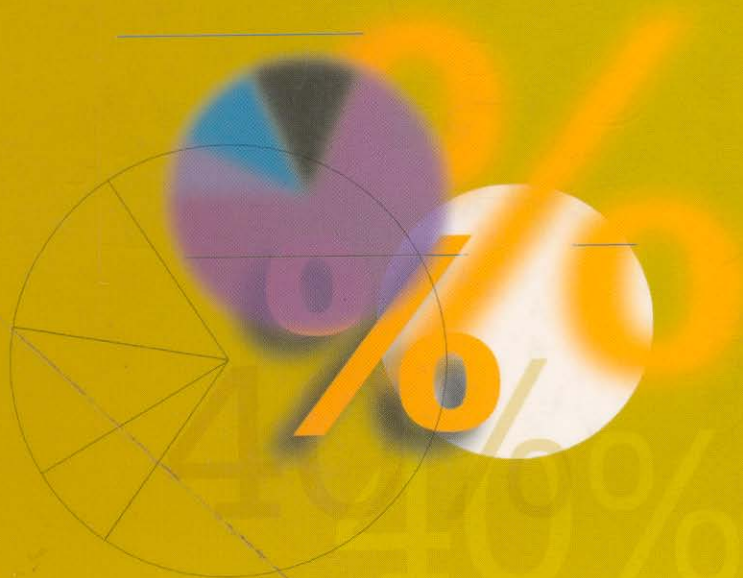


Métodos de Pesquisas de Survey

Earl Babbie



EDITORA
UFMG

As ciências sociais brasileiras, em particular a sociologia e a ciência política, são há muito assombradas por problemas de método, objeto de permanente mitificação e ocasionais mistificações. Apesar do tempo e do esforço dedicados ao tema, ele costuma ser tomado em termos de palavrosas discussões de filosofia da ciência e do confronto das "posturas" correspondentes — enquanto o trabalho substantivo de nossos sociólogos e cientistas políticos continua a caracterizar-se amplamente seja pelo velho ensaísmo, seja por relatos em que a pobreza analítica é freqüentemente erigida em virtude, reclamando-se uma suposta perspectiva "historiográfica", "antropológica" ou mesmo "jornalística".

Este volume de Earl Babbie trata de maneira a um tempo singela e competente do feijão-com-arroz envolvido naquilo que veio a ser a forma mais canônica de coleta e tratamento de dados nas ciências sociais, ou seja, a pesquisa de tipo *survey*. Os temas cobertos se estendem da inserção do *survey* no contexto da pesquisa científica ao desenho ou a concepção dos projetos, a coleta e o processamento dos dados e os recursos aí utilizados, até o instrumental envolvido na análise apropriada do material coletado.

Se o treinamento na análise de dados de *survey* representa, por si mesmo, provavelmente a melhor forma de introduzir o estudante na lógica do trabalho científico em geral, além de aparelhá-lo com um conjunto específico de instrumentos, o livro de Babbie tem o mérito de dar ênfase justamente à explicitação cuidadosa da lógica básica do *survey* e da análise multivariada. Feita de maneira simples e acessível, essa explicitação certamente fornece ao estudante um seguro ponto de partida para o estudo posterior de técnicas mais avançadas na mesma linha e boa referência crítica para a avaliação de perspectivas e técnicas alternativas.

19

dem

**MÉTODOS DE
PESQUISAS
DESURVEY**

545147

Earl Babbie

**MÉTODOS DE
PESQUISAS
DE SURVEY**

Tradução de
Guilherme Cezarino

2ª reimpressão

Belo Horizonte
Editora UFMG
2003

B.0071533

Universidade de Brasília
 e. Ponks LIVROS
 23/01/04 R\$ 36,00
 04/00468-0

545147

3: 001, 8
B 112 m
= 690

Copyright© 1997 by International Thomson Publishing Inc.
 Título original: *Survey Research Methods* (second edition)
 Copyright© 1999 da tradução brasileira by Editora UFMG
 2001 - 1ª reimpressão
 2003 - 2ª reimpressão.

Todos os direitos reservados
 Nenhuma parte deste livro pode ser reproduzida ou transmitida por qualquer forma ou meio, eletrônico ou mecânico, incluindo fotocópia, impressão ou através de qualquer sistema de arquivo e recuperação, sem autorização escrita do Editor.

PROJETO GRÁFICO E CAPA: Glória Campos - *Mangá*
 EDITORAÇÃO DE TEXTO: Ana Maria de Moraes
 REVISÃO DE TEXTO E NORMALIZAÇÃO: Olga Maria Alves de Sousa
 REVISÃO DE PROVAS: Lilian Valderez Felício e Maria Aparecida Ribeiro
 REVISÃO TÉCNICA DA TRADUÇÃO: Ronaldo de Noronha
 COORDENAÇÃO: Renarde Freire Nobre
 APOIO À COORDENAÇÃO: Magda de Pinho Tavares
 PRODUÇÃO GRÁFICA: Warren de Marilac Santos
 FORMATAÇÃO: Eduardo Ferreira, Warren de Marilac Santos e Marco Severo
 EDITORA UFMG
 Av. Antônio Carlos, 6627 - Ala direita da Biblioteca Central - Térreo
 Campus Pampulha - 31270-901 - Belo Horizonte/MG
 Tel. (31) 3499-4650
 Fax: (31) 3499-4768
 e-mail: editora@ufmg.br
 www.editora.ufmg.br

UNIVERSIDADE FEDERAL DE MINAS GERAIS
 REITORA: Ana Lúcia Almeida Gazzola
 VICE-REITOR: Marcos Borato Viana
 CONSELHO EDITORIAL

Titulares: Antônio Luiz Pinho Ribeiro, Beatriz Rezende Dantas, Carlos Antônio Leite Brandão, Heloísa Maria Murgel Starling, Luiz Otávio Fagundes Amaral, Maria das Graças Santa Bárbara, Maria Helena Damasceno e Silva Megale, Romeu Cardoso Guimarães, Wander Melo Miranda (Presidente)

Suplentes: Cristiano Machado Gontijo, Denise Ribeiro Soares, Leonardo Barci Castriota, Lucas José Bretas dos Santos, Maria Aparecida dos Santos Paiva, Maurílio Nunes Vieira, Newton Bignotto de Souza, Reinaldo Martiniano Marques, Ricardo Castanheira Pimenta Figueiredo

B112m Babbie, Earl
 Métodos de pesquisas de Survey/Earl Babbie; tradução de Guilherme Cezarino.- Belo Horizonte: Ed. UFMG, 1999.
 519p. - (Coleção Aprender)
 Tradução de : Survey research methods
 1. Ciências Sociais - Metodologia 2. Pesquisa social - Metodologia
 I. Título II. Série

CDD: 300
 CDU: 303.4

Catálogo na publicação: Divisão de Planejamento e Divulgação da Biblioteca Universitária - UFMG

ISBN: 85-7041-175-8

A escolha do livro *Survey Research Methods*, de Earl Babbie, para publicação em língua portuguesa, deu-se a partir de uma preocupação compartilhada por alguns colegas sociólogos que, na condição de professores, sentiam falta de um material didático mais completo para o ensino da pesquisa quantitativa nas ciências sociais, em especial o *survey*. Portanto, o livro visa atender a uma expectativa claramente didática, tendo sido avaliado como um bom manual de *survey*, que aborda diversos aspectos metodológicos do modelo e contém exemplificações acompanhadas de boa ilustração (gráficos, figuras, tabelas). Espero que tenhamos feito uma escolha de fato útil para uma área de pesquisa de reconhecida validade e recorrente uso nas ciências sociais.

Esta tradução que se oferece ao público de língua portuguesa é resultado de um projeto do Departamento de Sociologia e Antropologia e do Mestrado em Sociologia, da Universidade Federal de Minas Gerais — UFMG, com financiamento da CAPES, através do Programa PROIN/97, e sob gerenciamento financeiro da Fundação de Desenvolvimento da Pesquisa - FUNDEP. Agradeço às várias pessoas que estiveram envolvidas no processo, em particular à profa. Solange de Deus Simões, pelos contatos iniciais e entendimentos quanto ao direito de tradução.

O coordenador

Dedicado a Sam Stouffer, Paul Lazarsfeld,
Charles Glock e aos velhos turcos

S u m á r i o

LISTA DE TABELAS	17
LISTA DE FIGURAS	19
PREFÁCIO	21
PREFÁCIO À PRIMEIRA EDIÇÃO	29
O CONTEXTO CIENTÍFICO DA PESQUISA DE <i>SURVEY</i>	35
CAPÍTULO 1	
A LÓGICA DA CIÊNCIA	37
A Perspectiva Tradicional	37
A Desmistificação da Ciência	39
A Ciência na Prática	41
O que É Ciência?	43
Resumo	54
Leituras Adicionais	55
CAPÍTULO 2	
A CIÊNCIA E AS CIÊNCIAS SOCIAIS	57
A Busca de Regularidades Sociais	58
As Características das Ciências Sociais	62
Métodos de Pesquisas Científico-Sociais	67
Resumo	75
Leituras Adicionais	76
CAPÍTULO 3	
PESQUISA DE <i>SURVEY</i> COMO MÉTODO DAS CIÊNCIAS SOCIAIS	77
Breve História da Pesquisa de <i>Survey</i>	78

Características Científicas da Pesquisa de <i>Survey</i> ...	82
Comparação do <i>Survey</i> com Outros Métodos	86
A Pesquisa de <i>Survey</i> É Realmente Científica?	87
Leituras Adicionais	91
DESENHO DE PESQUISA DE <i>SURVEY</i>	93
CAPÍTULO 4	
TIPOS DE DESENHOS DE PESQUISAS	95
Finalidades da Pesquisa de <i>Survey</i>	95
Unidades de Análise	98
Desenhos Básicos de <i>Surveys</i>	101
Variações dos Desenhos Básicos	105
Escolhendo o Desenho Adequado	108
Resumo	110
Leituras Adicionais	110
CAPÍTULO 5	
A LÓGICA DA AMOSTRAGEM DO <i>SURVEY</i>	113
Lógica da Amostragem de Probabilidade	118
Conceitos e Terminologia de Amostragem	120
Teoria da Amostragem Probabilística e	
Distribuição Amostral	125
Populações e Molduras de Amostragem	131
Tipos de Desenhos de Amostragem	134
Amostragem e Ponderação Desproporcionais	148
Amostragem Não-Probabilística	152
Usos Não- <i>Survey</i> dos Métodos de Amostragem	156
Resumo	157
Leituras Adicionais	158
CAPÍTULO 6	
EXEMPLOS DE DESENHOS DE AMOSTRAGEM	159
Amostragem de Estudantes Universitários	159
Amostragem dos Professores de Escolas	
de Medicina	161

Amostragem de Mulheres de Igrejas	
Episcopais	162
Amostragem de Casas em Oakland	165
Resumo	176
Leituras Adicionais	177
CAPÍTULO 7	
CONCEITUAÇÃO E DESENHO DE	
INSTRUMENTOS	179
Lógica da Conceituação	179
Um Quadro de Referências de	
Operacionalização	182
Tipos de Dados	185
Níveis de Medição	187
Guias para a Elaboração de Questões	189
Qualidade das Medições	194
Formato Geral dos Questionários	198
Ordenando as Questões no Questionário	205
Instruções	206
Reprodução do Questionário	209
Resumo	210
Leituras Adicionais	211
CAPÍTULO 8	
CONSTRUÇÃO DE ÍNDICES E ESCALAS	213
Índices Versus Escalas	214
Construção de Índices	217
Construção de Escalas	233
Tipologias	240
Resumo	243
Leituras Adicionais	244
COLETA DE DADOS	
	245
CAPÍTULO 9	
QUESTIONÁRIOS AUTO-ADMINISTRADOS	247
Distribuição e Retorno de Correspondência	248

Opções Postais e Custos Relativos	249
Monitorando os Retornos	250
Correspondências de Acompanhamento	251
Taxas Aceitáveis de Resposta	253
Um Caso de Pesquisa	254
Resumo	257
Leituras Adicionais	258
CAPÍTULO 10	
<i>SURVEYS</i> POR ENTREVISTAS	259
Importância do Entrevistador	259
Regras Gerais para Entrevistar	261
Treinamento do Entrevistador	266
A Operação de Entrevistar	270
Resumo	275
Leituras Adicionais	276
CAPÍTULO 11	
PROCESSAMENTO DOS DADOS	279
Computadores na Pesquisa de <i>Survey</i>	280
Codificação	288
Construção do Livro de Códigos	291
Opções de Codificação e Inserção de Dados	293
Pré-Codificação para a Inserção de Dados	297
Limpeza dos Dados	299
Resumo	300
Leitura Adicionais	301
CAPÍTULO 12	
PRÉ-TESTES E ESTUDOS-PILOTO	303
Fazendo os Pré-Testes	303
Fazendo Estudos-Piloto	310
Avaliando Pré-Testes e Estudos-Piloto	313
Resumo	321

ANÁLISE DA PESQUISA DE <i>SURVEY</i>	323
CAPÍTULO 13	
LÓGICA DA MEDIÇÃO E DA ASSOCIAÇÃO	327
A Imagem Tradicional	327
Intercambialidade de Índices	331
Implicações	333
Resumo	334
Leituras Adicionais	335
CAPÍTULO 14	
CONSTRUINDO E COMPREENDENDO TABELAS	337
Análise Univariada	337
Descrições de Subgrupos	344
Análise Bivariada	349
Análise Multivariada	357
Resumo	360
Leituras Adicionais	361
CAPÍTULO 15	
O MODELO DE ELABORAÇÃO	363
História do Modelo de Elaboração	363
O Paradigma de Elaboração	370
Elaboração e Hipótese <i>Ex Post Facto</i>	379
Resumo	381
Leituras Adicionais	382
CAPÍTULO 16	
ESTATÍSTICA SOCIAL	383
Estatística Descritiva	383
Estatística Inferencial	391
Resumo	405
Leituras Adicionais	407
CAPÍTULO 17	
TÉCNICAS MULTIVARIADAS AVANÇADAS	409
Análise de Regressão	409
Análise de Trajetória	415
Análise Fatorial	417

Análise de Variância	421
Análise Discriminante	424
Modelos Log-Lineares	428
Resumo	430
Leituras Adicionais	432

CAPÍTULO 18

O RELATÓRIO DA PESQUISA DE <i>SURVEY</i>	433
Algumas Considerações Básicas	433
Organização do Relatório	436
Diretrizes para Relatar as Análises	441
Resumo	443
Leituras Adicionais	443

A PESQUISA DE <i>SURVEY</i> NO CONTEXTO SOCIAL	445
--	-----

CAPÍTULO 19

A ÉTICA NA PESQUISA DE <i>SURVEY</i>	447
Participação Voluntária	448
Sem Prejuízo para os Entrevistados	450
Anonimato e Sigilo	451
Identificação da Finalidade e do Patrocinador	454
Análise e Relatório	455
Um Código de Ética Profissional	456
Ética – Ilustrações Relevantes	460
Resumo	463
Leituras Adicionais	463

CAPÍTULO 20

O CONSUMIDOR INFORMADO DE PESQUISA DE <i>SURVEY</i>	465
Desenho da Pesquisa	466
Medição	467
Amostragem	469
Análise dos Dados	470
Relatório dos Dados	471
Leituras Adicionais	471

BIBLIOGRAFIA	473
--------------------	-----

APÊNDICES	481
-----------------	-----

A. Tabela de Números Aleatórios	482
B. Erro Estimado de Amostragem para uma Binomial (Nível de Confiança de 95%)	486
C. Distribuição do Qui-Quadrado	488
D. Áreas Normais de Curva	490

GLOSSÁRIO	491
-----------------	-----

ÍNDICE REMISSIVO	511
------------------------	-----

Lista de Tabelas

TABELA 5-1. Resultados das pesquisas de boca de urna: novembro de 1988	116
TABELA 5-2. Pesquisas de opinião política antes da eleição de 1988	117
TABELA 6-1. Forma usada na lista das igrejas	163
TABELA 6-2. Estratificação de quarteirões com crescimento na área de pobreza D	171
TABELA 6-3. Procedimento de seleção dos quarteirões	172
TABELA 6-4. Amostra de folha de listagem	173
TABELA 8-1. Validação do índice de orientação científica	230
TABELA 8-2. Escalonamento da orientação científica	237
TABELA 8-3. Escores de índice e escala	239
TABELA 8-4. Uma tipologia política	240
TABELA 8-5. Nove estilos de vida americanos	242
TABELA 14-1. Uma ilustração de análise univariada	339
TABELA 14-2. Atitude com relação às Nações Unidas: "Como a ONU está resolvendo os problemas que ela tem que enfrentar?"	346
TABELA 14-3. Fundindo categorias extremas	347
TABELA 14-4. Omitindo os "não sei"	349
TABELA 14-5. "Você concorda ou discorda da proposição de que homens e mulheres devem ser tratados igualmente em todos os aspectos?"	350
TABELA 14-6. Direção oposta da porcentagem	350
TABELA 14-7. "Você aprova ou desaprova a proposição geral de que homens e mulheres devem ser tratados igualmente em todos os aspectos?"	357

TABELA 14-8. "Você aprova ou desaprova a proposição geral de que homens e mulheres devem ser tratados igualmente em todos os aspectos?" ...	359
TABELA 14-9. "Você concorda ou discorda da proposição geral de que homens e mulheres devem ser tratados igualmente em todos os aspectos?"	360
TABELA 15-1. Resumo dos dados de Stouffer sobre escolaridade e aceitação do recrutamento	366
TABELA 15-2. Relação hipotética entre escolaridade e dispensa de amigos	367
TABELA 15-3. Relação hipotética entre a dispensa de amigos e a aceitação do próprio recrutamento	368
TABELA 15-4. Dados hipotéticos relacionando escolaridade à aceitação do recrutamento, pelo fator de ter amigos dispensados	369
TABELA 15-5. O paradigma da elaboração	371
TABELA 15-6. Região de origem, região de treinamento e atitudes para com o Medicare	373
TABELA 15-7. Classe social e envolvimento médio com a igreja entre mulheres episcopais	375
TABELA 15-8. Classes sociais e a ocupação de cargos em organizações seculares	375
TABELA 15-9. Envolvimento na igreja por classe social e ocupação de cargos seculares	376
TABELA 16-1. Matriz de dados brutos parciais	384
TABELA 16-2. Dados brutos hipotéticos sobre escolaridade e preconceito	385
TABELA 16-3. Dados hipotéticos relacionando sexo e emprego	386
TABELA 16-4. Dados hipotéticos relacionando classe social e preconceito	388
TABELA 16-5. Escala de diferenciação semântica	390
TABELA 16-6. Associações gamma entre os itens de diferenciação semântica da escala de santificação	390
TABELA 16-7. Ilustração hipotética de qui-quadrado	401
TABELA 17-1. Análise fatorial: atos de delinqüência (brancos)	420

