
- b. kohort demograficznych 1
- b. niereaktywne 20, **340–36**
- b. panelowe **123, 124**
- b. przekrojowe **121, 122**
- b. sondażowe 19, 55, **267–543**, 544
- – elastyczność b.s. **301, 302**
- – mocne i słabe strony b.s. **302**
- – nowe technologie w b.s. **–300**
- – porównanie metod zbie danych w b.s. **300, 301**
- – rzetelność b.s. **302**
- – sztuczność b.s. **302**
- – tematy odpowiednie dl **268, 269**

* Numery stron oznaczone grubą czcionką wskazują fragmenty, w których autor bezpośrednio omawia dane zagadnienie lub temat.

– – trafność b.s. 302
– b. społeczne
– – zastosowania b.s. 24, 132, 133, 537–546
– b. terenowe 19, 20, 78, 308–339, 544
– – etyka w b.t. 312, 313, 336
– – mocne i słabe strony b.t. 334–336
– – rzetelność w b.t. 335, 336
– – tematy odpowiednie dla b.t. 309–312
– – trafność w b.t. 334, 335
– b. trendów 122, 123
– b. uczestniczące 323–325
– b. zbliżone do dynamicznych 125–127, zob. także: badania dynamiczne
– cele b. zob. cele badań
– plan b. zob. plan badań
– wymiar czasowy b. zob. wymiar czasowy badań
błąd 31–33
– b. ekologiczny rozumowania 119, 120, 355
– b. losowy zob. błąd z próby
– b. standardowy zob. błąd z próby
– b. statystyczny zob. błąd z próby
– b. z próby (losowy, statystyczny, standardowy) 218, 498, 499, 503
– – b. z p. w doborze wielostopniowym zob. dobór próby (wielostopniowy)
– b. systematyczny 498
– błędy w rozumowaniu na temat związków przyczynowych zob. przyczynowość
błędne wnioskowanie o jednostkach analizy zob. jednostki analizy (błędne wnioskowanie)
bodziec eksperymentalny 247

CAPI 298
CATI 297, 298, 433
cele badań 46, 110–113
– c.b. eksploracyjne 111, 112
– c.b. opisowe 112, 113
– c.b. wyjaśniające 113
chi-kwadrat zob. test chi-kwadrat
częstości brzegowe zob. marginesy
czterogrupowy schemat Salomona zob. schemat eksperymentalne
czynniki 494, 495
czyszczenie danych 434, 435
– cz. dotyczące dopuszczalnych kodów 434, 435
– cz. logiczne 435

dane
– analiza d. zob. analiza (danych)
– d. ilościowe 48–50, zob. także: analiza ilościowa
– d. jakościowe 48–50, zob. także: analiza jakościowa
– gromadzenie d. 107–396
– przetwarzanie d. 132
– – p.d. jakościowych 406–422
– redukcja d. 481
– wprowadzanie d. 433
darwinizm społeczny zob. paradygmat darwinizmu społecznego
debriefing zob. sesja poeksperymentalna
dedukcja
– metoda dedukcyjna 15
– teoria dedukcyjna 46–48
– – budowa t.d. 75–78
– rozumowanie dedukcyjne 67, 70–75
definicja 16, 146–148, 151–153
– d. nominalna 146, 148
– d. operacyjna 146–148
– d. realna 146, 148
– d. w badaniach opisowych 151–153
– d. w badaniach wyjaśniających 153
determinizm 85–88
dobór próby 17, 18, 131, 132, 200–244, 542
– celowy (arbitralny) d.p. 205
– d.p. metodą kuli śnieżnej 205, 206
– d.p. oparty na dostępności badanych 204, 205
– gronowy d.p. zob. dobór próby (wielostopniowy grupowy)
– grupowy d.p . zob. dobór próby (wielostopniowy grupowy)
– historia d.p. 202–204
– kwotowy d.p. 206, 207
– losowy d.p. 208–239
– – l.d.p. w eksperymencie 250, 251
– – prosty l.d.p. 224, 225
– – teoria i logika l.d.b. 208–220
– – typy l.d.p. 223–239
– nielosowy d.p. zob. dobór próby (nieprobabilistyczny)
– nieprobabilistyczny d.p. 204–208
– nieproporcjonalny d.p. 236–239
– probabilistyczny d.p. zob. dobór próby (losowy)
– proporcjonalny d.p. 234–236
– systematyczny d.p. 225
– – cykliczność w s.d.p. 226, 227
– warstwowy d.p. 227–230
– wiązkowy d.p. zob. dobór próby (wielostopniowy grupowy)
– wielostopniowy grupowy d.p. 230–239
– – liczba grup w w.g.d.p. 232–234
– – warstwowanie w w.g.d.p. 234
– – w.g.d.p. a błąd z próby 231–234
dobrowolne uczestnictwo 50, 515, 516
dojrzewanie w eksperymencie 255
dopasowywanie 18, 251–253
dopytywanie 293, 295
dyferencjał semantyczny 192, 193