L i s t a d e f i g u r a s

FIGURA 5-1. Porcentagem de alunos que aprovam o código estudantil	127
FIGURA 5-2. Porcentagem de alunos que aprovam o código estudantil	127
FIGURA 5-3. Porcentagem de alunos que aprovam o código estudantil	127
FIGURA 7-1. Formatos de respostas	200
FIGURA 7-2. Você já fez parte da associação local da APP?	201
FIGURA 7-3. Você já ouviu falar do Programa Cidade Modelo? ..	202
FIGURA 7-4. Você trabalhou alguma vez durante a semana passada? (Inclua trabalho parcial, mas não trabalho em casa.)	203
FIGURA 7-5. Formato de questão matricial	204
FIGURA 8-1. Índices versus escalas	216
FIGURA 8-2. Relações bivariadas entre itens sobre orientação científica	221
FIGURA 8-3. Relações trivariadas entre os itens sobre orientação científica	223
FIGURA 8-4. Relação trivariada hipotética entre itens de orientação científica	225
FIGURA 11-1. Cartão perfurado padrão para registrar dados	282
FIGURA 11-2. Exemplo parcial de um livro de códigos	292
FIGURA 11-3. Uma folha parcial de transferência de códigos	294
FIGURA 11-4. Ilustração da inserção de dados computadorizada	295

FIGURA 11-5. Ilustração da pré-codificação de um questionário	298
FIGURA 13-1. A imagem tradicional da ciência	329
FIGURA 13-2. A intercambialidade de índices	332
FIGURA 14-1. Três “médias”	341
FIGURA 14-2. Percentuando uma tabela	354
FIGURA 15-1	370
FIGURA 15-2	370
FIGURA 16-1. Uma população hipotética de homens e mulheres que são favoráveis ou se opõem à igualdade sexual	395
FIGURA 16-2. Uma amostra representativa	396
FIGURA 16-3. Uma amostra não representativa	398
FIGURA 16-4. Uma amostra representativa de uma população na qual as variáveis se relacionam	399
FIGURA 17-1. Mapa simples dos valores de x e y	410
FIGURA 17-2. Gráfico dos valores de duas variáveis com linha de regressão (hipotética)	411
FIGURA 17-3. Diagrama das fontes religiosas do anti-semitismo	417
FIGURA 17-4. Dois padrões de distribuição de renda de Republicanos e Democratas	422
FIGURA 17-5. Seis escritores: três escrevem a mão e três com computadores	425
FIGURA 17-6. Projeção gráfica dos seis escritores em termos de idade e renda	425
FIGURA 17-7. Apenas renda é suficiente para predizer o método de escrever	426
FIGURA 17-8. Padrão ligeiramente mais complicado	426
FIGURA 17-9. Separando canetas dos computadores	427
FIGURA 17-10. Projeção dos seis escritores em nova dimensão	428
FIGURA 19-1. Código de ética e práticas profissionais	456

P r e f á c i o

Este livro é a segunda edição do original publicado há dezesseis anos. Foi meu primeiro manual e sempre ocupou um lugar especial no meu coração, apesar de há muito tempo não lhe dedicar atenção.

Comecei a lecionar pesquisa de *survey* em 1968, na Universidade do Havá. Na época, aprendi o que outros professores da matéria já sabiam: não havia nenhum manual realmente bom para se dar este curso. Escolhi usar um livro britânico de C. A. Moser, mas meus alunos freqüentemente não conseguiam se identificar com os exemplos britânicos, nem com o modo esquisito dos britânicos usarem a língua inglesa! (Só bem mais tarde me ocorreu que estudantes não-americanos poderiam ter problema semelhante ao usarem textos americanos.)

Quando comecei a conhecer outros professores da mesma disciplina (é incrível o número destes professores que passam pelo Havá a caminho de algum outro lugar), fui descobrindo um padrão interessante. Sempre nos fazíamos duas perguntas, e quase sempre as respostas eram as mesmas.

A primeira era: “Você vai fazer pesquisa na sua sala de aula este ano?” A resposta em geral era uma versão de: “Não. Fiz uma recentemente e tão cedo não quero passar pela mesma experiência.”

Segunda pergunta: “Que manual você está usando?” Os livros até que variavam, mas a resposta sempre terminava com: “... porém não gosto dele.”

Minha própria frustração neste aspecto fez com que, certa vez, passasse um dia inteiro imaginando o índice “ideal” de um manual sobre pesquisas de *survey*. O índice que então esbocei é essencialmente o mesmo deste livro.

O título que dei àquela obra imaginária reflete uma frustração mais específica minha com o acervo mais amplo de textos publicados sobre pesquisas de *survey*. Alguns eram

bem abstratos, ficando o leitor perdido sobre como de fato *fazer* uma pesquisa. Outros eram justamente o contrário, dando receitas concretas de passos a serem dados na realização de uma pesquisa específica. Tudo bem com esta última abordagem, contanto que você quisesse fazer exatamente a mesma pesquisa que o autor do texto. Mas se seus desejos e/ou circunstâncias fossem diferentes, você ficaria perdido.

Intitulei meu livro fictício de *A Survey Research Cookbook and Other Fables* (Livro de Receitas de Pesquisas de Survey e Outras Fábulas), título que apareceu em esboços de capítulos e correspondências que troquei durante a criação do livro. Quando o manuscrito finalmente foi para a gráfica, o título já fora enterrado no cemitério das prosas nunca usadas — o qual, na minha experiência, é o merecido descanso final de um acúmulo quase infinito de desperdício verbal.

Então eu já havia imaginado o índice de um livro ainda fictício, com um título um tanto jocoso. Pus tudo isto de lado, por um dia só, pois no dia seguinte recebi uma carta de Jack Arnold, editor de Sociologia da *Wadsworth Publishing Company*, dizendo do interesse da editora em publicar um manual de pesquisas de *survey*. Além do mais, um de seus consultores editoriais, Rod Stark, havia sugerido que fosse eu o autor deste livro.

Dois anos antes, Rod e eu éramos amicíssimos no curso de pós-graduação da Universidade da Califórnia, em Berkeley, e considerei a hipótese de que a carta de Jack não passasse de uma brincadeira. Sempre desconfiei que Rod era o tipo capaz deste tipo de piada, embora nunca houvesse feito nenhuma assim comigo. Apesar da desconfiança, respondi à carta da Wadsworth, anexando o índice que havia imaginado para o livro. Pouco tempo depois, assinei o contrato do livro e me lancei ao trabalho de escrever os capítulos, seguindo o índice esboçado.

Escrever o livro foi uma experiência ambivalente. Às vezes, as palavras pareciam fluir rápida e naturalmente, bastando apenas colocar papel em branco na máquina datilográfica. Nestas ocasiões, me sentia mais lendo do que escrevendo o livro. No dia mais produtivo, em meio às minhas atividades normais como diretor do Escritório de Pesquisas de *Survey*, redigi dois capítulos, sem quase nenhuma necessidade de edição.

Já redigir outros capítulos mais parecia uma tortura. Gastava horas, garimpando dolorosamente as palavras, como um bêbado procurando pôr um pé atrás do outro numa linha

reta. Como o bêbado, caí muito. Às vezes, tropeçava em cada palavra, levando horas para criar um só parágrafo, depois jogava tudo fora e recomeçava. Não era uma visão agradável.

Um menino chega atrasado na escola e a professora diz: “Johnny, por que o atraso?”

Johnny diz: “Havia tanto gelo na rua hoje de manhã, que eu dava um passo à frente e escorregava dois pra trás.”

“Ah, é?”, diz a professora. “Então, como você chegou até aqui?”

“Desisti e tentei voltar pra casa.”

[Rufar de tambores, e o ruído da bola no aro da cesta.]

O pior foi o capítulo sobre estatística. Mais parecia um buraco negro, sugando toda a alegria da vida. Embora os pareceristas da pré-publicação elogiassem o manuscrito (“Obrigado, mamãe”, eu escrevia nas margens dos pareceres), consideraram unanimemente o capítulo sobre estatística um monstro malformado que deveria desaparecer de vez. Então, revisei. E revisei mais.

Podei o capítulo sobre estatística até restarem apenas três cálculos: *lambda*, *gama* e *qui-quadrado*. Na primeira edição do livro, dois dos três estavam corretos. Mas este não foi o fim dos problemas estatísticos.

Havia decidido que seria útil incluir uma tabela de números aleatórios como apêndice, para auxiliar os leitores na seleção de amostras. Em vez de comprar uma tabela já existente de números aleatórios, decidi usar minha recém-adquirida capacidade de programação para criar minha própria tabela. Depois de diversas tentativas fracassadas, consegui fazer chegar à Wadsworth uma tabela de números aleatórios pronta para impressão em *photo-offset*, para não correr o risco dos tipógrafos estragarem tudo.

Descobri mais tarde um problema, quando comecei a receber cartas perguntando: “Meus alunos me indagaram, e não soube responder: por que a tabela de números aleatórios não tem nenhum 9?” Numa das primeiras tentativas de criar a tabela, eu havia notado a ausência de 0s e descobrira que o sistema que eu estava usando procurava números aleatórios entre 1 e 9. Descobri como estabelecer o limite inferior em 0, mas fiz isso abaixando o limite superior em um algarismo. Em vez de escolher números aleatórios entre 1 e 9, o sistema agora escolhia entre 0 e 8. É claro que ninguém falava nada sobre a abundância de 0s, 1s, 2s etc., mas todos reclamavam do 9 ausente.

Recomendei à Wadsworth imprimir uma folha de Errata cheia de 9s, com a instrução “inserir aleatoriamente”. Em vez, compramos de alguém uma tabela de números aleatórios.

Publicado em meados de 1973, o livro foi muito bem recebido. Logo, a Wadsworth estava recebendo sugestões de instrutores para que eu escrevesse um manual metodológico mais geral sobre pesquisas, e não somente pesquisas de *survey*. Foi assim que publiquei, em 1975, *The Practice of Social Research*, que se tornou o foco de meus escritos sobre métodos de pesquisa nos anos subseqüentes.

Muitos dos instrutores que inicialmente adotaram *Survey Research Methods* mudaram para *The Practice of Social Research*, como esperávamos, mas as vendas do primeiro livro continuaram boas nos anos seguintes. Periodicamente, a Wadsworth e eu considerávamos revisar o livro, mas sempre um outro projeto acabava sendo prioritário. Finalmente, em 1988, nos comprometemos a fazer a revisão no ano seguinte, apesar do que mais pudesse acontecer.

O primeiro obstáculo foi tecnológico. Há dez anos eu redigia em microcomputadores, com todos os materiais em disquete, para facilitar a revisão, mas *Survey Research Methods* fora escrito numa máquina datilográfica antiga, sem memória, antes do advento dos sistemas de processamento de texto.

Inicialmente, a Wadsworth concordou que um digitador passasse o livro original para disquetes, para eu poder fazer a revisão num processador de texto. Acabamos tentando uma solução mais *high-tech*: escanear eletronicamente o livro. Neste processo, cada página seria eletronicamente fotografada, mas, enquanto uma fotocopiadora imprimiria a foto em papel, o *scanner* armazenava em disquete. Um programa de tradução converteria letras e palavras num texto que pudesse ser editado como qualquer arquivo de processamento de texto.

Escaneamento de textos, do tipo que acabei de descrever, certamente será uma das tecnologias mais importantes no futuro. Mas descobrimos que ainda não era uma tecnologia do presente. A editora contratou uma empresa para o trabalho. Uma vez prontos, os arquivos tinham, no dizer orgulhoso desta empresa, “85% de exatidão”. Vale dizer, um sétimo do texto saía errado! (Posteriormente, descobri que outras empresas e sistemas são mais exatos, mas mesmo assim escaneamento de textos ainda não pode ser considerado “uma ciência exata”.)

Os heróicos esforços de minha secretária, Adrienne Alexander, finalmente me permitiram ter em 1988 um conjunto de disquetes com a versão de 1973 do livro, me permitindo iniciar a revisão. Caso o leitor tenha trabalhado com a edição anterior do livro, aqui vão alguns comentários sobre as principais mudanças nesta edição.

O livro de 73 foi um produto da época, repleto de linguajar sexista. Mais especificamente, o uso de pronomes masculinos na terceira pessoa quando o gênero era indefinido. Exemplo do texto original: “Um pesquisador usa amostragem estratificada quando *ele* quer...” Eu estava consciente deste problema na edição de 73, mormente ao escrever sobre entrevistadores e perfuradores de cartão, dois grupos predominantemente femininos nas pesquisas de *survey*. Ousadamente, usei pronomes femininos em relação a estes dois grupos, chamando os diretores de pesquisas de “*ela*”, acrescentando uma longa nota de rodapé explicando o procedimento e criticando a exclusão geral de mulheres da direção de projetos.

Nos anos subseqüentes, a questão da linguagem sexista esquentou e, às vezes, eu culpava “eles” por haverem criado a convenção da terceira pessoa masculina. Curioso de saber quem eram “eles”, descobri não haver nenhum grupo formal com o poder de legislar sobre a língua inglesa (como a Academia Francesa, encarregada de manter a pureza lingüística do *français*). No caso do inglês, estas questões são mais ou menos frouxamente controladas por professores de inglês, por editores em geral e editores de dicionários em particular e, bem, por pessoas como eu. Conseqüentemente, me comprometi a eliminar linguagem sexista dos meus textos.

Desde 1973, me tem sido fácil evitar linguagem sexista. Ofereço abaixo algumas soluções e o incentivo a adotá-las. Antes, eis algumas “soluções” das quais não gosto e que tenho evitado:

- Substituir *he* (*ele*) por *she* (*ela*) nada resolve, pois simplesmente troca um problema discriminatório por outro.
- Fazer usos aleatórios ou alternativos de *ele* e *ela* acaba confundindo, sugerindo haver uma razão para se usar o feminino ou o masculino.
- Usar repetitivamente ‘*ele ou ela*’ é desajeitado, embora seu uso parcimonioso possa ser bom.
- Mutantes lingüísticos como *ele/a* (em inglês, *s/he*) são horrorosos, além de impronunciáveis. Além do mais, parecem uma concessão sarcástica ao que o escritor poderia considerar uma preocupação irrazoável: como referências sarcásticas a *personagem* em lugar de *management*, *personal* em vez de *manual* etc.

Eis o que para mim *tem* funcionado. Usar a primeira e/ou segunda pessoa, evitando a terceira, produz uma comunicação mais pessoal e evita o dilema sexista. “Se quisermos, você e eu podemos escolher uma amostragem sistemática...” é um exemplo. Quando preciso acrescentar uma terceira pessoa, uso nomes andróginos como Jan e Pat. Outra solução são os pronomes plurais, como: “pesquisadores optam por amostragem sistemática quando *they* desejam...” (*they* é tanto *eles* ou *elas*).

A linguagem sexista também pode ser evitada pelo uso coloquial, geralmente aceito no inglês oral, da terceira pessoa do plural (*them*), quando o gênero é indefinido. Por exemplo: “quando chega uma pessoa a meu gabinete, faço com que *them* sentem numa cadeira”, se bem que revisores dificilmente deixam passar despercebida esta construção.

Depois de encontrar esta e outras soluções para evitar linguajar sexista, finalmente descobri a melhor solução: parar de *pensar* em termos sexistas. Quando pensava num pesquisador, não mais pensava na figura de um homem — nem só em mulheres para entrevistadoras. Conseqüentemente, já quase não me meto mais em armadilhas sexistas que exigem acrobacias lingüísticas para delas escapar.

Além de eliminar o linguajar sexista da edição de 73, tive também de lidar com alguns exemplos desatualizados de pesquisas. Exemplo: em vez de falar do sucesso dos pesquisadores que previram a eleição de Richard M. Nixon para a presidência dos EUA em 1968, usei dados de boca de urna da eleição presidencial de 1988 para ilustrar o poder da amostragem de *survey*.

O fato de só haver comprado meu primeiro computador pessoal seis anos após a publicação da primeira edição indica a atualização necessária na área de tecnologia de pesquisa. A edição atual não mais supõe pesquisadores usando cartões perfurados nem contadores-classificadores ou computadores “queimadores de lenha”. O foco agora é no uso de microcomputadores, entrevistas assistidas por computador e coisas assim. Até diria que minha descrição da tecnologia de *survey* está provavelmente um pouco à frente das normas atuais — contudo, por pouco tempo. O que hoje parece de ponta, bem logo será de uso geral e, em pouco tempo, se desatualizará.

Além do *hardware* de informática usado em pesquisas, muitas outras coisas mudaram nos dezesseis anos desde a primeira edição. Pesquisadores aprenderam truques novos, desenvolveram novas técnicas etc. Embora seja impossível fazer justiça a toda ela num só livro, tentei abordar a maioria da nova literatura.

Ensinando e escrevendo, acho que, com o tempo, melhorei as explicações de diversos aspectos do processo de pesquisa, e esta nova edição me dá oportunidade de compartilhar o que aprendi. Particularmente, as novas tecnologias da informática me auxiliaram a mostrar graficamente muitos conceitos, que é a forma como os compreendo mentalmente. Sempre converti imagens e diagramas em palavras, mas hoje estou mais apto a apresentá-las *au naturel*.

Finalmente, o livro novo dá mais atenção explícita à pesquisa de mercado. Pelo que sei, a edição de 73 foi muito usada como referência supletiva em cursos e agências de *marketing*, já que, em minha opinião, manuais sobre pesquisas de *marketing* tendem a ser mais teóricos que práticos. Certamente pesquisa de *survey* é uma técnica amplamente utilizada por pessoas de *marketing*. Se bem que não pretendendo fazer deste livro o texto principal de um curso de pesquisa de mercado, tentei fazê-lo mais relevante e útil como suplemento.

Como na edição original do livro, gostaria de agradecer diversas pessoas, importantes para sua criação. Inicialmente rededico-o a Samuel A. Stouffer, Paul F. Lazarsfeld e Charles Y. Glock, em quem vejo minha linhagem na pesquisa de *survey*. O comentário estranho sobre “os velhos turcos” é uma referência à dedicatória de Paul Lazarsfeld para sua obra *The Language of Social Research* (1955) a “Charles Y. Glock e seus ‘jovens turcos’ no Escritório de Pesquisa Social Aplicada da Universidade de Columbia”.

Pelos comentários tão sábios quanto simpáticos dos rascunhos da edição revista, sou grato aos colegas Alan Acock, da Louisiana State University, e Tom Guterbock, da Universidade da Virgínia.

Serina Beauparlant, a nova editora de sociologia na Wadsworth, herdou o manto que Steve Rutter vestira dezesseis anos antes. Seu trabalho foi excepcional, como demonstra este produto acabado. Nesta tarefa, foi auxiliada por Donna Linden, editora de produção, e pelo *designer*-executivo Andrew Ogus.

Finalmente, gostaria de agradecer a minha esposa Sheila, que inicialmente teria sido co-autora da primeira edição. Apesar de ter resolvido não fazê-lo, continua sendo minha eterna parceira e colaboradora. Fonte permanente de inspiração, constantemente me incentivou a olhar acima e além das montanhas, do outro lado das esquinas e debaixo das pedras. Convido-o, leitor, a investigar e procurar comigo nas páginas que se seguem.

Prefácio à Primeira Edição

Pesquisa de *survey* é hoje provavelmente o método de pesquisa mais conhecido e amplamente usado nas ciências sociais. Cresce diariamente seu uso no mundo acadêmico. É lecionada e usada em departamentos de sociologia, ciência política, psicologia, administração de empresas, saúde pública e geografia, dentre outros. Cada vez mais, alunos de pós-graduação nas ciências sociais são incentivados a fazer *surveys* como requisito de pesquisa original para teses e dissertações. Estudantes de graduação freqüentemente fazem *surveys*, e professores anualmente publicam relatórios de centenas de *surveys*.

Fora do mundo acadêmico, quase todo mundo ouviu falar de pesquisas de opinião pública, de previsões eleitorais, de estudos do mercado de consumo e de censos. Praticamente todo mundo já foi afetado por pesquisas. Campanhas e sonhos políticos deslancham ou são destruídos com base em pesquisas eleitorais. Empresas produzem em massa ou interrompem produção baseadas em pesquisas de mercado. Programas sociais de ajuda federal muitas vezes dependem dos resultados de estudos de população em cidades e estados carentes. Atividades tanto profissionais como não profissionais de pesquisas são freqüentes. Clubes pesquisam seus sócios sobre políticas a seguir e datas de piqueniques. Bibliotecas e lanchonetes levantam preferências dos usuários sobre serviços e horários. Estações de rádio convidam ouvintes a telefonarem votando em temas comunitários.

O uso tão difundido e a aceitação tão ampla da pesquisa de *survey* parece indicar que se trata de técnica de fácil

aprendizado e utilização. Todos que leram um relatório de pesquisa de opinião no jornal local provavelmente acham que eles mesmos podem fazer uma. Afinal, qualquer um pode perguntar e contar as respostas.

Por Que Este Livro Enfatiza Lógica e Capacidades

Este livro aborda três questões relacionadas à percepção errônea de que é fácil fazer pesquisas de *survey*. Em primeiro lugar, a popularidade destes métodos tem levado a muita pesquisa malfeita. Às vezes, técnicas de pesquisa são mal-usadas; muitas vezes *surveys* são realizados quando algum outro método de pesquisa seria mais adequado. Espera-se que a discussão detalhada da lógica e das capacitações necessárias às pesquisas de *survey* ajude a melhorar a qualidade das mesmas.

Em segundo lugar, o mau uso e o uso excessivo levaram à rejeição total de pesquisas de *survey* por parte de muitas pessoas, inclusive um número crescente de jovens cientistas sociais. Neste aspecto, espera-se que este livro demonstre que pesquisa de *survey* pode ser um método extremamente útil de investigação científica em certas situações. Usadas corretamente em situações apropriadas, *surveys* podem gerar informações necessárias, difíceis de se obterem através de qualquer outro método.

Em terceiro lugar, afirmar que determinada pesquisa foi mal feita pressupõe a existência de um acervo comprovado e estabelecido de padrões científicos, à luz dos quais avaliar atividades de pesquisa. Lamentavelmente, isto não ocorre. Certamente há padrões científicos muito rigorosos em alguns aspectos da pesquisa de *survey*, como amostragem e manipulação estatística de dados. Mas os padrões não são tão claramente definidos em outros aspectos, tais como a redação de questões e a codificação de respostas. Padrões existentes em geral não são formalizados e, ainda por cima, tendem a ser transmitidos como tradição oral, de pesquisador para aluno-aprendiz. Contudo, pesquisadores iniciantes sem acesso a este sistema de aprendizado podem não entender nem mesmo os padrões informais. Este livro procura revelar estes padrões informais e ordená-los dentro do contexto lógico da investigação científica.

O próprio título, *Survey Research Methods*, reflete a orientação geral do texto. Como instrutor de métodos de pesquisa de *survey* e consultor de *surveys*, eu sentia a necessidade premente de um só volume que oferecesse a estudantes e a pesquisadores em potencial um guia prático de pesquisas de *survey*. O mercado

está cheio de livros e ensaios que discutem a lógica teórica da investigação científica, sem especificar como normas científicas são aplicadas na prática. No outro extremo, há um punhado de livros do tipo receita-de-cozinha no mercado, com diretrizes passo a passo sobre como pesquisar, mas sua utilidade diminui rapidamente à medida que a situação de campo do pesquisador se diferencia daquelas discutidas diretamente no texto. Quando o pesquisador encontra um problema de campo não coberto no livro-receita, ficará provavelmente perdido. Este livro procura uma solução para este problema.

Este texto é focado na *lógica* e nas *capacitações* da pesquisa de *survey*. Relativamente pouca atenção é dada à estatística, já que há vários textos excelentes sobre ela. Ao cobrir os diversos aspectos do desenho e da análise de *survey*, usei o seguinte formato. Primeiro, discute-se a lógica teórica, por exemplo, da amostragem. É necessário que o leitor entenda a lógica básica de se selecionar uma amostra de entrevistados cujas respostas podem representar a população da qual foram selecionados. Então, discutem-se, passo a passo, os métodos mais típicos de amostragem. A esta altura, o leitor deve entender não apenas *como* selecionar uma amostra convencional, mas também *por que* métodos convencionais de amostragem advêm da lógica teórica da amostragem. É apresentado, em seguida, um leque de problemas práticos freqüentemente encontrados por pesquisadores de campo e discutem-se métodos típicos de se lidar com tais problemas, dentro do contexto da lógica da amostragem do *survey*. Aqui, o leitor deve entender por que as soluções sugeridas oferecem a melhor correspondência com a lógica e os métodos ideais da amostragem. Não há como abordar todos os problemas práticos que os leitores possam porventura vir a encontrar, mas se espera que eles compreendam, na prática, a lógica da amostragem, de forma a poderem chegar às suas próprias boas soluções. Mesmo quando a situação da pesquisa não permite uma *boa* solução, o leitor deve ser capaz de avaliar o significado da decisão que tomar, a respeito das conclusões a serem tiradas do seu estudo.

Nenhum *survey* satisfaz plenamente os ideais teóricos da investigação científica. Cada um representa um conjunto de compromissos entre o ideal e o possível. A meta primordial deste texto é ajudar os leitores a chegarem aos melhores compromissos possíveis. *Surveys* perfeitos podem não ser possíveis, mas bons *surveys* podem e devem ser realizados.

Este livro pode parecer um tanto elementar, focalizando quase exclusivamente a lógica e as capacitações *básicas*. Por exemplo, relativamente pouca atenção é dada às técnicas de

escala, enquanto muita atenção é devotada à construção de índices simples. Discussões de análise de *survey* são primordialmente focalizadas em tabelas de percentagens. Há duas razões para esta orientação. A literatura metodológica existente cobre extensamente técnicas avançadas, enquanto as técnicas mais básicas têm sido abordadas menos extensamente. Mais importante, uma base firme na lógica e nas capacidades para a pesquisa de *survey* é pré-requisito para o pleno entendimento de técnicas mais avançadas. Assim, se o leitor conseguir compreender plenamente a lógica de análise através de tabelas de percentagens, sinto que estará mais equipado para lidar com correlações, regressões, análise fatorial e assim por diante. Lamentavelmente, hoje em dia, muitos alunos saltam diretamente para formas mais complicadas de análise, sem entenderem bem a lógica das pesquisas em geral.

Como Este Livro Pode Ser Usado

Este livro foi feito para uma ampla faixa de leitores: qualquer pessoa interessada em realizar um *survey* ou em usar e/ou avaliar os resultados de um *survey*. Mais especificamente, destina-se a três audiências distintas: estudantes de metodologia, pesquisadores principiantes e consumidores de pesquisas.

Primeiro, o livro se dirige a alunos de graduação cursando pela primeira vez métodos de pesquisas. Enquanto diversos professores preferem enfatizar pesquisa de *survey* no curso geral, outras metodologias podem ser acrescentadas. Por isso o livro é publicado em capa mole, para permitir um uso flexível dos materiais. A organização do material surgiu de minhas próprias experiências em sala de aula, lecionando metodologias de *survey* e um esboço preliminar do texto foi testado neste ambiente. Além dos materiais no corpo da obra, a maioria dos capítulos vem acompanhada de uma lista de leituras adicionais, que podem ser usadas para um exame mais detalhado de tópicos específicos.

Em segundo lugar, o livro pretende ser um guia prático e realista para o pesquisador principiante, sem muita experiência em pesquisas de *survey*. O livro aborda os prováveis problemas, decisões e compromissos que constituem a substância quotidiana da pesquisa de campo. Procurei oferecer o leque mais amplo possível de exemplos tirados de minhas próprias pesquisas e consultorias, numa ampla variedade de projetos e condições de pesquisa, bem como das experiências de colegas e outros pesquisadores. Espero que o número de exemplos aumente a

probabilidade de um pesquisador encontrar orientação no texto. Simultaneamente, busquei apresentar tais exemplos dentro da lógica básica da investigação científica. Em vez de pedir que o leitor memorize técnicas específicas — mesmo um leque variado de técnicas —, peço que entenda por que determinadas técnicas são recomendadas e usadas. Assim, quando o pesquisador precisar aceitar um compromisso ou aproximação, deve estar apto a saber o que está comprometendo e que conseqüências este provavelmente acarretará.

Finalmente, espero que os consumidores de pesquisa de *survey* encontrem neste livro o solo metodológico que os ajude a fazer avaliações críticas e esclarecidas. Espero ter apresentado um quadro suficientemente completo das realidades práticas da pesquisa de campo, para melhorar a perspectiva dos críticos do *survey*. Discutindo honestamente os méritos e usos possíveis da pesquisa de *survey*, espero converter alguns críticos em apoiadores esclarecidos. Também discuto as desvantagens, insuficiências e maus usos das pesquisas de *survey*, com a esperança de ajudar alguns usuários ardentes demais a se tornarem um pouco mais críticos.

Agradecimentos

Tive muita sorte ao ter a oportunidade de aprender pesquisa de *survey* num ambiente de aprendizado, tutelado por Charles Glock. Numa época em que professores são acusados de explorar estudantes, Charlie consistentemente me tratava como um colega mais jovem, me incentivando a embarcar em projetos próprios e me tornar mais independente, podendo recorrer a ele para orientação mais avançada. Um reconhecimento rotineiro num prefácio de livro é muito pouco para a dívida que tenho com Charlie Glock. Apesar de depois ter aprendido coisas sobre pesquisa de *survey* que ele não me ensinou, ainda me sinto um constrangido ao afirmar que este livro é meu.

Ao mesmo tempo, devo muito a outros professores e colegas, principalmente a Rod Stark, que agia como irmão mais velho nos meus primeiros anos na Universidade de Berkeley. Mais tarde, em Berkeley, tive a sorte de conhecer Gertrude Selznick, que me obrigou a examinar ainda mais profundamente os alicerces lógicos da análise de *survey*, e nossas comparações de anotações sobre pesquisa empírica me deram um entendimento bem melhor do que é de fato investigação científica. Outros mestres e colegas que aprimoraram

minha compreensão da pesquisa de *survey* enquanto ciência incluem George von Bekesy, Lin Freeman, Dave Gold, Ren Likert, Bill Nicholls, Jay Palmore, Steve Steinberg e Charlie Yarbrough. Sou especialmente grato pelos comentários detalhados e profundos de três revisores do manuscrito: Andy Anderson, Joseph Spaeth e Billy J. Franklin.

Igualmente reconheço minha dívida para com os muitos pesquisadores com quem trabalhei em consultoria. Alguns me ensinaram novas técnicas, outros revisaram e comentaram os primeiros esboços deste texto. Dentre eles cito Jim Dannemiller, Dave Ford, Dennis Hall, Dave Johnson, Jan LeDoux, Heung-soo Park, Dinny Quinn, Françoise Rutherford, Yongsock Shin, Dave Takeuchi, Chuck Wall e Choon Yang.

Além do mais, o livro jamais teria se concretizado sem a assistência de diversas pessoas envolvidas com o manuscrito. Na editora Wadsworth, Jack Arnold, Steve Rutter e Rod Stark organizaram e providenciaram importantes insumos editoriais. Barbara Higa assumiu um extenso projeto bibliográfico. Pat Horton me ajudou a encontrar tempo para redigir, além de ter datilografado o longo manuscrito.

Dado seu conhecimento especial de coleta de dados e experiência geral com pesquisas de *survey*, minha esposa Sheila teria sido co-autora do livro, tendo papel importante na organização inicial da obra. Pouco depois do início da redação, tornou-se impossível para Sheila dedicar tempo suficiente ao livro, mas ela continuou ativamente participando como consultora, confidente e crítica.

O Contexto Científico da Pesquisa de *Survey*

CAPÍTULO 1

A LÓGICA DA CIÊNCIA

CAPÍTULO 2

A CIÊNCIA E AS CIÊNCIAS SOCIAIS

CAPÍTULO 3

PESQUISA DE *SURVEY* COMO MÉTODO DAS CIÊNCIAS SOCIAIS

Este livro aborda o tema da pesquisa científica. Sua finalidade é ajudá-lo a fazer suas próprias pesquisas e compreender as pesquisas feitas por outros. Embora o livro focalize um determinado método de pesquisa — o *survey* —, deve ser lido dentro do contexto geral da ciência. A pesquisa de *survey* emprega técnicas específicas peculiares a este método, que podem se encaixar nas normas gerais da investigação científica.

Apesar de termos ciência há séculos e inúmeros livros terem sido escritos sobre pesquisa científica, o assunto ainda hoje é enigmático para não-cientistas e cientistas iniciantes. A Parte 1 deste livro pretende oferecer uma visão geral da lógica da ciência, para que, posteriormente, as técnicas específicas da pesquisa de *survey* possam fazer mais sentido. Três capítulos são dedicados a este tópico.

O Capítulo 1 considera a lógica básica da ciência. Começa discutindo a imagem tradicional da ciência, imagem que creio ser em geral enganosa e inapropriada no contexto da atual pesquisa científica. Neste aspecto, meu propósito não é desmistificar, mas tornar a ciência relevante e realista.

Na prática, a ciência não é nem mística nem pura. Como todas as atividades humanas, a pesquisa científica é um compromisso entre o ideal e o possível. Na prática, é freqüentemente guiada pela emoção, pelo erro e pela não-racionalidade. Todavia, a pesquisa científica é bem diferente de outras atividades humanas. Tentarei elucidar tais diferenças no Capítulo 1.

O Capítulo 2 aborda o tópico ainda mais complicado da *ciência social*, sobre o qual muito se escreveu nos últimos anos. Alguns autores insistem que a ciência social é tão “científica” quanto a ciência natural. Outros não aceitam tal noção. Ainda outros pensam que a pesquisa social está progredindo rumo ao *status* de ciência, mas ainda não chegou lá.

Sugiro que este debate não faz sentido. Assim como é impossível uma afirmação definitiva sobre o que “ciência” realmente é, é igualmente impossível afirmar que a ciência social é ou não científica. O que faremos é considerar as diferenças e semelhanças entre ciências sociais e naturais, explorando as implicações destas diferenças e semelhanças para a lógica e as técnicas da pesquisa científica social.

O Capítulo 3 é dedicado a um exame específico da pesquisa de *survey*. O capítulo começa com um apanhado histórico dos *surveys* e passa a abordar modos como a pesquisa de *survey* se encaixa nas normas gerais de ciência e de ciência social.

O propósito principal destes capítulos é prepará-lo para o entendimento das bases lógicas subjacentes às competências e técnicas específicas da pesquisa de *survey*. Em termos mais gerais, espero que você adquira um embasamento melhor na lógica da ciência *per se*. Isto é particularmente importante porque toda pesquisa científica é um compromisso entre o ideal e o possível. A maior parte deste livro lida com estes compromissos.

Inicialmente examinamos o que deve ser feito em circunstâncias ideais, consideramos em seguida quais compromissos são mais apropriados quando o ideal não pode ser atingido. A menos que você compreenda plenamente a lógica básica da ciência, não conseguirá entender por que um compromisso é aceito no lugar de um outro. Compreendendo esta lógica, ao começar sua própria pesquisa você será capaz de conseguir o melhor compromisso possível dentro das reais condições de campo.

Capítulo 1

A Lógica da Ciência

Ciência é uma palavra caseira. Todos a usam, pelo menos ocasionalmente, mas imagens da “ciência” diferem muito. Para alguns, é sinônimo de matemática, para outros significa aventais brancos e laboratórios, outros a confundem com tecnologia, e há aqueles para quem a palavra equivale a matérias difíceis em cursos colegiais ou universitários.

Evidentemente, ciência não é nenhuma destas coisas *per se*. Porém é difícil especificar exatamente o que é ciência. Este capítulo começa resumindo a imagem de ciência frequentemente apresentada a estudantes em cursos de introdução às ciências. Em seguida, examinamos a imagem negativa que algumas pessoas têm das ciências. Finalmente, descrevemos os componentes lógicos da ciência tal como é *praticada*.

A Perspectiva Tradicional

Cursos de introdução às disciplinas científicas frequentemente apresentam uma imagem de Ciência que a faz parecer direta, precisa e até rotineira. Posteriormente, neste capítulo, criticarei esta perspectiva, mas ela é útil como ponto de partida.

Teoria Científica

Cientistas, é o que se diz, começam com um interesse por algum aspecto do mundo ao redor. O interesse pode ser saber como a pressão arterial é regulada no corpo, por que

um grão de arroz é mais duro que outro, o que determina o caminho dos cometas ou o que causa câncer. Então, eles examinam rigorosamente esse tópico e estrutura em termos lógicos e abstratos.