efekt Hawthorne 249
eksperyment 18, 19, 245–266, 543
– dobór uczestników w e. 250–253
– e. Asha 63, 64, 521
– e. klasyczny 247–250
– e. kontrolowany zob. eksperyment laboratoryjny
– e. laboratoryjny (kontrolowany) 18, 262
– e. naturalny 246, 247, 262–264
– e. podwójnie ślepy 249, 250
– historia w e. zob. historia w eksperymencie
– mocne i słabe strony e. 264, 265
– schematy eksperymentalne zob. schemat
– sztuczność e. 264
– tematy odpowiednie dla e. 246, 247
– trafność w badaniach eksperymentalnych 248, 255–259, zob. także: nietrafność
eksplanacja w analizie rozbudowanej 465–470
eksploracja zob. cele badań (eksploracyjne)
ekstrapolacja 490
element 211
epistemologia 29
EPSEM zob. metoda równego prawdopodobieństwa wyboru

estymator 217, 218
etnografia 55, 316
– e. instytucjonalna 322, 323
etnometodologia zob. paradygmat etnometodologiczny
etyka
– e. badawcza 11, 12, 14, 23, 24, 50, 51, 513–527, zob. także: kontrola
– e. w badaniach ewaluacyjnych zob. badania ewaluacyjne (etyka)
– e. w badaniach terenowych zob. badania terenowe (etyka)
– kodeksy e. zawodowej 51, 523, 524
ewaluacja zob. badania ewaluacyjne
ewolucjonizm zob. paradygmat darwinizmu społecznego

feminizm zob. paradygmat (p. feministyczne)
fenomenologia 317
fokus zob. grupa fokusowa
funkcjonalizm strukturalny zob. paradygmat funkcjonalizmu strukturalnego

gamma 483–485, 489, zob. także: miary siły związku
generalizacje zob. uogólnianie
Generalny Sondaż Społeczny (*General Social Survey*) 16, 21, 22, 44, 303, 439, 454
grupa
– g. eksperymentalna 248, 249
– g. fokusowa 330, 331
– g. kontrolna 248, 249
GSS zob. Generalny Sondaż Społeczny

hipoteza(y) 67–72
– h. formułowane post factum 474, 475
– h. zerowa 72, 503
– testowanie (sprawdzanie) h. 15, 67–70, 72–75
historia w eksperymencie 255

idiograficzne wyjaśnianie 45–49
– przyczynowość w i.w. 45, 88–91
indeks 17, 173–190
– braki danych w i. 184–186
– budowa i. 176–183
– i. a skala 173–176
– punktacja i. 183, 184
– walidacja i. 186–189
indukcja
– i. analityczna 350
– metoda indukcyjna 15
– rozumowanie indukcyjne 67, 70–75
– teoria indukcyjna 46–48
– – budowa teorii indukcyjnej 78, 79
informator 13, 207
– wybór i. 207, 208
interpolacja 490
interpretacja 13, zob. także: analiza danych
– i. w analizie rozbudowanej 470
interwał (odstęp) losowania 225
interwencja społeczna 20, 370
istotność statystyczna 498–503, zob. także: testy istotności statystycznej

jakościowe badania terenowe zob. badania terenowe
jednostka losowania 212
jednostki analizy 16, 113–121
– błędne wnioskowanie o j.a. 119–121
– j.a. w analizie istniejących danych statystycznych zob. analiza istniejących danych statystycznych (jednostki analizy w a.i.d.s.)
– j.a. w analizie treści zob. analiza treści (jednostki analizy w a.t.)