Eles identificam todos os fenômenos relevantes ao assunto estudado. Com base no conhecimento existente, inter-relacionam estes fenômenos numa rede de *relações causais*¹ — quais elementos causam ou influenciam outros elementos. Deste modo, desenvolvem uma *teoria*, um conjunto de *proposições* lógicas, inter-relacionadas, que explicam a natureza do fenômeno estudado.

Para testar a validade da teoria, supostamente derivam *hipóteses*, que são predições sobre o que acontecerá em condições especificadas. Muitas vezes, essas hipóteses são da forma se — então. Se o Evento A acontecer, segue-se então o Evento B. Por exemplo, se você estudar bem neste curso, terá uma vida muito feliz. (É uma hipótese, lembre-se, não necessariamente o que as coisas são.) Já que esta relação causal é sustentada pela teoria geral, se o Evento B não se seguir ao Evento A, a validade da própria teoria seria questionada.

Operacionalização

Teorias por natureza são abstratas e gerais. Hipóteses, embora mais específicas, são também em geral um tanto abstratas. (No exemplo acima, não especifiquei que tipo de estudo ou que tipo de felicidade.) Assim, hipóteses devem ser convertidas em termos operacionais, isto é, nas operações concretas feitas para testá-las — processo chamado de *operacionalização*. Você deve especificar quais fenômenos no mundo real constituem o Evento A e quais constituem o Evento B.

Se a hipótese é sobre o efeito da temperatura na taxa de crescimento de uma planta, você especifica como temperatura e crescimento serão medidos, ou seja, que operações correspondem a esses conceitos abstratos. Especificadas as operações, você descreve um *experimento* para testar a hipótese, especificando a duração do experimento, a frequência de medição do crescimento e da temperatura, se e como a temperatura será artificialmente controlada, ou se variações naturais serão anotadas, e como a medição será registrada e analisada e assim por diante.

Baseado na sua hipótese, você pode prever taxas de crescimento correspondentes a diferentes temperaturas e especificar

o quanto as predições devem ajustar-se às taxas observadas de crescimento para confirmar ou rejeitar a hipótese. Deixar claro tudo isso de antemão pode evitar muita discussão mais tarde.

Experimento

Finalmente, as operações especificadas são efetuadas. Dados são coletados e manipulados como prescritos e a hipótese é testada. Se o experimento confirmar a hipótese, valida-se a teoria geral da qual ela foi derivada. Se a hipótese não for confirmada, a teoria geral é questionada. Qualquer que seja o resultado, presume-se que você publicará seus achados, o mundo se tornará um lugar um pouco melhor para se viver e você começará a pensar noutros tópicos para conquistar.

Imaginação e brilho parecem necessários para construir uma teoria, mas não para a coleta e análise de dados. Especificados de antemão os métodos experimentais e as operações, assistentes técnicos supostamente podem realizar e interpretar o experimento.²

Resumo

Já que cientistas operam com procedimentos *racionais e objetivos*, suas conclusões têm supostamente mais qualidade do que as impressões subjetivas e preconceitos do leigo. Cientistas lidam com fatos e números e diz-se que os números não mentem.

A Desmistificação da Ciência

Em anos recentes, dá-se a alguns estudantes uma imagem da ciência e dos cientistas bem diferente da perspectiva tradicional apresentada. Esta visão mais negativa da ciência tem várias dimensões.

Primeiro, salienta-se que cientistas são motivados pelas mesmas emoções humanas e limitados pelas mesmas fraquezas humanas de todas as pessoas. Diz-se que cientistas freqüentemente selecionam seus objetos de estudo com base em tendências pessoais, e que alguns podem dedicar todas as suas energias a “provar” algum palpite de estimacão. Em vez de planejar e executar experimentos objetivamente, a pesquisa deles é uma busca contínua de dados para corroborar um preconceito.

Cultos e panelinhas do mundo científico têm sido mostrados ao público. Afirma-se que um trabalho científico apresentado para publicação pode ser julgado mais com base no *pedigree* acadêmico do pesquisador (diplomas, escola etc.) do que nos méritos intrínsecos do próprio trabalho. Um editor de periódico que estudou com o Professor X pode rejeitar todos os trabalhos apresentados por alunos do Professor Y.

Além do mais, freqüentemente se enfatiza que “idéias aceitas” nas disciplinas científicas muito dificilmente podem ser questionadas. Um trabalho de pesquisa apresentando uma perspectiva radicalmente nova sobre um assunto antigo e supostamente já decidido pode nunca ser publicado.

Também diz-se que muitos pesquisadores avaliam a perspectiva de um projeto de pesquisa mais pela probabilidade de receber financiamentos do que por sua possível contribuição ao conhecimento.

Já que tanta pesquisa científica é realizada nas universidades, é relevante anotar as críticas à norma “publique ou pereça”, atribuída a muitos departamentos acadêmicos. A acusação é que professores que se destacam no magistério podem ser despedidos se não realizarem e publicarem pesquisas suficientemente. Pesquisas efetuadas sob tais condições de pressão provavelmente não manifestam a curiosidade intelectual e a busca da verdade que associamos à imagem idealizada da ciência.

Além do mais, pesquisar por coerção às vezes produz fraudes científicas. Ocasionalmente, você lê na imprensa reportagens sobre algum pesquisador que falsificou dados para parecer um estudioso produtivo. Além de errado, este tipo de trapaça às vezes desorienta pesquisas de outros, desperdiçando recursos, atrasando descobertas científicas, além de outras perturbações causadas ao processo de pesquisa.

As críticas à ciência foram alimentadas por diversas narrativas francas de pesquisas, publicadas nos últimos anos por cientistas de renome.³ Cada vez mais cientistas procuram apresentar honestamente seus projetos de pesquisa, pôr seus achados numa perspectiva adequada e oferecer melhor orientação a aspirantes a pesquisadores. Como estas narrativas revelam erros, omissões e outros problemas práticos, muitos críticos contemporâneos da ciência as tomam como confissões de que esta não passa de mistificação.

Alguma desmistificação da ciência é até boa de vez em quando. É fácil pensar na ciência como um empreendimento místico e os cientistas como praticantes-mágicos infalíveis. Se a ciência é fundamentalmente uma atividade racional e objetiva, deve ser capaz de resistir a uma avaliação racional e objetiva.

Os aspectos que não sobreviverem à crítica provavelmente não deverão continuar fazendo parte do empreendimento científico.

O grande perigo na crítica a pesquisas malfeitas e até desonestas é oferecer uma fuga fácil à dificuldade de entender a ciência como esta deve ser. Estudantes às vezes acham mais simples ignorar a ciência, como se ela fosse uma bobagem ritualista, do que aprender estatística ou a lógica da pesquisa científica. É mais fácil considerar toda a ciência ruim do que se tornar um bom cientista.

Evidentemente, neste tópico sou tendencioso. Acredito que a ciência é uma atividade humana significativa. Enquanto muitas atividades rotuladas de “científicas” de fato não o são, em minha opinião muitas atividades científicas diferem de formas importantes de outras atividades humanas. É importante entender estas diferenças, tanto para quem pesquisa como para quem lê sobre elas e tem suas vidas afetadas por elas.

O problema primordial tem a ver com a inexactidão da perspectiva tradicional do método científico, tal como é convencionalmente apresentada a alunos principiantes de ciência, tanto nas ciências sociais como nas naturais. Na prática, a ciência não corresponde exatamente à sua imagem tradicional, mas, ao mesmo tempo, não é tão ruim como argumentam seus críticos mais severos. A seção seguinte procura descrever a ciência na prática, que é distinta de sua perspectiva tradicional. Em seguida, abordarei os aspectos que distinguem a ciência das outras atividades humanas.

A Ciência na Prática

A perspectiva tradicional sugere que os cientistas passam diretamente de uma curiosidade intelectual sobre alguns fenômenos para a derivação de uma teoria. Raramente, talvez nunca, isto acontece. O interesse inicial num fenômeno muitas vezes se origina em alguma pesquisa empírica anterior, talvez em alguns achados inconsistentes gerados pela sua própria pesquisa ou no trabalho de outros. De certo modo, você pode começar com a “resposta” e partir para a descoberta da “questão”.

Você pode iniciar com uma observação específica: crianças de lares sem pai e/ou mãe têm taxa de delinqüência mais alta do que as outras. Em seguida, você tenta desenvolver um entendimento mais geral de por que isto acontece.

Quase nunca teorias resultam de processos totalmente dedutivos. Mais comumente, teorias são o resultado final de uma longa cadeia de *dedução* e *indução*. Em certo ponto,

você pode ter uma explicação preliminar de uma relação empírica, pode testar esta explicação preliminar através da coleta de mais dados, usar os novos resultados para modificar a explicação, coletar novos dados e assim por diante. Construção de teoria, então, envolve interação de observações e explicações.

Portanto, raramente as teorias são confirmadas num determinado momento. Na ciência, são relativamente poucos os “experimentos críticos”, ou seja, experimentos que levam toda a teoria a se sustentar ou desabar. Ao contrário, evidências montadas pouco a pouco apóiam uma teoria continuamente modificada. Em algum momento, alguma forma da teoria pode se tornar geralmente aceita, mas raramente é possível identificar o momento exato em que a teoria foi “provada” ou aceita. Além do mais, toda teoria continua a passar por modificações. Nenhum cientista já descobriu ou descobrirá “A Verdade”.

A operacionalização de *conceitos* nunca é tão clara e direta como sugere a imagem tradicional da ciência. Este assunto será abordado mais profundamente no Capítulo 7, mas deve-se observar que a maioria dos conceitos científicos são passíveis de diversas interpretações. Assim, você pode especificar operacionalizações preliminares desses conceitos e usar os resultados dos experimentos tanto para avaliar tais operacionalizações quanto para testar hipóteses. Se melhorias num novo bem de consumo aparentemente não aumentarem a satisfação dos consumidores, você pode questionar se está medindo satisfação adequadamente.

Mesmo quando conceitos são operacionalizados de forma aceitável, resultados experimentais raramente são conclusivos em sentido absoluto, mesmo com relação a hipóteses específicas. Tipicamente, se confirma ou se rejeita uma hipótese *até certo ponto*, quase nunca completamente. Já comentamos que teorias são normalmente aceitas com base no peso da evidência de diversos experimentos. Se um grande número de observações empíricas é mais bem explicado pela Teoria A do que por qualquer outra teoria disponível, então a Teoria A provavelmente será aceita.⁴

Finalmente, é totalmente incorreta a impressão de que testes empíricos são uma atividade rotineira. A imagem tradicional da ciência sugere que criatividade e até mesmo brilho são necessários à derivação de teorias e ao desenho de experimentos, mas que a execução propriamente dita dos experimentos é monótona e sem imaginação. Na prática, a execução de um experimento, a coleta de dados empíricos, exige inúmeras decisões críticas. Surgem situações inesperadas. Observações bizarras são registradas, sugerindo erros de medida. Dados

podem ser perdidos ou falsificados. (Como se verá neste livro, esses problemas são particularmente comuns em *surveys*.) Além disto, a operacionalização de conceitos nunca é totalmente não-ambígua e deve ser especificada mais ainda durante o experimento. Cada uma destas situações requer decisões que influenciarão o resultado do experimento e, por extensão, a avaliação da hipótese e a teoria da qual ela pode ter sido derivada.

Isto é importante. Se você trabalhou como assistente de pesquisa num projeto, ou conhece outros que o fizeram, provavelmente conhece a disparidade entre o desenho teórico do projeto e o trabalho quotidiano nele. Especialmente se foi um projeto mal supervisionado, você pode ter uma visão um tanto negativa da sua descrição, aparentemente tão científica numa publicação acadêmica. Neste caso, você pode ficar tentado a concluir que toda ciência é “não-científica”.

Enfatizo este ponto por dois motivos. Primeiro, saliento (e farei isso no livro todo) que a qualidade de um projeto de pesquisa depende em grande parte das decisões e atividades aparentemente mundanas que se dão na coleta e no processamento de dados. O diretor de projeto que não se envolve intimamente com tais atividades corre o sério risco de dirigir um projeto sem sentido. Mas, ao mesmo tempo, as implicações destas decisões — mesmo o desleixo — nem sempre aparecem prontamente. O que ao leigo pode parecer um procedimento desleixado pode de fato ser insignificante para o valor final dos dados. Tomar boas decisões e avaliar as implicações das más depende de uma compreensão sólida da lógica da ciência enquanto atividade distinta. É este último aspecto que abordaremos agora.

O que É Ciência?

Basicamente, toda ciência pretende entender o mundo ao redor. Três componentes principais constituem esta atividade: descrição, a descoberta de regularidades e a formulação de teorias e leis. Primeiro, cientistas *observam e descrevem* objetos e eventos que aparecem no mundo. Isto pode envolver a medição da velocidade de um objeto que cai, o comprimento de onda de emissões de uma estrela distante ou a massa de uma partícula subatômica. Tais descrições são guiadas pelos objetivos da exatidão e da utilidade.

Segundo, cientistas procuram *descobrir regularidades e ordem* no caos por vezes alucinante e estonteante da experiência. Em parte, isto pode envolver a coincidência ou *correlação* de

certas características ou eventos. Assim, por exemplo, você pode notar que pressão atmosférica se correlaciona com altitude ou que a aplicação de força a um objeto resulta na modificação de sua velocidade.

Terceiro, cientistas tentam *formalizar e generalizar as regularidades* descobertas em teorias e leis. São exemplos disto a lei da gravidade de Newton e as teorias gerais e especiais da relatividade de Einstein. Teorias e leis são, em geral, enunciados lógicos de relações entre características e eventos que oferecem explicações para uma ampla faixa de ocorrências empíricas.

Vale a pena observar que *não*-cientistas buscam estas mesmas três metas. Todos nós observamos e descrevemos o mundo ao redor. Procuramos achar regularidades: o funcionário de escritório pode descobrir que chegar atrasado ao trabalho resulta em repreensão. Tentamos formular leis e teorias que ofereçam orientação geral na vida diária, tais como perspectivas religiosas que sustentam que aderir a ensinamentos religiosos resultará em recompensas neste e/ou no outro mundo.

Para enfatizar este ponto, vejamos as semelhanças nas respectivas atividades de um cientista descuidado e de um racista. Ambos fazem observações sobre o mundo e podem relatá-las a outros. Por exemplo, o cientista observa que membros de uma determinada tribo primitiva gozam de saúde dental relativamente boa; o racista observa que um lojista chinês trapaceou numa transação comercial. O cientista conclui preliminarmente que a dieta da tribo pode ser responsável pela saúde dental dos seus membros, enquanto o racista conclui que a raça do lojista é responsável pelas práticas comerciais antiéticas.

Tanto o cientista como o racista procuram novas observações para reforçar suas conclusões preliminares. O cientista checa a saúde dental de outras tribos primitivas com dietas semelhantes, enquanto o racista fica de olho em lojistas chineses. É importante notar que tanto o cientista como o racista são *seletivos* em suas observações subseqüentes. Concentrando-se na dieta, o cientista pode ignorar o ambiente climático da tribo, a estrutura econômica e assim por diante. Por sua vez, o racista não dará atenção à educação dos lojistas, sua classe social etc.

Além do mais, tanto o cientista descuidado como o racista podem tender a ignorar observações que contradizem as conclusões a que chegaram. O cientista pode ignorar relatórios de boa saúde dental em tribos com dietas radicalmente diferentes e relatórios sobre tribos que seguem a dieta em questão,

mas cujos dentes apodrecem e caem. O racista pode ignorar todos os lojistas chineses honestos e todos os não-chineses desonestos.

O cientista descuidado pode justificar o fato de ignorar tais relatórios atribuindo-os a trabalho de campo malfeito. (Evidentemente, o cientista cuidadoso lidaria com tais casos.) O racista pode referir-se vagamente às “exceções que provam a regra”.⁵

Ao se depararem com um excesso de observações que desconfirmam suas hipóteses, tanto o cientista descuidado como o racista procurarão informações adicionais que possam alinhar as observações perturbadoras às suas conclusões. No caso da tribo com dentes bons mas dieta diferente, o cientista pode intensificar o trabalho de campo na tentativa de descobrir que a dieta desta tribo é mais semelhante àquela em questão do que parecia inicialmente. Confrontado a um lojista não-chinês desonesto, o racista poderia começar a buscar um avô chinês do bisavô deste lojista ou paixão por comida chinesa.

Apesar de anomalias empíricas, nosso cientista pode acabar com uma teoria relacionando uma certa dieta a saúde dental, e o racista, uma teoria de que chineses são desonestos. As duas teorias serão aparentemente corroboradas por observações empíricas e explicações lógicas.

A comparação acima visa salientar dois pontos. Primeiro, não há diferença mágica entre atividades científicas e não científicas. Vimos duas linhas de investigação bem semelhantes. Segundo, as atividades de “cientistas” variam em qualidade “científica”. Faz mais sentido falar de atividades mais ou menos científicas do que dicotomizar entre atividades científicas e não científicas. Assim, uma linha de investigação conduzida por um físico profissional pode ser relativamente não científica, enquanto um bombeiro hidráulico pode fazer pesquisas do mais alto calibre científico.

No restante desta seção, daremos atenção às características que tornam uma determinada atividade mais ou menos científica. Procuraremos compreender as características *ideais* da ciência, entendendo que nenhuma atividade está de pleno acordo com estes ideais, seja conduzida por um cientista profissional ou por um leigo.

Ciência É Lógica

Ciência é fundamentalmente uma atividade racional, e explicações científicas devem fazer sentido. Religiões podem se fundar em revelações, costumes, tradições, apostar na fé. A ciência, porém, deve se fundamentar na razão lógica.

Lógica é um ramo difícil e complexo da filosofia, e um delineamento completo dos sistemas da lógica excede em muito o escopo deste livro. Talvez alguns exemplos ilustrem o que significa a ciência ser lógica. Por exemplo, um determinado evento não pode logicamente causar um outro que ocorreu antes dele. Suas atitudes sociais com relação a, digamos, relações raciais não podem determinar a região do país onde você nasceu, mas o inverso pode ser verdade.

Assim, a ciência adota uma abordagem diferente das visões *teleológicas* de algumas religiões. Por exemplo, alguns cristãos acreditam que Jesus estava destinado a ser crucificado e que, portanto, este destino fez com que ele fosse traído e julgado. Tal ponto de vista não pode ser aceito dentro da lógica da ciência.

Na lógica da ciência é impossível um objeto ter duas qualidades *mutuamente excludentes*. Ao jogar uma moedinha, não pode dar simultaneamente cara e coroa. Em contraste, podemos notar que muitas pessoas preconceituosas dizem que os armênios são “clânicos” (recusam misturar-se com outras nacionalidades) e “caras-de-pau” (impõem sua presença aos não-armênios). Face a essas assertivas conflitantes, a lógica da ciência sugere que ou uma ou outra, ou ambas, destas caracterizações dos armênios não é verdadeira ou que as duas características estão definidas de tal forma que não são mutuamente excludentes.

Um evento também não pode ter resultados mutuamente excludentes. Assim, educação superior não pode fazer uma pessoa mais rica e mais pobre ao mesmo tempo. Educação universitária pode fazer uma pessoa mais rica e outra pessoa mais pobre, assim como alguns armênios podem ser descritos como clânicos e outros como caras-de-pau, mas resultados ou descrições contraditórias não têm sustentação lógica e são intoleráveis para a ciência.

Nada disto diz que, na prática, a ciência está totalmente isenta de enunciados ilógicos. Você deve saber que os físicos atualmente consideram a luz como partículas e como ondas, apesar de estas descrições da natureza da luz serem contraditórias. Esta contradição particular existe na ciência porque a luz se comporta como partículas sob certas condições e como ondas sob outras. Conseqüentemente, físicos continuam a usar as duas conceituações contraditórias segundo parecem apropriadas em condições dadas. Todavia, isto representa uma tensão para a lógica da ciência.

Indo um pouco além da noção de “senso comum” da lógica, dois sistemas lógicos distintos, mencionados antes, são

importantes para a busca científica: *lógica dedutiva* e *lógica indutiva*. Beveridge as descreve assim:

Os lógicos distinguem raciocínio indutivo (de instâncias particulares para princípios gerais, de fatos para teorias) e raciocínio dedutivo (do geral para o particular, aplicando teoria a um caso particular). Na indução, inicia-se com dados observados e se desenvolve uma generalização que explica a relação entre os objetos observados. Por outro lado, no raciocínio dedutivo se começa com alguma lei geral, que é aplicada a uma instância particular.⁶

O exemplo clássico da lógica dedutiva é o *silogismo* familiar: “Todos os homens são mortais; Sócrates é homem; portanto, Sócrates é mortal.” Um pesquisador poderia dar seqüência a este exercício dedutivo testando empiricamente a mortalidade de Sócrates. Esta é essencialmente a abordagem discutida antes como a “perspectiva tradicional da ciência”.

Usando a lógica indutiva, você pode começar notando que Sócrates é mortal e observar também diversos outros homens. Você nota que todos os homens observados eram mortais e, a partir daí, arrisca a conclusão de que todos os homens são mortais.

A lógica dedutiva é um sistema muito antigo, remontando no mínimo a Aristóteles, que predominou na filosofia ocidental até os séculos XVI ou XVII. O nascimento da ciência moderna foi marcado pelo surgimento da lógica indutiva em vários contextos científicos. Cada vez mais conclusões gerais derivadas de observações cuidadosas contradiziam os postulados gerais que ancoravam muitos sistemas dedutivos.

Na astronomia, por exemplo, as observações de Ptolomeu levaram-no a variações do seu modelo, desenvolvendo um modelo epicíclico no qual as estrelas e planetas rotavam em círculos ao redor de pontos no espaço que, por sua vez, rotavam em círculos ao redor da Terra estacionária. À medida que se observaram variações adicionais, o sistema se tornou cada vez mais complicado, com níveis novos de epiciclos sendo acrescentados para manter as crenças-chaves no movimento circular e na Terra estacionária.

Copérnico atacou o sistema de Ptolomeu sugerindo que o Sol, e não a Terra, era o centro do universo. Ele derivou essa perspectiva radicalmente nova da observação do movimento celestial, e não de um compromisso inicial de que a Terra era o centro do universo. Mas Copérnico não desafiou o suposto do movimento circular. “Um astrônomo posterior,

Kepler, disse que Copérnico não conseguiu vislumbrar as riquezas ao seu alcance, contentando-se em interpretar Ptolomeu em vez de interpretar a natureza.⁷ Kepler, por outro lado, estava determinado a interpretar a natureza sob a forma de volumosos dados empíricos que herdara do astrônomo dinamarquês Tycho Brahe. Continua Butterfield:

Sabemos com que colossal gasto de energia ele testou hipótese após hipótese, descartando-as até alcançar o ponto em que já tinha um vago conhecimento da forma exigida — aí decidindo que, para fins de cálculo, uma eclipse poderia lhe dar resultados aproximados e descobrindo que de fato uma eclipse dava certo.⁸

Este exemplo ilustra o surgimento da lógica indutiva na ciência. Dramas semelhantes ocorreram em outros campos de investigação durante os férteis séculos XVI e XVII. Mais ou menos um século mais tarde, a pesquisa indutiva e científica de Charles Darwin entrou em conflito com outra tradição.

Não se deve concluir destes exemplos históricos que a lógica dedutiva é inerentemente incorreta ou está ultrapassada. Um exercício de lógica dedutiva é tão bom quanto sua consistência interna e a verdade de seus supostos iniciais. Por outro lado, a lógica indutiva é tão boa quanto sua consistência interna e a exatidão de suas observações.

Na prática, a pesquisa científica envolve tanto o raciocínio indutivo quanto o dedutivo, na medida em que os cientistas vão e vêm incessantemente entre teoria e observações empíricas.

A Ciência É Determinística

A ciência se baseia no suposto de que todos os eventos têm causas antecedentes sujeitas à identificação e ao entendimento lógico. Para o cientista nada “simplesmente acontece” — acontece por uma razão. Se alguém gripa, se chove hoje, se uma bola parece rolar morro acima, o cientista supõe que cada um destes eventos é susceptível de explicação racional.

Como veremos no capítulo seguinte, esta característica da ciência traz uma dificuldade especial para as ciências sociais, que competem com noções de senso comum sobre comportamento social. Você pode afirmar que fez alguma coisa, por exemplo, votou num candidato, simplesmente porque decidiu agir assim, mas o cientista social provavelmente argumentaria que seu voto foi determinado por vários eventos e condições prévias. A decisão de votar pode ser atribuída à

sua classe social, à região do país onde você mora etc., mesmo que você porventura negue a influência destes fatores.

Mas deve-se tomar diversos cuidados neste aspecto. Primeiro, cientistas não conhecem, nem pretendem conhecer, as causas específicas de todos os eventos. Simplesmente supõem que tais causas existem e podem ser descobertas. Segundo, a ciência aceita causação múltipla. Qualquer evento pode ter várias causas; sua decisão de votar pode ter sido causada por diversos fatores. Um evento, também, pode ter uma causa, enquanto outro evento semelhante pode ter uma causa diferente. Duas pessoas podem votar no mesmo candidato por razões diferentes, mas se supõe que em cada caso existe uma razão.

Finalmente, grande parte da ciência se fundamenta numa forma *probabilística* de determinismo. Assim, o Evento A pode resultar no Evento B 90% das vezes ou 70% de todos os Republicanos podem votar num determinado político, enquanto somente 23% dos Democratas o fazem. Neste sentido, diz-se que filiação político-partidária *determina* comportamento eleitoral, ainda que de forma incompleta. Outros fatores poderiam ser introduzidos para explicar as discrepâncias.

A Ciência É Geral

A ciência busca entendimento geral mais do que explicação de eventos individuais. Tipicamente, o cientista se interessa menos por que determinada bola cai para baixo, quando solta do alto, e mais por que *todas* as bolas tendem a se comportar assim. Do mesmo modo, o cientista se interessa menos em explicar por que você votou da forma que o fez e mais em entender por que os eleitores em geral votaram da forma como o fizeram.

Esta característica da ciência se relaciona com seu determinismo probabilístico. É concebível que possamos explicar *completamente* as razões de um evento — por exemplo, por que certa pessoa votou no candidato X. É concebível que consigamos descobrir cada fator individual que influenciou a decisão do voto. Caso o façamos, podemos supostamente prever com perfeita exatidão o comportamento eleitoral de pessoas *idênticas*. Mas, a longo prazo, esta capacidade não geraria muita informação útil sobre o comportamento eleitoral em geral. Primeiro, é duvidoso que encontrássemos outra pessoa com exatamente as mesmas características da primeira. Mais importante ainda, nossas descobertas poderiam ser de pouquíssima ajuda na compreensão do padrão de voto de pessoas com outras características. Ficamos, então, satisfeitos

com menos de 100% de entendimento, se pudermos entender o comportamento eleitoral em geral.

Este é o sentido em que o cientista e o historiador diferem em suas abordagens do mesmo tema. O historiador procura entender tudo sobre um determinado evento específico, enquanto o cientista se interessa mais no entendimento geral de uma classe de eventos semelhantes, mas não idênticos. Assim também o psicólogo e o terapeuta diferem na abordagem do comportamento humano. O psicólogo examina o comportamento esquizofrênico de vários indivíduos, procurando chegar a uma compreensão geral da esquizofrenia, enquanto o terapeuta aproveita o conhecimento geral já existente para procurar ajudar um indivíduo específico.

Portanto, a capacidade de generalização é uma característica importante das descobertas científicas. Descobrir que bolas vermelhas caem na Terra a uma certa aceleração é menos útil do que descobrir que bolas de todas as cores fazem isso. Também é menos útil saber que bolas caem com uma determinada aceleração ao nível do mar do que saber que a aceleração de todas as bolas em queda pode ser determinada por sua altitude.

A Ciência É Parcimoniosa

Portanto, os cientistas procuram descobrir fatores determinantes de tipos de eventos. Ao mesmo tempo, procuram descobrir os fatores *não* determinantes de eventos. Assim, ao determinarmos a aceleração de um objeto em queda, descartamos sua cor como irrelevante.

Em termos mais gerais, os cientistas tentam descobrir as razões dos eventos usando o mínimo possível de fatores explicativos. Na prática, é claro, o número de fatores explicativos considerados aumenta caracteristicamente o grau de determinação conseguida. Um cientista político pode conseguir um certo grau de explicação do comportamento eleitoral usando apenas dois fatores, por exemplo, filiação partidária e classe social. Outro cientista político poderia alcançar um entendimento mais completo levando em consideração outros fatores, como raça, região onde a pessoa foi criada, sexo, educação etc. Frequentemente, cientistas são forçados a escolher entre simplicidade de um lado e grau de explicação de outro. Em última análise, tentam otimizar o equilíbrio entre explicação e simplicidade, a fim de conseguir o máximo de explicação com um número mínimo de fatores. Esta *parcimônia* é uma qualidade bem ilustrada pela elegância da famosa equação de Einstein: $e = mc^2$.

Já observamos que a ciência é geral, no sentido de procurar chegar a descobertas e leis de aplicabilidade geral. Entretanto, a maioria dos conceitos gerais são sujeitos à diversidade de interpretações. Por exemplo, ao procurar explicar as fontes do preconceito em geral, você percebe que o preconceito assume muitas formas diferentes. Portanto, ao desenhar, realizar e relatar sua pesquisa, você precisa ser preciso em seus métodos de medir o conceito.

Ao realizar um projeto de pesquisa sobre preconceito, você tem de operacionalizar especificamente o conceito de preconceito, por exemplo, concordância com várias afirmações num questionário que parecem indicar preconceito. No relatório da pesquisa, você deve ter cuidado ao descrever suas operacionalizações detalhadamente, para que o leitor saiba exatamente como o conceito foi medido. Mesmo que alguns leitores possam discordar de sua operacionalização, pelo menos saberão qual ela é.

Freqüentemente, a generalizabilidade de uma descoberta é alcançada pelo uso de diversas operacionalizações diferentes dos conceitos envolvidos. Se determinado conjunto de fatores causa preconceito, independentemente de como o preconceito é medido, você pode concluir que estes fatores resultam em preconceito em geral.

A Ciência É Empiricamente Verificável

No ápice de sua elegância, a ciência resulta na formulação de leis ou equações gerais, descrevendo o mundo ao redor. Mas tais formulações só são úteis se puderem ser verificadas pela coleta e manipulação de dados empíricos. Uma teoria geral do preconceito seria inútil se não sugerisse modos através dos quais dados pudessem ser coletados e não previsse os resultados que seriam obtidos na análise dos dados.

Há, porém, outra forma de ver esta característica. De certo modo, nenhuma teoria científica pode ser provada. Consideremos o exemplo da gravidade. Os físicos dizem que um corpo cai para baixo por causa da atração geral entre os corpos físicos e que esta relação é afetada pela massa dos corpos envolvidos. Já que a massa terrestre é vasta, uma bola lançada de uma janela se movimenta em direção à Terra.

Tal explicação da gravidade é empiricamente verificável. Um pesquisador pode jogar uma bola pela janela e observá-la caindo. Mas isto não *prova* a verdade da teoria da gravidade.

O que o pesquisador faz é especificar que se a bola *não* cair, a teoria da gravidade é incorreta. Já que se observa a bola se comportando como esperado, a teoria da gravidade *não foi desconfirmada*.

Assim sendo, quando afirmamos que uma explicação científica deve sujeitar-se ao teste empírico, queremos dizer, mais exatamente, que o pesquisador deve ser capaz de especificar as condições nas quais a teoria seria desprovada. Neste sentido é que cientistas falam da *desconfirmabilidade* das teorias. Se você falha consistentemente em desprovar sua teoria, fica crescentemente confiante na correção dela. Mas é importante entender que você nunca vai conseguir *prová-la*.

Continuando com o exemplo anterior, outro teórico poderia notar que a bola experimental era da mesma cor do chão no qual caía, sugerindo que corpos da mesma cor são mutuamente atraídos por qualquer razão que ele pudesse pensar. O experimento inicial daria confirmação às duas teorias em competição. A “teoria da atração pela cor” sugere, porém, que se uma bola de cor *diferente* do chão for atirada pela janela, não deveria cair. Um experimento apropriado resultaria (esperamos) na desconfirmação empírica da teoria.

A Ciência é Intersubjetiva

Com frequência se afirma que a ciência é “objetiva”, mas tal afirmação tipicamente resulta em muita confusão quanto ao que seja “objetividade”. Além do mais, nota-se crescentemente nos últimos anos que nenhum cientista é completamente objetivo em seu trabalho. Todos os cientistas são “subjetivos” até certo ponto — influenciados por suas motivações pessoais. Ao afirmar que a ciência é *intersubjetiva*, queremos dizer que dois cientistas com orientações subjetivas diferentes chegariam à mesma conclusão se cada um deles conduzisse o mesmo experimento. Um exemplo de ciência política deve esclarecer este conceito.

A tendência dos intelectuais nos EUA se alinharem mais com o Partido Democrata do que com o Republicano leva muitas pessoas a supor que os Democratas, como grupo, são mais escolarizados do que os Republicanos. É razoável supor que esta afirmação satisfaria um cientista Democrata e aborreceria um Republicano. Mesmo assim, seria possível que os dois cientistas concordassem sobre o desenho de um projeto de pesquisa para coletar dados no eleitorado americano, referente a filiação partidária e nível educacional. Os dois cientistas poderiam então realizar estudos independentes

desse assunto, e ambos descobririam que os Republicanos, como um todo, têm um nível educacional mais alto do que os Democratas. (Devido ao fato de que o Partido Democrata também atrai uma proporção maior dos eleitores da classe trabalhadora, enquanto os homens de negócios são mais atraídos pelo Partido Republicano.) Os dois cientistas políticos — com orientações subjetivas opostas — chegariam à mesma conclusão empírica.

Cientistas muitas vezes discordam entre si. Podem oferecer explicações claramente diferentes de um evento. Mas, em geral, tais desacordos envolvem questões de conceituação e definição. Assim, um cientista social pode relatar que religiosidade se relaciona positivamente com preconceito, enquanto outro discorda. O que discorda provavelmente sugerirá que uma ou ambas as variáveis foram incorretamente medidas. Você poderia realizar seu próprio estudo, medindo diferentemente as duas variáveis, e relatar uma relação negativa entre elas. Mas, se os primeiros pesquisadores houvessem relatado precisa e detalhadamente o desenho e a execução dos seus estudos e você conseguisse replicá-los exatamente, chegaria ao mesmo achado. É isto que significa a intersubjetividade da ciência.

A Ciência É Aberta a Modificações

A seção anterior deve ter deixado claro que a “ciência” não oferece uma seqüência de etapas fáceis para atingir “A Verdade”. Dois cientistas, ambos aderindo às características previamente discutidas de ciência, podem chegar a explicações bem diferentes de um fenômeno. Além disso, pode não haver, num dado momento, como avaliar os méritos relativos delas. Se duas explicações se contradizem, as duas não podem, presumivelmente, estar corretas. Ou se demonstra que uma das duas ou ambas estão incorretas ou se descobre que as duas explicações, não são, afinal de contas, mutuamente excluídas, devido a uma mudança de paradigma, por exemplo.

Inúmeras teorias “científicas” do passado foram mais tarde desprovadas e substituídas por novas teorias. Tudo que “sabemos” hoje era antes previamente “conhecido” diferentemente e, às vezes, consideramos ingênuas, tolas ou estúpidas estas antigas visões. Vale a pena lembrar, porém, que tudo que “conhecemos” hoje provavelmente será mudado no futuro e o povo do futuro — nossos arrogantes descendentes — nos considerarão ingênuos, tolos ou estúpidos. (Se isto o perturba, talvez o console lembrar-se de que eles também sofrerão a mesma sorte.)

A ciência não busca a verdade definitiva, mas a *utilidade*. Teorias científicas não devem ser julgadas por sua verdade relativa, mas pela medida de sua utilidade em melhorar nosso conhecimento do mundo ao redor.

Em última análise, as características da ciência discutidas neste capítulo oferecem um conjunto de diretrizes que aumentam a utilidade de descobertas e teorias. Investigações que procuram se pautar por tais características produzirão, a longo prazo, mais descobertas úteis do que investigações de outros tipos. Assim, uma pessoa pode ser capaz de prever o tempo com mais exatidão baseada no seu joelho reumático do que todos os meteorologistas científicos do mundo, mas, a longo prazo, os cientistas contribuirão mais para o nosso conhecimento geral da natureza do clima.

Resumo

Este capítulo começou revendo a imagem tradicional da ciência, primariamente como um conjunto de etapas que inevitavelmente conduzem à "Verdade". Esta visão foi contrastada com outra, mais recente, crítica da ciência, que sugere que cientistas não são tão diferentes dos leigos. Na parte principal do capítulo, tentamos mostrar que, embora a investigação científica não seja infalível, ela *difere* em aspectos importantes de outras atividades humanas. Cientistas estão certamente sujeitos a todas as fraquezas humanas das demais pessoas, mas a ciência provê um conjunto de diretrizes que pode aumentar a utilidade das suas investigações.