kodeksy etyki zawodowej zob. etyka (kodeksy e. zawodowej)
kodowanie 428–434, zob. także: arkusz kodowy, książka kodowa
– k. brzegowe 433
– k. danych jakościowych 406–410
– – jednostki kodowania w k.d.j. 407
– k. otwarte 408
– k. w analizie treści zob. analiza treści (kodowanie w a.t.)
– sposoby k. 432–434
– tworzenie kategorii kodowych 348, 429–431, 408–410
komunikacyjny model oczekiwań 260
koncepcje 141, 142
konceptualizacja 16, 17, 130, 144–150
kontrola 165, 523
– k. pracy ankieterów zob. ankieter (kontrola pracy a.)
korelacja 92–95, 485, zob. także: współczynnik korelacji liniowej Pearsona
krzywa normalna 217
książka kodowa 431, 432
kwantyfikacja 48, 49, 424–437
kwestionariusz 19, 270
– budowa k. 19, 24, 276–282
– k. do samodzielnego wypełniania 282–290, 300, zob. także: ankieta pocztowa
– k. wypełniany przez ankietera 290–295, zob. także: wywiad ankieterski
– format k. 276
– instrukcje do k. 280, 281
– pilotaż k. 281, 282
– pytania k. zob. pytania kwestionariusza
– uwagi do k. 294
kwotowy dobór próby zob. dobór próby (kwotowy)

lambda 482, 483, 489, zob. także miary siły związku
liczebność
– l. obserwowana 504
– l. oczekiwana 503
linia regresji zob. regresja (linia r
listy przypominające w ankiecie pocztowej 286, 287
logika 36, 75, 78
losowanie 212
– l. systematyczne zob. dobór próby (systematyczny)

ładunki czynnikowe 494, 495

macierz
– m. korelacji 485
– m. kwotowania 206
makroteoria 57
mapa pojęciowa 412
marginesy (częstości brzegowe) 4
mediana 441–443
metoda(y)
– m. badawcza 26
– – wybór m.b. 130
– m. ilościowe zob. badania ilościowe
– m. interpretatywna zob. m analizy rozbudowanej

– m. jakościowe zob. badania jakościowe
– m. kuli śnieżnej zob. dobór próby metodą kuli śnieżnej
– m. Lazarsfelda zob. model analizy rozbudowanej
– m. połówkowa **165**
– m. rozszerzonych przypadków **320–322**
– m. równego prawdopodobieństwa wyboru (EPSEM) **211**
– m. stałego porównywania **402, 403**
– m. szkoły Columbia zob. model analizy rozbudowanej
– m. teorii ugruntowanej (MTU) **402, 403**, zob. także: teoria ugruntowana
metodologia **29**
miary siły związku **481–486**, 489
– m.s.z. dla zmiennych interwałowych i ilorazowych **485, 486**, 489
– m.s.z. dla zmiennych nominalnych **482, 483**, 489
– m.s.z. dla zmiennych porządkowych **483–485**, 489
miary złożone 13, 16, **172–199**, zob. także: wskaźniki proste i złożone
mikroteoria **57**, 60
modalna **441, 442**
model analizy rozbudowanej 22, 23, **459–477**
modernistyczne ujmowanie rzeczywistości **33, 34**

naturalizm **316, 317**
nauka 26, 29, 30, 33, 36–44, 46, 47, 56–58
– tradycyjny model n. **67–70**
nieprobabilistyczny dobór próby zob. dobór próby (nieprobabilistyczny)
nierównoważna grupa kontrolna zob. schematy quasi-eksperymentalne
nietrafność
– n. wewnętrzna 18, **255–257**
– n. zewnętrzna 18, **258, 259**
nomotetyczne wyjaśnianie **45–49**
– przyczynowość w n.w. **89–91**
notatki **411, 412**, 331–334
– n. kodowe **411**
– n. operacyjne **411**
– n. teoretyczne **411**
– notatka integrująca **411, 412**
– notatka podstawowa **411**
– notatka porządkująca **411**
nowe technologie w badaniach sondażowych zob. badania sondażowe (nowe technologie)