Este capítulo abordou a ciência em geral e o próximo focalizará especificamente a *ciência social*. Veremos que a ciência social está presa às mesmas regras que os outros tipos de investigação científica. Ao mesmo tempo, porém, a temática especial das ciências sociais apresenta problemas especiais — e oportunidades especiais.

Notas

¹ Palavras e frases em "itálico" são definidas e discutidas no Glossário no fim do livro. Se você não entender bem um termo, vale a pena verificar imediatamente no Glossário.

² Em minha experiência, isto é mais típico da psicologia do que da sociologia, como se vê, por exemplo, em propostas de teses de doutorado. Um estudante de pós-graduação em psicologia cuja proposta de PhD foi aceita já se encontra quase no final da tese, enquanto o de sociologia está apenas começando.

³ Ver, por exemplo, WATSON, James D. *The Double Helix*. New York: The New American Library Inc., 1968, e o acervo de biografias de pesquisas sociais em HAMMOND, Phillip (Ed.). *Sociologists at Work*. New York: Basic Books, 1964, e GOLDEN, M. Patricia (Ed.). *The Research Experience*. Itasca, IL: F. E. Peacock, 1976.

⁴ Não suponha que isto ocorre automaticamente. Muito tem sido escrito sobre a tenacidade com que os cientistas às vezes se apegam a "paradigmas" estabelecidos (modelos ou pontos de vista), mesmo em face de evidência em contrário e paradigmas adversários. O livro clássico neste tópico é de KUHN, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1970.

⁵ Originalmente, *provar*, nesta expressão, significava "testar", como quando se diz que grandes obstáculos "testaram severamente" o compromisso de alguém. Vale dizer, exceções eram vistas como ameaças ou desafios à regra.

⁶ BEVERIDGE, W. I. B. *The Art of Scientific Investigation*. New York: Vintage Books, 1950. p.113.

⁷ BUTTERFIELD, Herbert. *The Origins of Modern Science*. New York: The MacMillan Company, 1960. p.24.

⁸ Ibidem. p.64.

Leituras Adicionais

BEVERIDGE, W. I. B. *The Art of Scientific Investigation*. New York: Vintage Books, 1950.

BUTTERFIELD, Herbert. *The Origins of Modern Science*. New York: The MacMillan Company, 1960.

GOLDEN, M. Patricia (Ed.). *The Research Experience*. Itasca, IL: F. E. Peacock, 1976.

IRVINE, William. *Apes, Angels and Victorians*. New York: Meridian Books, 1959.

KAPLAN, Abraham. *The Conduct of Inquiry*. San Francisco: Chandler Publishing Co., 1964.

KUHN, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1970.

TOBEN, Bob. *Space-Time and Beyond*. New York: E. P. Dutton, 1975.

WATSON, James D. *The Double Helix*. New York: The New American Library, Inc., 1968.

WHITEHEAD, Alfred North. *Science and the Modern World*. New York: The MacMillan Company, 1925.

Capítulo 2

A Ciência e as Ciências Sociais

Um dos mais vivos debates acadêmicos dos últimos anos diz respeito ao *status* “científico” das disciplinas englobadas sob a rubrica de ciências sociais — tipicamente incluindo sociologia, ciência política, psicologia social, economia, antropologia, pesquisa de mercado e, às vezes, áreas como geografia, história, comunicações e outros campos compostos ou especializados. A questão básica é se o comportamento humano pode ser submetido ao estudo “científico”. Já que o capítulo anterior chamou a atenção para a confusão sobre o termo *ciência* em geral, não é de surpreender que os acadêmicos discordem também sobre as ciências sociais.

Oposição à idéia de ciências sociais tem surgido tanto dentro como fora dos campos das ciências sociais. Dentro dos campos, o movimento para as ciências sociais tem significado um redirecionamento e, em alguns casos, uma renomeação da tradição acadêmica estabelecida. Cada vez mais, departamentos de Governo têm sido substituídos por departamentos de Ciência Política, departamentos de Fala se tornam de Comunicação. Em muitos casos, o movimento para as ciências sociais significa uma passagem da ênfase na descrição para a explicação sistemática. Na ciência política, isto significa maior ênfase em explicar comportamento político do que em descrever instituições políticas. Em antropologia, tem representado diminuição na ênfase em etnografia. O crescimento de subcampos como a econometria tem tido efeito semelhante na economia, como ocorre com a historiografia na história. Alguns geógrafos vêm mudando da enumeração de importações e exportações para modelos matemáticos de migração. Compreensivelmente, profissionais treinados e com experiência nos métodos mais tradicionais destes campos objetam às novas orientações.

Fora dos departamentos de ciência social, oposição semelhante vem das ciências físicas — de físicos, biólogos, químicos etc. Guiados pela imagem tradicional da ciência discutida acima, alguns cientistas físicos objetam que o “método científico” não pode ser aplicado ao comportamento social humano.

Freqüentemente, os defensores das ciências sociais abasteceram o debate emulando cegamente os protocolos e rituais das ciências estabelecidas. Essa emulação toma muitas formas: fascínio por equipamentos de laboratório, uso muitas vezes inapropriado da estatística e da matemática, o desenvolvimento de uma terminologia obscura e a adoção em bloco de teorias e terminologias das ciências físicas.

Na maioria, estes erros parecem advir da aceitação da imagem tradicional da ciência e da falta de compreensão da lógica da ciência na prática. Aspirantes a cientistas sociais muitas vezes buscam obter entendimento com métodos que não funcionam mesmo para os cientistas físicos. O resultado, freqüentemente, é serem ridicularizados pelos cientistas físicos, por colegas de profissão e por leigos.

É um firme suposto deste livro que o comportamento social humano pode ser submetido a um estudo “científico” tão legitimamente quanto átomos e células. Este é, porém, um suposto que deve ser entendido no contexto da discussão anterior da ciência na prática. Desta ótica, nenhuma diferença significativa parece existir entre ciências físicas e sociais. Como os cientistas físicos, os cientistas sociais procuram descobrir regularidade e ordem. Cientistas sociais buscam regularidade no comportamento social através da observação e medição cuidadosas, descoberta de relações e elaboração de modelos e teorias.

A Busca de Regularidades Sociais

Medindo Fenômenos Sociais

O primeiro tijolo da ciência é a “medição” ou a observação sistemática. Não há razão fundamental pela qual os cientistas sociais não possam medir fenômenos relevantes às suas investigações. Por exemplo, idade e sexo dos atores sociais podem ser medidos. Local de nascimento e estado civil podem ser medidos de diversas maneiras, variando em exatidão e economia.

Comportamento social agregado também pode ser medido sistematicamente. O cientista político consegue determinar o comportamento na urna em dia de eleição de todo o eleitorado

ou de zonas eleitorais individuais. A quantidade de tráfego num determinado trecho de rodovia pode ser medida em diferentes pontos do tempo. Pesquisadores de mercado podem medir volumes de vendas.

Atitudes também podem ser medidas, embora este seja um ponto de muito desacordo. Por exemplo, preconceito contra mulheres pode ser medido determinando a aceitação ou a rejeição individual de crenças e perspectivas que representam tal preconceito. Religiosidade, liberalismo político e conservadorismo, autoritarismo e variáveis semelhantes também podem ser medidos de maneira semelhante.

Medir atitudes é freqüentemente questionado como “não-científico”; embora aborde este tema mais adiante neste e em capítulos subseqüentes, devo fazer alguns comentários aqui. Deve-se reconhecer que todas estas medidas (todas as medidas, aliás) são basicamente arbitrárias. O cientista social não pode descrever uma pessoa inequivocamente como “alienada” e outra como “não alienada”. Pessoas serão, ao invés, descritas como *relativamente mais ou menos alienadas* — ou seja, comparando uma com a outra. Esta característica, entretanto, não é prerrogativa das ciências sociais, como demonstram a “escala de dureza” usada nas ciências físicas, a “escala Richter” para terremotos etc. Ninguém pode dizer em termos absolutos que um determinado metal é “duro” ou um terremoto “severo”, apenas que é “mais duro” ou “mais severo” que outro.

Descobrimo Regularidades Sociais

As pessoas tendem a considerar os objetos das ciências físicas como mais regulares que os das ciências sociais. Um objeto pesado cai para baixo sempre que é solto, enquanto uma pessoa pode votar num determinado candidato numa eleição e contra o mesmo candidato na eleição seguinte. Assim também, o gelo se derrete quando aquecido, enquanto pessoas aparentemente religiosas nem sempre vão à igreja. Embora estes exemplos sejam geralmente verdadeiros, há o perigo de se acabar descartando totalmente a existência de regularidades sociais. A existência de normas sociais observáveis contradiz esta conclusão.

Algumas normas sociais são prescritas pelas instituições formais de uma sociedade. Por exemplo, só pessoas de uma certa idade ou mais podem votar em eleições. Antes de 1920, homens americanos podiam votar e mulheres não. Tais prescrições formais, então, regulam ou regularizam o comportamento social.

Além de prescrições formais, podemos observar outras normas sociais. Republicanos registrados votarão mais provavelmente

em candidatos Republicanos do que Democratas. Professores universitários tendem a ganhar mais dinheiro do que trabalhadores não especializados. Mulheres tendem a ser mais religiosas do que homens.

Relatos de regularidades por cientistas sociais estão frequentemente sujeitos a três tipos de crítica. Primeiro, o relato pode ser acusado de trivialidade, de que todo mundo já sabia disso. Segundo, podem ser citados casos contraditórios, indicando que a observação não é totalmente verdadeira. Terceiro, pode-se argumentar que as pessoas envolvidas podem, se realmente quiserem, perturbar a regularidade observada.

A acusação de que muitas descobertas feitas por cientistas sociais são triviais ou já bem conhecidas levou muitos aspirantes a cientistas a procurarem descobertas esotéricas ou obscuras que provassem que a ciência social é mais do que senso comum pretensioso. Esta resposta é inapropriada sob diversos pontos de vista. Para começar, tantas contradições são evidentes no vasto corpo do “senso comum”, que se torna essencial expurgar sistematicamente os erros nele existentes. Mesmo quando uma proposição não é questionada pelos leigos, ela deve ser testada empiricamente.

Muitos instrutores de metodologia das ciências sociais iniciam suas aulas revelando um conjunto de “descobertas importantes” das ciências sociais, derivadas de estudos conduzidos por Samuel A. Stouffer durante a II Guerra Mundial.¹ Tais “descobertas” incluem os seguintes achados:

- Soldados negros ficavam mais felizes em campos de treinamento no Norte do que no Sul dos Estados Unidos.
- Soldados da Força Aérea americana, onde as promoções eram rápidas, tinham mais probabilidade de achar que seu sistema de promoções era justo do que os soldados da Polícia Militar, onde elas eram muito lentas.
- Soldados mais escolarizados tinham mais probabilidade de se ressentir com o alistamento militar forçado do que soldados menos escolarizados.

Quando os alunos começam a fazer pouco caso da obviedade das “descobertas importantes”, o instrutor revela que *cada uma delas foi desprovada* pela pesquisa de Stouffer e explica por que a relação observada faz sentido quando examinada abaixo

do nível das aparências.² Em suma, “documentar o óbvio” é uma função valiosa de qualquer ciência, física ou social, e não é uma crítica legítima de qualquer empreendimento científico. (Darwin cunhou o termo “experimento de tolo” numa referência irônica a grande parte de sua própria pesquisa.)

A crítica de que certas generalizações da ciência social estão sujeitas à desconfirmação em casos específicos também não é um desafio suficiente ao caráter “científico” das investigações. Assim, não basta notar que um certo homem é mais religioso do que uma certa mulher. Regularidades sociais representam padrões probabilísticos, e uma relação geral entre duas variáveis não precisa ser verdadeira em 100% dos casos observáveis.

A ciência física não está isenta desta crítica. Na genética, por exemplo, o cruzamento de uma pessoa de olhos azuis com outra de olhos castanhos *provavelmente* resultará numa criança de olhos castanhos. Mas se a criança nascer de olhos azuis, isto não desafia a regularidade observada, já que a genética afirma apenas que a probabilidade dos olhos serem castanhos é maior e, mais ainda, que olhos castanhos devem ser esperados num certo percentual dos casos. O cientista social faz uma previsão probabilística semelhante — que, no cômputo geral, mulheres têm mais probabilidade de serem mais religiosas do que homens. Com instrumentos de medição testados adequadamente, podemos prever o percentual de mulheres que se mostrarão mais religiosas do que homens.

Finalmente, a acusação de que regularidades sociais observadas podem ser perturbadas pela vontade consciente dos atores não é desafio suficiente à ciência social, mesmo se coisa semelhante parece não existir nas ciências físicas. (Supostamente, um objeto não consegue resistir e não cair no chão “porque não quer”.) Isto não é negar que um racista branco de direita pode, se quiser perturbar os cientistas políticos estudando eleições, votar num negro esquerdista radical. Numa eleição, todos os votantes podem, de repente, mudar e votar no candidato azarão, frustrando as pesquisas eleitorais. Seguindo o mesmo raciocínio, todos os trabalhadores podem ir trabalhar mais cedo ou ficar em casa e não ir trabalhar, evitando que aconteça o congestionamento de tráfego esperado na hora do “rush”. Mas tais coisas não ocorrem tão frequentemente a ponto de ameaçar seriamente a observação de regularidades sociais. Ironicamente, é claro, se todos os trabalhadores ficarem em casa, isto também seria uma regularidade susceptível de explicação.

O fato é que normas sociais existem e podemos observá-las. Quando mudam com o tempo, podemos também observar e explicar as mudanças. Em última análise, regularidades sociais persistem

porque tendem a fazer sentido para os indivíduos nelas envolvidos. Podemos sugerir que é lógico esperar que um certo tipo de pessoa aja de certa maneira e esta pessoa pode muito bem concordar com a base lógica desta expectativa. Assim, mesmo se direitistas brancos podem votar num negro radical de esquerda, eles acham estúpido votar assim, assim como eleitores negros de esquerda acham estúpido votar num candidato da Ku Klux Klan.

Criação de Teorias Sociais

Cientistas sociais ainda não criaram teorias de comportamento social comparáveis às teorias desenvolvidas pelos cientistas físicos. Inúmeras teorias de comportamento social foram elaboradas há já vários séculos, mas sua adequação não é mais defendida seriamente. Evidentemente, muitas teorias sobre o mundo físico foram também descartadas. O abandono da teoria dos epiciclos de Ptolomeu não nega o caráter científico da astronomia contemporânea. Nem saber que as teorias contemporâneas da física serão mais tarde superadas nega o *status* científico desse campo.

De qualquer maneira, atualmente as ciências sociais não têm teorias formais comparáveis às existentes em outros campos. Em parte, isto se deve ao fato de que métodos sistemáticos, “científicos”, não têm sido aplicados ao comportamento social tanto tempo quanto aos fenômenos físicos. Ao mesmo tempo, a relutância em admitir a susceptibilidade do comportamento social ao estudo científico tem limitado os recursos disponíveis para o desenvolvimento das ciências sociais.

Além do mais, este livro tem origem na convicção de que o desenvolvimento científico das ciências sociais foi seriamente prejudicado pela incompreensão da natureza lógica da ciência em geral, especificamente pelo compromisso com a imagem tradicional, exclusivamente dedutiva, da ciência, oposta à compreensão da ciência na prática. Em vista disto, discutiremos agora as características das ciências sociais, paralelamente à discussão do Capítulo 1, que abordou a ciência em geral.

As Características das Ciências Sociais

A Ciência Social É Lógica

As ciências sociais visam entender racionalmente o comportamento social. Isto não significa dizer que todo comportamento social é racional. Certos comportamentos sociais são irracionais,

alguns são não racionais, mas os cientistas sociais devem ser relativamente racionais ao procurar compreender todas as formas de comportamento.

O cientista social está preso a muitas das mesmas restrições lógicas que o cientista físico. Um evento não pode causar outro evento ocorrido antes. Um objeto não pode ter características mutuamente excludentes, e um evento ou característica não pode ter resultados mutuamente excludentes. Também, tanto a lógica dedutiva quanto a indutiva são apropriadas para as ciências sociais, como já se discutiu a respeito da ciência em geral no Capítulo 1.

A Ciência Social É Determinística

Como os cientistas físicos, também os cientistas sociais supõem que eventos têm razões, que as coisas “não acontecem simplesmente”. Todo evento ou situação tem determinantes antecedentes. Esta característica das ciências sociais muitas vezes parece entrar em choque com o senso comum, como vimos antes. O cientista social pode concluir que um grupo de pessoas se comporta de determinada maneira devido a alguns eventos e condições anteriores, como no caso do comportamento eleitoral do racista branco de direita. Neste sentido, as condições de raça, preconceito e orientação política *determinam* o comportamento eleitoral. Isto não significa dizer que estes eleitores não poderiam votar num candidato negro de esquerda; apenas é improvável que o façam.

A postura determinista das ciências sociais representa o afastamento mais significativo dos estudos mais tradicionais, humanistas, do comportamento social. Enquanto o observador mais humanisticamente orientado considera a busca da alma e as agonias através das quais uma pessoa pesa os méritos e deméritos relativos de uma ação, para chegar a uma decisão bem pensada, o cientista social mais tipicamente busca os determinantes gerais de tal decisão entre diferentes agregados de pessoas. Enquanto o humanista argumenta que a decisão de cada indivíduo resulta de um processo idiossincrático, o cientista social diz que ela se encaixa num padrão bem mais simples e geral.

A Ciência Social É Geral

Vimos que as ciências sociais visam à observação e ao entendimento de padrões gerais de eventos e correlações. A utilidade de uma teoria social ou correlação social cresce em função de sua “generalizabilidade”. Quanto mais fenômenos são explicados,

maior a utilidade. Assim, uma teoria do comportamento do consumidor aplicável apenas a jovens é menos útil do que outra aplicável a consumidores de todas as idades. Uma teoria do comportamento político aplicável apenas a americanos é menos útil que outra aplicável aos povos de todas as nações.

Embora muitas vezes possamos começar tentando explicar uma faixa mais limitada de comportamento social ou o comportamento de um subconjunto limitado da população, nossa meta, normalmente, é expandir o poder explicativo de nossas descobertas a outras formas de comportamento e a outros subconjuntos da população.

A Ciência Social É Parcimoniosa

Como o cientista físico, o cientista social busca atingir o maior poder explicativo com o menor número de variáveis. Em muitos casos, a consideração adicional de novas variáveis acrescenta poder explicativo e preditivo, mas resulta também num modelo mais complicado. Na prática, o acréscimo de mais variáveis muitas vezes reduz a generalizabilidade da explicação, já que certas variáveis podem ter um efeito sobre membros de um subconjunto da população e efeito diferente sobre membros de outros subconjuntos.