obciążenie próby zob. próba (obciążenie)
obiektywność 14, **63–66**, 314, 315 zob. także: rzeczywistość (obiektywna)
– o. a ideologia **528–531**
obserwacja 12, 14–16, 28, 31–36, 47, 48, 55, 58, 67–70, 73–75, 78, 132
– o. uczestnicząca **19**, 20, 24, 312
– o. wybiórcza **32**
– rejestrowanie o. **331–334**
obserwator
– rola o. **312–314**
ochrona badanych zob. badani (ochrona b.)
odchylenie standardowe 218, **444**
oferta badawcza **134, 135**
operacjonalizacja 16, 17, 19, 67–69, 130, **153–162**
operat losowania 203, **220–223**
opis 21, 41 zob. także: cele badań (opisowe), statystyka opisowa
oszukiwanie badanych zob. badani (oszukiwanie)

paradygmat(y) 12, 14, 15, **55–66**, 316–325, 465–474
– p. analizy rozbudowanej **465–474**
– p. darwinizmu społecznego **58**
– p. etnometodologiczny 57, **60, 61**, 63, **317, 318**
– p. feministyczne **62, 63**
– p. funkcjonalizmu strukturalnego **14, 61, 62**
– p. konfliktu 15, 57, **59, 60, 62**
– p. pozytywistyczny **57, 58**, 63–66
– p. symbolicznego interakcjonizmu 15, 57, **60**
– p. w jakościowych badaniach terenowych **316–325**
parametr **212**
parametryczne testy istotności **498**
PGSS zob. Polski Generalny Sondaż Społeczny
placebo **249**
plan badań 16, 32, **109–138**, 541, 542
pojęcia 16, 41, 49, 59, 60, **67**, 76, **141–145**
– p. jako konstrukty **142–144**
– wyjaśnianie pojęć zob. konceptualizacja
polityka 23, **527–533**
Polski Generalny Sondaż Społeczny **303**
pomiar 12, 13, 16, 68, 100–102, **139–171**, 542
– p. ilorazowy **159, 160**
– p. ilościowy **12**, zob. także: badania ilościowe, dane ilościowe
– p. interwałowy (przedziałowy) **157, 158**
– p. jakościowy **12**, zob. także: badania jakościowe, dane jakościowe
– p. nominalny **156**
– p. porządkowy **156, 157**
– p. przedziałowy zob. pomiar interwałowy
– poprawność p. **163**
– poziomy p. **156–160**
– precyzja p. **162, 163**
– rzetelność p. **163–166, 168, 169**
– techniki p. 13, 16
– trafność p. **166–169**
populacja 131, 132, 211, **220–223**
– p. badana **211, 212**
porównywanie podgrup **445–448**
postmodernistyczne ujmowanie rzeczywistości **33–35**, 65, 66
posttest **247, 248**
poufność **518–520**, 522
powtarzanie badań **32**, 165
poziom
– p. istotności **503**
– p. pomiaru zob. pomiar
– p. realizacji próby zob. **poziom** zwrotów
– p. ufności **219, 220**, 497
– p. zwrotów **286–288**
– – monitorowanie p.z. **286**
– – wystarczający p.z. **287, 288**
pozorna zależność zob. zależność pozorna
pozytywizm zob. paradygmat pozytywistyczny
PPS (*probability proportional to size*) zob. dobór próby (proporcjonalny)
prawa naukowe **66**