Note-se que a parcimônia das ciências sociais, à semelhança de sua postura determinista, as abre à crítica das pessoas que sustentam uma visão mais humanista. Enquanto o humanista tende a explorar as profundidades dos fatores idiossincráticos que resultam numa decisão ou ação de um ator social, o cientista social conscientemente tenta limitar esta investigação.

O pesquisador de mercado procura explicar preferências por morar na cidade, nos subúrbios ou no campo baseado em, digamos, três ou quatro variáveis. Evidentemente, pode-se argumentar que todo mundo tem muitas outras razões individuais para a opção por seu estilo de vida e que um número limitado de variáveis não consegue explicar adequadamente a profundidade de uma decisão tomada por qualquer uma das pessoas estudadas.

O humanista tem razão ao afirmá-lo, mas o cientista social tem uma meta bem diferente da do humanista. O cientista social procura conscientemente obter o máximo de compreensão com o menor número de variáveis. Nem o cientista social nem o humanista está mais correto do que o outro; simplesmente, as metas dos dois são diferentes. Temos que entender plenamente a meta do cientista social, para reconhecer a não validade desta crítica.

A Ciência Social É Específica

Cientistas sociais, como os físicos, devem especificar seus métodos de medição. Isto é particularmente importante nas ciências sociais, que lidam com conceitos vagamente definidos no discurso comum. Enquanto o físico define “aceleração” mais rigorosamente que um leigo, a definição científica não difere tanto assim do entendimento comum do termo. Já conceitos como “alienação”, “liberalismo” e “preconceito” possuem significados tão variados na linguagem comum que suas definições rigorosas não são prontamente aparentes.

Cientistas sociais podem submeter tais conceitos a exame científico rigoroso, mas para isto devem especificar claramente a natureza das medições efetuadas em cada caso. Também as definições serão avaliadas com base na utilidade — sua contribuição para a explicação e o entendimento generalizados — e não na “Verdade” absoluta.

A Ciência Social É Empiricamente Verificável

Para serem úteis, proposições e teorias científicas sociais devem ser testadas no mundo real. Assim, afirmar que educação se associa negativamente com preconceito é inútil sem sugerir modos para medir as duas variáveis e testar a proposição. Como nas ciências físicas, o cientista social deve ser capaz de descrever condições empíricas nas quais uma dada proposição seria julgada incorreta, isto é, os modos pelos quais ela seria desprovada.

Crenças religiosas, como a da existência de Deus, por exemplo, não são susceptíveis de verificação empírica. Do mesmo modo, a afirmação de que membros de um grupo religioso ou racial são desleais “em seus corações”, mesmo quando parecem agir de forma leal, não pode ser submetida à verificação empírica. O mesmo se aplica a proposições sobre como seria o comportamento humano caso o Sol não aparecesse numa determinada manhã.

A Ciência Social É Intersubjetiva

Se um determinado exame científico social for descrito adequadamente em seus detalhes específicos, qualquer outro cientista social, de quaisquer convicções pessoais, deve poder replicar o exame com resultados semelhantes. Descobertas científico-sociais opostas são mais freqüentemente conseqüências de

desacordos sobre o desenho mais apropriado da pesquisa — incluindo definição e medição de conceitos — do que de resultados obtidos com um mesmo desenho.

Isto é válido tanto para pesquisas sobre tópicos altamente emocionais como religião, política e preconceito quanto para o estudo da aceleração de objetos em queda. Evidentemente, na prática, pesquisadores raramente realizam estudos quando acreditam que seu desenho é incorreto. Por exemplo, o cientista social conservador provavelmente não definirá conservadorismo de forma negativa para os conservadores.

A Ciência Social É Aberta a Modificações

Provavelmente nenhuma teoria social sobreviverá indefinidamente. Ou um crescente peso de evidências contrárias a derrubará ou se encontrará uma substituta, mais nova e parcimoniosa. Em qualquer caso, não se pode esperar que um achado de ciência social, a longo prazo, resista ao teste do tempo.

Claro que, na prática, o cientista social lida com fenômenos também afetados por ideologias — religiosas, políticas, filosóficas —, e ideologias são menos abertas a modificações que a ciência. Quando cientistas sociais explicam religiosidade em termos de variáveis estruturais, eles desafiam crenças religiosas básicas sobre comportamento moral, sistemas religiosos de recompensas e castigos etc. Quando cientistas políticos concluem que a classe operária nos EUA é mais autoritária do que a classe média, desafiam a ideologia política da esquerda. A postura determinista das ciências sociais em geral não aceita a imagem filosófica do livre arbítrio dos seres humanos, de tão antiga história na civilização ocidental.

O perigo é que alguns cientistas sociais possam estar tão pessoalmente comprometidos com certas posições ideológicas, que este compromisso os impeça de manter aberta sua prática científica. Assim, o cientista político comprometido com a esquerda pode não estar disposto a considerar, realizar ou aceitar pesquisas que possam levar à conclusão de que a classe operária é mais autoritária do que a classe média.

Já foi comentado que esta situação não é privilégio das ciências sociais. Investigações nas ciências físicas desafiaram e continuam desafiando sistemas estabelecidos de crenças ideológicas e alguns cientistas físicos têm sido prejudicados por compromissos ideológicos que reduziram a abertura de suas atividades científicas.

Métodos de Pesquisas Científico-Sociais

Apesar deste livro abordar especificamente um só método de pesquisa social, é útil colocar este método no contexto dos outros métodos disponíveis ao cientista social. Faço isto, em parte, para sugerir que o exame de um determinado fenômeno social freqüentemente é mais bem-sucedido usando-se vários métodos diferentes — ponto particularmente importante a enfatizar numa época em que a pesquisa de *survey* goza de tão grande popularidade. A pesquisa de *survey* tem vantagens especiais, mas veremos neste livro que também tem limitações e não é o método apropriado para estudar certos tópicos. Pesquisadores sociais que se restringem a um só método, *survey* ou qualquer outro, limitam gravemente sua capacidade de entender o mundo ao seu redor.

Ao mesmo tempo, é importante compreender que todos os métodos de pesquisa social são norteados pelas características gerais da ciência delineadas neste e no capítulo anterior. É útil, portanto, examinar as forças e fraquezas relativas de cada método.

O Experimento Controlado

Em muitos aspectos, o experimento controlado representa o exemplo mais claro de pesquisa científica, pelo menos na imagem popular da ciência. O desenho experimental tem muitas variações, mas aqui nos limitaremos ao desenho antes/depois com um único grupo de controle.

Suponha que você esteja interessado em métodos para reduzir o preconceito racial. Suponha também que você acredita na hipótese de que o preconceito antinegro poderia ser reduzido com uma consciência maior do papel importante dos negros na História americana. Para testar esta hipótese, você pode alugar ou mesmo produzir um filme documentando a história dos negros nos EUA. Este filme representaria o *estímulo* para o experimento.

Em seguida, você seleciona dois grupos de sujeitos. Na prática, você provavelmente procura voluntários e pode pagá-los para participarem do experimento. Se você for pesquisador universitário, provavelmente estes participantes serão alunos universitários. O mais importante na seleção dos participantes é criar dois grupos emparelhados, isto é, dois grupos de sujeitos tão parecidos um com o outro quanto

possível. Você pode conseguir isto pelo emparelhamento cuidadoso de características (sexo, idade, raça etc.), ou designando os sujeitos aos dois grupos *aleatoriamente*.

Um dos grupos será chamado *grupo experimental* e o outro *grupo de controle*. Os dois grupos serão testados para medir os níveis iniciais de preconceito contra negros. Por exemplo, ambos preencherão um questionário do tipo concordo/discordo a diversas frases manifestando preconceito antinegro. Espera-se que os dois grupos alcancem aproximadamente o mesmo escore geral neste *pré-teste*.

Em seguida, o grupo experimental assistirá ao filme documentário da história dos negros nos EUA. O grupo de controle *não* assistirá ao filme. Depois, os dois grupos serão novamente testados quanto a preconceito contra negros. A hipótese do pesquisador será confirmada se o grupo experimental demonstrar preconceito significativamente menor do que o grupo de controle no *pós-teste*.

O papel do grupo de controle em tal experimento é crítico. Serve à função de isolar o estímulo experimental como única fonte da mudança nos sujeitos do experimento. Se decorrer um longo tempo entre o pré e o pós-teste, o preconceito dos sujeitos pode diminuir graças a fatores externos ao experimento. O filme poderia então ser irrelevante para a redução observada do preconceito. Se isto ocorrer, o preconceito do grupo de controle deve diminuir também. A hipótese será confirmada somente se o preconceito do grupo experimental diminuir mais do que o do grupo de controle.

De forma semelhante, o grupo de controle ajuda o pesquisador a se precaver contra o efeito do experimento por si mesmo. É possível que o ato de testar e retestar torne os sujeitos mais sensíveis aos propósitos do estudo. Enquanto eles podem parecer relativamente preconceituosos no pré-teste, o próprio teste pode alertá-los para o fato de que o pesquisador quer descobrir quão preconceituosos são. Já que pouca gente quer ser identificada como racista, os sujeitos podem ter mais cuidado ao responderem o questionário pós-teste, procurando responder de forma a evitar aparecerem como racistas. Mas este fator deve operar igualmente no grupo de controle e no experimental, e o declínio diferencial no preconceito observado é o teste da hipótese.

Isolar as variáveis experimentais é a vantagem chave do experimento controlado. Ele apresenta também diversas desvantagens, contudo. Primeiro, o experimento controlado tipicamente

não fornece dados descritivos úteis. Se 20% dos dois grupos concordarem com uma afirmação preconceituosa, nada aprendemos sobre o percentual da população em geral que concordaria com ela, já que os sujeitos não são, via de regra, selecionados dessa população por métodos de amostragem aleatória. Se eles forem tirados de uma subpopulação especial, como estudantes, seu valor descritivo se reduz ainda mais.

Segundo, o experimento controlado representa um teste artificial da hipótese. A relevância do experimento para o mundo real está sempre sujeita a questionamento. No exemplo anterior, suponhamos que o filme documentário pareça reduzir significativamente o preconceito antinegros quando assistido como parte de um experimento científico num laboratório especial, com os sujeitos conscientes de participarem de um experimento. O filme não teria necessariamente o mesmo efeito se assistido pelo grande público na televisão ou em cinemas.

Finalmente, os achados podem não ter aplicabilidade generalizável a outros grupos da população. É concebível que o filme reduzisse preconceito entre universitários — se os sujeitos forem extraídos de estudantes —, mas não teria impacto em não-estudantes. Um exemplo não-experimental ilustra esta possibilidade. Durante anos, acreditou-se que, enquanto o preconceito contra negros era maior na classe operária do que nas classes média e alta, o anti-semitismo aumentava à medida que se subia na escala social. Tal conclusão se baseava numa série de estudos sobre preconceito realizados com estudantes universitários. Estudantes de famílias de classe alta pareciam consistentemente um pouco mais anti-semitas do que os de classes sociais relativamente mais baixas. Esse achado se deveu ao fato de que os alunos nos estudos provinham todos de uma faixa relativamente estreita de famílias de classes mais altas. O verdadeiro achado seria, então, que estudantes de classes mais altas eram ligeiramente mais anti-semitas do que alunos de classe média-alta. Estudos subseqüentes da população em geral indicaram, porém, que os respondentes de classe operária eram mais anti-semitas, assim como eram também mais antinegros.

As carências do experimento controlado podem ser reduzidas pela sofisticação das pesquisas, variação do desenho experimental e replicação em grupos muito diferentes de participantes. Além do mais, o experimento controlado pode ser especialmente valioso quando combinado com outros métodos, visando a um único tópico de pesquisa.

Análise de Conteúdo

Alguns tópicos de pesquisa são susceptíveis ao exame sistemático de documentos, como romances, poemas, publicações governamentais, música etc. Este método de pesquisa chama-se *análise de conteúdo*. A título de ilustração, suponha que você queira examinar as mudanças nas atitudes oficiais dos soviéticos para com os EUA. Você pode limitar o período do estudo aos anos entre 1950 e 1990 e decidir tomar os editoriais do *Pravda* como indicadores de política oficial.

Você pode, então, ou obter cópias de todos os editoriais do *Pravda* durante aquele período ou fazer um desenho amostral selecionando, digamos, cada décimo editorial. Cada editorial será examinado e receberá um escore correspondendo a posição favorável ou crítica aos EUA — ou a ser irrelevante para os EUA. Esta atividade exige um método sistemático de atribuir escores. Você tem de especificar que tipos de referências aos EUA serão consideradas e quantificadas como favoráveis e que tipos consideradas e quantificadas como críticas. É concebível que você deseje ponderar referências diferencialmente, em termos da força relativa do apoio ou da crítica. Neste caso, é necessário especificar como, nesta ponderação, os pesos diferenciais serão atribuídos.

Tendo atribuído escores aos editoriais, você pode agora agregar os escores de forma a caracterizar diferentes períodos temporais. Talvez você combine os escores por ano, relatando, por exemplo, que 12% dos editoriais em 1950 eram favoráveis aos EUA, 8% dos de 1951 eram a favor, e assim por diante. O padrão das diferenças ao longo do tempo representará a atitude oficial soviética para com os EUA.

Se bem que estes dados sirvam a uma finalidade descritiva útil, talvez você deseje ir além de descrições para explicar as flutuações observadas na atitude oficial soviética para com os EUA. Suponha que você acredite que as atitudes soviéticas foram, em grande parte, determinadas pelo montante de críticas que a URSS recebia dos EUA; isto é, quando os EUA criticavam a URSS, esta reagia acirrando as críticas aos EUA.

Você pode examinar esta possível explicação analisando o conteúdo de discursos e debates nos EUA. Você decidiria quais formas de comunicação nos EUA melhor refletiriam o estímulo hipotetizado e examinaria e daria escores a todos ou a uma amostra deles quanto a seu conteúdo de crítica anti-soviética. Os métodos de escore teriam de ser especificados e os

escores dados teriam de ser agregados para o mesmo período de tempo usado no estudo inicial. Você então compararia os dois padrões de flutuações, para determinar se a crítica americana à URSS era tipicamente seguida ou não de pronunciamentos anti-americanos no *Pravda*.

Análise de conteúdo tem a vantagem de fornecer um exame sistemático de materiais em geral avaliados de forma mais impressionística. Um repórter pode, por exemplo, ler os editoriais do *Pravda* ao longo do tempo — anotando mentalmente os que mencionam os EUA e talvez observando os editoriais que apóiam ou criticam fortemente os EUA —, criando uma impressão geral das flutuações na postura oficial. Através da análise de conteúdo sistemática, contudo, você se previne contra vieses inadvertidos que pode projetar no estudo. Você pode começar suspeitando que a URSS foi relativamente pró EUA durante um certo período, podendo, inconscientemente, dar mais atenção a conteúdos de editoriais favoráveis neste período, descontando os editoriais negativos descobertos. Seguindo rigorosamente um sistema preestabelecido de amostragem e atribuição de escores, você reduziria a influência de seus vieses.

Como todo método de pesquisa, a análise de conteúdo tem pontos fracos. Primeiro, o tipo de documentos selecionados para exame pode não ser a medida mais apropriada da variável estudada: é possível que editoriais do *Pravda* não sejam o melhor indicador da atitude soviética oficial para com os EUA. Discursos públicos de autoridades governamentais selecionadas ou pronunciamentos soviéticos no saguão das Nações Unidas poderiam ser mais apropriados. Na maioria dos casos, não há como determinar qual fonte representa o melhor foco do estudo. Este inconveniente pode ser minimizado examinando sistematicamente diferentes fontes e determinando se cada fonte leva à mesma conclusão.

Segundo, métodos de escore quase sempre contêm um elemento arbitrário. Alguns comentários em editoriais nos EUA são tão ambíguos que atribuir-lhes escores é difícil, talvez impossível. Outros comentários podem ser julgados favoráveis por um observador e críticos por outro. Talvez o pesquisador consiga reduzir esse problema pedindo avaliações independentes de outros pesquisadores experientes no assunto. Se vários observadores independentes concordarem sobre o escore dos comentários num editorial, a utilidade deste sistema de quantificação cresce. Em última análise, não há como garantir, com segurança absoluta, que comentários editoriais estão recebendo o escore certo. No lugar dessa confiabilidade absoluta, você

deve ser o mais específico possível na criação, execução e relato do sistema de escore. Os leitores do seu relatório da pesquisa devem saber, pelo menos, o que os escores representam exatamente, mesmo se discordarem da adequabilidade do sistema usado.

Análise dos Dados Existentes

Pesquisa científica não é igual a coleta e análise de dados originais. De fato, alguns tópicos de pesquisa podem ser estudados analisando dados já coletados e compilados. O exemplo clássico é o estudo do suicídio de Émile Durkheim.³ Interessado em descobrir as razões primárias do suicídio, Durkheim conduziu sua investigação sem coletar um só dado original. Ele testou um amplo espectro de hipóteses examinando taxas publicadas de suicídio em diferentes áreas geográficas. Por exemplo, examinou taxas diferentes de suicídio de protestantes e católicos, comparando as taxas de áreas predominantemente protestantes com as de áreas predominantemente católicas. Examinou os efeitos do clima, comparando taxas de regiões quentes com as de regiões mais frias.

A análise de dados agregados existentes tem a grande vantagem da economia. O pesquisador não precisa arcar com os custos de amostragens, entrevistas, codificações, recrutamento de sujeitos experimentais etc. Mas esta forma de pesquisa social tem duas importantes desvantagens.

Primeiro, o pesquisador fica limitado a dados já coletados e compilados, que podem não representar adequadamente as variáveis que o interessam. No entanto, este problema pode ser resolvido com uma boa dose de engenhosidade. Quando Samuel Stouffer examinou os efeitos da Grande Depressão nas famílias americanas e fora dos EUA, considerou vários indicadores possíveis de uma quebra hipotética de normas familiares tradicionais.⁴ Taxas de divórcio forneceram um indicador, mas Stouffer foi bem além, considerando as taxas de casamentos inter-religiosos, cerimônias civis em oposição a religiosas, casamentos fora dos Estados onde viviam os casais etc. Durkheim exibiu engenhosidade semelhante no estudo do suicídio.