prawdopodobieństwo 30, 38, 39
– p. wyboru 210–212
– rachunek p. 212–220
preeksperymentalne schematy badawcze zob. schematy (s. preeksperymentalne)
premodernistyczne ujmowanie rzeczywistości 33
pretest 247
prezentacja wyników zob. raportowanie wyników
procentowanie w tabeli 450–453
proces atrybucji 260
proporcja losowania 225
proporcjonalna redukcja błędu (PRE) 482, 488
prosty dobór losowy zob. dobór próby (losowy prosty)
próba, zob. także: dobór próby
– błąd z próby zob. błąd (z próby)
– obciążenie p. 209, 210, 222, 223
– – o.p. w eksperymencie 256
– – o.p. w sondażach telefonicznych 295
– p. celowa zob. dobór próby (celowy)
– p. gronowa zob. dobór próby (wielostopniowy grupowy)
– p. kwotowa zob. dobór próby (kwotowy)
– p. losowa zob. dobór próby (losowy)
– – p. losowa prosta zob. dobór próby (losowy prosty)
– p. nielosowa zob. dobór próby (nieprobabilistyczny)
– p. nieprobabilistyczna zob. dobór próby (nieprobabilistyczny)
– p. reprezentatywna zob. reprezentatywność
– p. udziałowa zob. dobór próby (kwotowy)
– p. warstwowa zob. dobór próby (warstwowy)
– p. wiązkowa zob. dobór próby (wielostopniowy grupowy)
– p. zespołowa zob. dobór próby (wielostopniowy grupowy)
– rozkład z prób zob. rozkład (z prób)
– wielkość p. 218
przeciętna wartość zob. wartość przeciętna
przedział ufności 219–220, 497
przekrojowa analiza przypadków 401, 402
przewidywanie 30, 69
– p. wartości zmiennej 44
przyczynowość 12, 15, 30, 43, 45, 84–106
– błędy w rozumowaniu na temat związków przyczynowych 97–100
– kryteria p. 91–95
– modele przyczynowe 44, 491
– p. w idiograficznych modelach wyjaśniania zob. idiograficzne wyjaśnianie
– p. w nomotetycznych modelach wyjaśniania zob. nomotetyczne wyjaśnianie
– ustalanie związków przyczynowych 18, 21, 72, 91–95, 100–102
przyczyny 15, 45, 69, 72, 85–87, 96
– p. konieczne 96
– p. wystarczające 97, 98
pytania kwestionariusza 269–276
– format p. 270, 271, 276–279
– formy przeczące w p.k. 274
– kolejność p. 279, 280
– p. filtrujące 277
– p. i stwierdzenia 270
– p. obciążone 275, 276
– p. otwarte 270
– p. podwójne 271, 272
– p. sugerujące 275
– p. tabelaryczne 279
– p. warunkowe 277–279
– p. zamknięte 270, 271

quasi-eksperyment zob. schematy quasi-eksperymentalne

rachunek prawdopodobieństwa zob. prawdopodobieństwo (rachunek p.)
randomizacja 251–253
raportowanie (prezentacja) wyników 521, 523, 545
redukcja danych zob. dane
redukcjonizm 120, 121
regresja
– analiza r. 486–491
– – a.r. z przesunięciem czasowym 494
– – ograniczenia a.r. 490
– linia r. 486–488
– r. cząstkowa 489
– r. do średniej w eksperymencie 256
– r. krzywoliniowa 489
– r. liniowa 486–488
– r. wielokrotna 488, 489
– równanie r. 486–488
rekompensata w eksperymencie 256–257
replikacja zob. powtarzanie badań
– r. w analizie rozbudowanej 465, 466
reprezentatywność 210–212
respondent zob. badani
rozkład
– r. częstości 439–441
– r. z prób 213–219
– – odchylenie standardowe r. z p. 218
rozpowszechnianie lub naśladowanie w eksperymencie 256
rozproszenie 442, 443
równanie regresji zob. regresja (równanie r.)
rzeczywistość 33–36, 64–66, 141, 142
– rz. obiektywna 35, 65
rzetelność 163–166
– rz. a trafność 167–169
– rz. w analizie istniejących danych statystycznych zob. analiza istniejących danych statystycznych (rzetelność w a.i.d.s.)
– rz. w badaniach sondażowych zob. badania sondażowe (rzetelność b.s.)
– rz. w badaniach terenowych zob. badania terenowe (rzetelność b.t.)

schemat(y)
– s. eksperymentalne 253–259
– – czterogrupowy s. Solomona 258, 259
– – s.e. w badaniach ewaluacyjnych 377, 378
– s. preeksperymentalne 253–255
– – s. z dwukrotnym pomiarem w jednej grupie 253–255
– – statyczne porównanie międzygrupowe 253–255
– – studium przypadku z pojedynczym pomiarem 253–255
– s. quasi-eksperymentalne w badaniach ewaluacyjnych 378–383
– – nierównoważna grupa kontrolna w s.q.-e. 380

– – s. z uwzględnieniem szeregów czasowych **378–380**
– – s. z uwzględnieniem wielu szeregów czasowych **380–383**
– s. bez pretestu z grupą kontrolną **259**
semiotyka **403–405**
sesja poeksperymentalna (*debriefing*) **521**
skala 17, 173–176, **190–196**
– budowa s. **190**
– s. a indeks **173–176**
– s. dystansu społecznego Bogardusa **190**, **191**
– s. Guttmana **193–196**
– s. Likerta **191**, 192, 270
– s. Thurstone'a **191**
sondaż zob. badania sondażowe
– pozorne s. 269, 296
– s. faksowe **298**, **299**
– s. online zob. sondaże przez Internet
– s. przez Internet (s. online) 99, 200, **301**
– s. telefoniczne **295–298**, 301, zob. także: CATI
– –s.t. a automatyczne sekretarki **296**
– s. wyborcze 17–19, 55
– sondażowe przepychanie **269**, 386, 532
specyfikacja w analizie rozbudowanej **470**, **471**, 473
statystyka 23, **478–510**
– s. jako estymator parametru **216**, 219
– s. opisowa **480–490**
– wnioskowanie statystyczne **497–508**
– – wnioskowanie jednozmiennowe **497**, **498**
stopnie swobody **504**, 505
stratyfikacja zob. dobór próby (warstwowy)
studium przypadku 40, **320**
subiektywność **14**
symboliczny interakcjonizm zob. paradygmat symbolicznego interakcjonizmu
syndrom Pigmaliona **260**
szeregi czasowe zob. analiza szeregów czasowych

średnia **441**, **442**
świadoma zgoda zob. zasada świadomej zgody

tabela krzyżowa (kontyngencji) **453–456**, zob. także: procentowanie w tabeli
tabele cząstkowe **465**
tablice liczb losowych **224**, 225
tendencja centralna **441–443**
teoria 14, 15, 38–48, **54–83**
– konstruowanie t. 66, 67, 75–79
– t. a badania 12, 14, 15, **79–81**, 400–406, 538, 539
– t. dedukcyjna zob. dedukcja (teoria dedukcyjna)
– t. *indukcyjna zob. indukcja* (teoria indukcyjna)
– t. sprawiedliwości dystrybutywnej Jasso **76–78**
– t. systemów społecznych zob. paradygmat funkcjonalizmu strukturalnego
– t. ugruntowana 78, **318–320**, zob. także: metoda teorii ugruntowanej
test
– t. chi-kwadrat **503–506**
– t. połówkowy zob. metoda połówkowa
– t. powtórny **164**, 165
– testy istotności statystycznej **498**, 499, 503–508
testowanie hipotez zob. hipoteza (testowanie h.)
tezy **67**, 76
tradycja **30**, 31
trafność **166–169**, zob. także: nietrafność
– t. a rzetelność zob. rzetelność (rz. a trafność)
– t. fasadowa **166**
– t. kryterialna **166**
– t. teoretyczna **166**, **167**
– t. treściowa **168**
– t. w analizie istniejących danych statystycznych zob. analiza istniejących danych statystycznych (trafność w a.i.d.s.)
– t. w badaniach eksperymentalnych zob. eksperyment (trafność w badaniach eksperymentalnych)
– t. w badaniach sondażowych zob. badania sondażowe (trafność b.s.)
– t. w badaniach terenowych zob. badania terenowe (trafność w b.t.)
treść
– t. jawna **347**, **348**
– t. ukryta **347**, **348**
twierdzenia 66, 67
typologia 17, **196**, **197**

uogólnianie (generalizacje) **32**, **46**
– u. w eksperymencie **250**

walidacja
– w. indeksu zob. indeks (walidacja i.)
– w. zewnętrzna **187–188**
wariancja
– w. całkowita **486**
– w. niewyjaśniona **488**
– w. wyjaśniona **489**
warstwowanie zob. dobór próby (warstwowy)
– w. w wielostopniowej próbie grupowej zob. dobór próby (wielostopniowy grupowy)
wartość
– w. przeciętna **441**, **442**
– w. zmiennej zob. zmienna (wartość z.)
ważenie **236**, **237**
wielostopniowy dobór grupowy zob. dobór próby (wielostopniowy grupowy)
wnioskowanie statystyczne zob. statystyka (wnioskowanie statystyczne)
wprowadzanie danych zob. dane (*wprowadzanie d.*)
wskaźniki **144–146**
– w. i wymiary **144**, **145**, 173
– w. proste i złożone **160**, 161
– w. społeczne 371, **391–394**
– wymienialność w. 145, 146, **165**
współczynnik korelacji liniowej Pearsona **485**, **486**, zob. także: miary siły związku
współczynniki ścieżkowe **491**
wyjaśnianie 15, 16, 22, 41, **45**, **46**, 55, 66, 67
– w. idiograficzne zob. idiograficzne wyjaśnianie
– w. nomotetyczne zob. nomotetyczne wyjaśnianie
wyjątek 32, 38
wymiar, zob. także: wskaźniki
– w. czasowy badań **121–127**
– w. pojęć 145, **146**

wymieranie grup w eksperymencie 256
wysyłka przypominająca zob. listy przypominające w ankiecie pocztowej
wytwory społeczne 118, 119
wywiad
– w. ankieterski (bezpośredni, osobisty) 290, 300, 301, zob. także: kwestionariusz wypełniany przez ankietera
– w. bezpośredni zob. wywiad ankieterski
– w. jakościowy 327–330
– w. osobisty zob. wywiad ankieterski
– w. osobisty wspomagany komputerowo zob. CAPI
– w. telefoniczny zob. sondaże telefoniczne
– w. telefoniczny wspomagany komputerowo zob. CATI

zakres
– z. międzykwartylowy 444
zależność zob. zmienna (związki między zmiennymi)
– z. cząstkowa 465, 469, 470
– z. początkowa 465
– z. pozorna 22, 93, 94, 466–470
– z. zerowa zob. zależność początkowa
założenia 33, 66, 67, zob. także: aksjomaty
zapisywanie notatek 411, 412
zasada świadomej zgody 517, 522
zastosowania badań społecznych zob. badania społeczne (zastosowania b.s.)
zmienna 13, 15, 21, 22, 40–45, 67, 68, 76, 152, 155, 156
– definiowanie z. 155, 156
– pomiar z. 100–102
– wartość (atrybut) z. 41, 42, 45, 67, 155, 156
– z. ciągła 444
– z. dyskretna (skokowa) 444
– z. kontrolna zob. z. testowa
– z. niezależna 18, 44, 465
– – z.n. w eksperymencie 247
– z. skokowa zob. z. dyskretna
– z. stratyfikacyjna zob. dobór próby (warstwowy)
– z. testowa (kontrolna) 22, 464 473
– – z.t. pośrednicząca 465
– – z.t. uprzednia 465
– z. tłumiąca 472
– z. zależna 18, 44, 465
– – z.z. w eksperymencie 247
– z. zniekształcająca 473
– związki (zależności) między z 22, 23, 41–45, 73, 74, 76, 100–102
– – miary siły z.m.z. zob. miary siły związku
zmienność 153–155
– zakres z. 153, 154
zniechęcenie w eksperymencie 257
zogniskowany wywiad grupowy zob. grupa fokusowa
związki przyczynowe zob. przyczynowość
zwrotność ankiet zob. poziom zwrotów