O segundo problema tem a ver com o que é chamado de *falácia ecológica*. Sempre que você correlaciona variáveis geradas a partir de dados agregados, torna-se difícil determinar se a mesma relação entre as variáveis permanece verdadeira no nível dos indivíduos. Por exemplo, quando Durkheim encontrou taxas

de suicídio consistentemente mais altas em áreas predominantemente protestantes do que em áreas predominantemente católicas, não teve como determinar se os protestantes estavam cometendo suicídio. É concebível que católicos residindo em áreas predominantemente protestantes apresentassem as taxas de suicídio mais altas de todas. Também, quando Stouffer encontrou mais casamentos “impulsivos” durante os anos da Depressão, não teve como determinar se os envolvidos eram os mais afetados pela Depressão. Um exame engenhoso e lógico dos dados pode ajudar a reduzir este risco. Incentivo a leitura dos dois livros citados, para ver algumas das formas usadas por Stouffer e Durkheim para resolver o problema.

Estudo de Caso

O *estudo de caso* é uma descrição e explicação abrangentes dos muitos componentes de uma determinada situação social. Um estudo de comunidade, como o de W. Lloyd Warner sobre a comunidade de “Jonesville”, é um exemplo de estudo de caso.⁵ Num estudo de caso, você busca coletar e examinar o máximo de dados possíveis sobre o seu tema. Num estudo de comunidade, você aprende a história da comunidade, seus aspectos religiosos, políticos, econômicos, geográficos, composição racial etc. Você pode procurar determinar a estrutura de classes da comunidade, isto é, quem são os cidadãos mais importantes e poderosos e quem está embaixo da pirâmide.

Em resumo, você procura a descrição mais abrangente possível da comunidade e tenta determinar as inter-relações lógicas dos seus vários componentes. Qual foi o impacto do fechamento da madeireira no início dos anos 30? Como, em 1960, o prefeito reformista mudou as relações de poder na comunidade? O que causou a invasão dos Yuppies nos anos 80?

É importante saber que, quanto aos objetivos científicos, a abordagem estudo de caso de pesquisa social difere radicalmente dos outros métodos vistos até aqui. Enquanto a maioria das pesquisas visa diretamente o entendimento generalizado, o estudo de caso, busca inicialmente o entendimento abrangente de um só caso idiossincrático. Se a maioria das pesquisas procura limitar o número de variáveis consideradas, o estudo de caso tenta maximizá-lo. Em última análise, o pesquisador, fazendo um estudo de caso, busca, tipicamente, conhecimentos geralmente aplicáveis além do caso único estudado, mas por si só o estudo de caso não garante esta generalizabilidade.

Por exemplo, um estudo de caso pode sugerir que a entrada de trabalhadores estrangeiros não qualificados na comunidade teve como efeito geral promover a ascensão ocupacional dos trabalhadores nativos, com muitos assumindo cargos de supervisão dos novos migrantes. Este achado pode levá-lo a supor que você descobriu um princípio geral de mobilidade ocupacional e que mudança semelhante ocorreria na maioria das comunidades que sofressem influxo de mão de obra estrangeira não-qualificada. Um único estudo de caso, contudo, não poderia confirmar esta hipótese, e estudos adicionais em outras comunidades seriam necessários para confirmá-la.

Observação Participante

O termo *observação participante* refere-se a um método de coleta de dados em que o pesquisador torna-se participante no evento ou grupo social estudado. Você pode juntar-se a uma marcha de protesto como forma de colher dados sobre os outros participantes, ou entrar num grupo religioso que deseja estudar.

Na prática, como observador participante, você pode ou não revelar seu papel de pesquisador; esta decisão tem importantes implicações metodológicas e éticas. Se você admitir abertamente aos demais participantes que está realizando um estudo científico do grupo, sua presença pode afetar o fenômeno que pretende estudar. Saber que suas ações podem ser publicadas pode afetar o modo como os participantes agem. Por outro lado, se você esconder suas atividades de pesquisa e fingir ser um membro típico do grupo, estará sujeito às questões éticas relativas a engano. Além do mais, filiação aparentemente genuína pode apresentar problemas científicos. O que você fará se for eleito presidente do grupo ou se pedirem sua opinião sobre o que o grupo deve fazer em seguida? Sua reação afetará o que você está tentando estudar. Já que situações e finalidades de pesquisas variam tanto neste aspecto, não se pode dar uma diretriz geral, mas você deve estar consciente das questões envolvidas.

Como o estudo de caso, a observação participante visa colher muita informação detalhada. Mergulhando nos eventos sociais em andamento, você estará em condições de atingir profundidade de conhecimento muito maior do que seria possível, por exemplo, na análise de conteúdo ou no experimento. Ao mesmo tempo, porém, o observador participante encontra grande dificuldade em manter procedimentos sistemáticos de pesquisa. Já que será humanamente impossível observar e documentar tudo o que acontece, você terá de selecionar seus dados. Tentar

observar e registrar tudo pode resultar numa situação na qual vieses inconscientes se tornam a base da seleção. Por exemplo, se você começa a firmar a conclusão de que estudantes mulheres estão se tornando mais atuantes na direção de marchas de protesto, pode, inconscientemente, ficar mais pronto a anotar as instâncias que sustentam esta conclusão. Em última análise, o perigo maior é não conseguir dizer ao leitor quais critérios você usou na seleção e relato das observações, para que ele possa avaliar a adequabilidade dos critérios usados e a generalizabilidade das conclusões.

Voltaremos a considerar brevemente estes métodos de pesquisa nas conclusões do Capítulo 3, que examina a pesquisa de *survey*. Após descrever a natureza, as forças e fraquezas dos métodos de *survey*, nós os compararemos com os outros métodos descritos neste capítulo. As conclusões a que chegaremos podem ser reveladas aqui: (1) situações diferentes de pesquisas sociais requerem métodos diferentes e (2) o melhor formato muitas vezes é o que envolve o uso de diversos métodos focados no mesmo tópico.

Resumo

Neste capítulo, consideramos a possibilidade de aplicar métodos de investigação científica ao comportamento social. Mesmo anotando algumas considerações particulares pertinentes a esta aplicação, não encontramos nenhum obstáculo fundamental às ciências sociais. Embora as características da ciência em geral possam ser trazidas para o estudo da sociedade, isto não pode ser feito emulando a imagem tradicional da ciência.

Vimos também que pesquisadores sociais podem usar vários métodos de pesquisa para observar e entender o comportamento social. O capítulo seguinte aborda especificamente a pesquisa de *survey* como mais um método de pesquisa social. Logo veremos que as características básicas da ciência em geral aplicam-se igualmente à pesquisa de *survey* em particular.

Notas

¹ STOFFER, Samuel A. et al. *The American Soldier*. Princeton, NJ: Princeton University Press, 1949.

² Estes achados são discutidos no Capítulo 15, "O Modelo de Elaboração".

³ DURKHEIM, Émile. *Suicide: a Study in Sociology*. Trad. George Simpsen. New York: Free Press, 1951.

⁴ STOUFFER, Samuel A. *Social Research to Test Ideas*. New York: Free Press, 1962. p.134-153.

⁵ WARNER, W. Lloyd. *Democracy in Jonesville*. New York: Harper & Row, 1949.

Leituras Adicionais

BABBIE, Earl. *Observing Ourselves: Essays in Social Research*. Belmont, CA: Wadsworth, 1986.

COOK, Thomas D., CAMPBELL, Donald T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally, 1979.

DURKHEIM, Émile. *The Rules of Sociological Method*. Trad. Sarah Solovay e John Mueller. New York: Free Press, 1962. (Editado por George Catlin.)

HOLSTI, Ole. *Content Analysis for the Social Sciences and Humanities*. Reading: Addison-Wesley, 1969.

HUNT, Morton. *Profiles of Social Research: The Scientific Study of Human Interactions*. New York: Russell Sage Foundation, 1958.

LOFLAND, John. *Analyzing Social Settings*. Belmont, CA: Wadsworth, 1984.

McCALL, George, SIMMONS, J. L. (Ed.). *Issues in Participant Observation: a Text and Reader*. Reading, MA: Addison-Wesley, 1969.

RAY, William, RAVIZZA, Richard. *Methods Toward a Science of Behavior and Experience*. Belmont, CA: Wadsworth, 1985.

WALLACE, Walter. *The Logic of Science in Sociology*. Chicago: Aldine-Atherton, Inc., 1971.

Capítulo 3

Pesquisa de *Survey* como Método das Ciências Sociais

Este livro aborda fundamentalmente a lógica e as capacidades de um método de pesquisa: o *survey*. Antes de continuar, descreveremos rapidamente os componentes do *survey* típico, se bem que veremos mais tarde que métodos de *survey* podem ser aplicados a uma vasta gama de tópicos e desenhos.

Suponha o interesse de estudar atitudes de estudantes de uma universidade. Seleciona-se uma *amostra* de centenas de estudantes da totalidade do corpo discente. Um *questionário* é elaborado para obter informações (por exemplo, atitudes) relevantes ao tema investigado. Os questionários são aplicados à amostra de alunos, através de entrevistas pessoais, por telefone ou por correio. As respostas de cada estudante da amostra são codificadas de forma padronizada e registradas de forma *quantitativa*. A seguir, transferem-se as respostas codificadas de cada aluno para disquetes de computador. Os registros padronizados de todos os alunos são submetidos a uma análise agregada, para fornecer descrições dos estudantes da amostra e determinar correlações entre diferentes respostas. As conclusões descritivas e explicativas obtidas pela análise são, então, generalizadas para a população da qual a amostra foi selecionada, neste caso todo o corpo estudantil.

Breve História da Pesquisa de *Survey*

Pesquisa de *survey* apresenta semelhanças suficientes com outros métodos de pesquisa para ter uma história bem longa. Em particular, *surveys* são muito semelhantes a censos, sendo a diferença principal entre eles que um *survey*, tipicamente, examina uma amostra de população, enquanto o *censo* geralmente implica uma enumeração da população toda. Censos remontam, no mínimo, à antiga civilização egípcia, onde os governantes julgavam útil colher dados empíricos descrevendo seus súditos. As funções políticas da pesquisa de *survey* continuam até hoje, com a continuação dos censos, a aparição de pesquisas políticas encomendadas por candidatos e usos feitos por sociólogos políticos. Um dos primeiros usos políticos do *survey* de atitudes ocorreu em 1880. Um sociólogo político alemão enviou questionários pelo Correio a 25.000 trabalhadores franceses, para averiguar em que grau eram explorados pelos patrões. Era um questionário bem longo, incluindo itens como:

Seu empregador ou representante dele lança mão de desonestidade para privá-lo de parte do seu salário?
Se você se abastece nos armazéns da empresa, a qualidade dos produtos serve de pretexto para deduções fraudulentas do seu salário?¹

O pesquisador em questão era Karl Marx. 25.000 questionários foram enviados e não há qualquer registro de que algum tenha sido retornado.

Informa-se também que Max Weber, um dos fundadores da sociologia moderna, usou métodos de pesquisa de *survey* no seu estudo sobre a ética protestante. Além do exame histórico comparativo do desenvolvimento econômico, ele também estudou operários protestantes e católicos, para obter dados confirmatórios em nível individual.²

Mas a maioria da pesquisa de *survey* contemporânea foi realizada neste século por pesquisadores americanos. O estado atual do método resulta de importante trabalho de desenvolvimento em três setores distintos da sociedade americana. Primeiro, o trabalho ininterrupto do U. S. Bureau of Census fez importantes contribuições aos campos de *amostragem e coleta de dados*. Apesar do Bureau ser mais conhecido pelo *recenseamento decenal* da população americana, a grande maioria das atividades da agência tem a ver com uma série contínua de *surveys* amostrais que constantemente atualizam dados demográficos e econômicos entre os recenseamentos.

O Bureau desempenhou um papel singularmente importante no desenvolvimento de definições padronizadas de amostragens e de métodos para a implementação destas definições no campo. Ao mesmo tempo, os dados gerados pelo Bureau constituíram um valioso recurso para o desenvolvimento de *desenhos de amostras em surveys específicos*. (O Capítulo 6 ilustra este ponto com detalhes.)

A segunda fonte de desenvolvimento foram as atividades de firmas comerciais de pesquisas de opinião, entre elas as organizadas por George Gallup, Elmo Roper, Louis Harris e outros. Estas firmas foram uma fonte contínua de fundos para apoiar o desenvolvimento e o uso de métodos de *survey*, principalmente nas áreas de *marketing de produtos e de pesquisas de políticas*. Em épocas em que não havia recursos para sustentar pesquisas acadêmicas de *survey*, estas empresas comerciais fizeram *experimentação com métodos de amostragem, redação de perguntas, técnicas de coleta de dados* e outros aspectos de pesquisa de *survey*. Além disso, pesquisas comerciais de opinião têm sido uma fonte de dados de valor incalculável para *análise secundária*, e numerosos livros e artigos acadêmicos foram publicados relatando análises mais detalhadas de dados coletados inicialmente para fins descritivos e comerciais.³

Terceiro, o refinamento científico da pesquisa de *survey*, sobretudo métodos sofisticados de análise, foi em grande parte o produto de algumas poucas universidades americanas. Mais especificamente, dos esforços de dois homens e, mais tarde, de centros de pesquisas de *survey*. Samuel A. Stouffer e Paul F. Lazarsfeld devem ser considerados pioneiros da pesquisa de *survey* como ela é conhecida hoje.

A obra pioneira de Stouffer consistiu, em grande parte, em tentativas de aplicar métodos empíricos de pesquisa social a problemas sociais, começando com análises dos efeitos da Depressão nos EUA e com coleta de dados sobre o *status* dos americanos negros nos anos 30.⁴ Na II Grande Guerra, Stouffer dirigiu o Departamento de Informação e Educação do Exército dos EUA, reunindo um grupo emergente de cientistas sociais para examinar temas relevantes ao sucesso americano na guerra.⁵ Durante a era McCarthy, no início dos anos 50, Stouffer realizou *surveys* nacionais para examinar os efeitos da cruzada anticomunista.⁶

Nestes e em inúmeros outros estudos, Stouffer procurou desenvolver métodos científicos de pesquisa empírica, adequados ao exame dos fenômenos sociais. Seus esforços contavam com um sólido treinamento científico pelos estatísticos britânicos

Karl Pearson e R. A. Fisher, bem como de sua criatividade inata. O legado de Stouffer é evidente em desenhos de pesquisas, métodos de amostragem, desenhos de questionários, lógica de análise e outros métodos de *survey* ainda hoje usados correntemente.

Paul Lazarsfeld trouxe para os EUA uma formação intelectual européia. Como Stouffer, interessava-se pelo estudo de fenômenos sociais como liderança, comunicações, comportamento econômico e profissões. No exame desses tópicos, Lazarsfeld, como Stouffer, desenvolveu técnicas rigorosas para aplicar métodos empíricos a questões sociais. Na área do comportamento político, ele achou importante examinar o voto como processo e não como evento singular. Para isso, projetou e realizou *estudos de painéis*, reentrevistando um determinado grupo de respondentes em ocasiões diferentes durante uma campanha política, acompanhando assim mudanças nas intenções de voto ao longo do tempo. Não contente em descrever, Lazarsfeld combinou este esforço com um exame cuidadoso dos fatores demográficos e sociais associados às mudanças observadas.⁷

As contribuições de Lazarsfeld para o desenvolvimento da pesquisa de *survey* foram tantas que não se poderia enumerá-las todas aqui; mas três merecem ser citadas. Primeiro, a carreira de Lazarsfeld correu paralela ao desenvolvimento técnico dos equipamentos mecanizados de processamento de dados — inicialmente os perfuradores e separadores de cartões e depois os computadores. Deve-se creditar a ele o reconhecimento do potencial de uso desses equipamentos na pesquisa social analítica e haver encaminhado jovens cientistas sociais para realizar esse potencial.

Segundo, Lazarsfeld usou equipamentos mecanizados de processamento de dados para elucidar e formalizar a lógica da análise de *survey*. Enquanto a engenhosidade de Stouffer o levou a sugerir razões para relações observadas nos estudos sobre o Exército americano na II Guerra Mundial, foi Lazarsfeld quem formalizou estas razões num modelo lógico destas relações, mostrando como o modelo podia ser implementado na prática.⁸ (Este tópico é discutido mais detalhadamente no *modelo de elaboração*, Capítulo 15.)

A terceira grande contribuição de Lazarsfeld foi a criação do centro permanente de pesquisas de apoio aos métodos de *survey*, começando com a organização do Bureau for Applied Social Research na Columbia University. À semelhança do grupo de pesquisas de Stouffer no Exército, o Bureau reuniu e treinou uma legião de jovens e brilhantes cientistas sociais;

o Bureau, porém, conseguiu sobreviver no pós-guerra. Organizações semelhantes, estabelecidas depois, incluem o Survey Research Center na Universidade da Califórnia em Berkeley, o National Opinion Research Center na Universidade de Chicago, o Institute for Social Research da Universidade de Michigan, o Survey Research Center na Universidade da Califórnia em Los Angeles, e o Survey Research Laboratory da Universidade de Wisconsin.

Além destas instituições, há atualmente institutos e centros semelhantes em campi universitários espalhados pelos EUA e pelo mundo, servindo a diversas funções. Primeiro, embora os departamentos acadêmicos freqüentemente ofereçam ensino em sala de aula sobre métodos de *survey*, os centros de *survey* caracteristicamente permitem que estudantes recebam treinamento prático de aprendizado, muitas vezes trabalhando como assistentes de pesquisa. Segundo, inúmeros *surveys* são realizados em tais centros, bancados por subvenções governamentais e de fundações ou encomendados por clientes comerciais. Os centros também dão consultorias e fazem outras prestações de serviços a pesquisadores usuários de metodologias de *survey*. Finalmente, esses centros desempenham um papel importante no desenvolvimento metodológico da pesquisa de *survey*. Enquanto um pesquisador individual conduzindo um projeto de pesquisa pode relutar em experimentar técnicas alternativas de coleta de dados, uma unidade organizada de pesquisas é capaz de fazê-lo em diversos estudos.

Antes de concluir esta breve história do *survey*, devemos mencionar o papel desempenhado pelas associações profissionais. Ao longo dos anos, associações profissionais, cujos membros freqüentemente utilizam métodos de *survey*, têm promovido discussões de novas técnicas e descobertas empíricas em suas reuniões e publicações especializadas. Talvez as três associações principais neste aspecto sejam a American Sociological Association, a American Political Science Association e a American Marketing Association. Também a AAPOR (American Association for Public Opinion Research) tem sido de grande relevância para os pesquisadores de *survey*, reunindo praticantes dos setores acadêmico, comercial e governamental. A *Public Opinion Quarterly*, publicação trimestral da AAPOR, é o veículo chave da pesquisa de *survey*.

Capítulos subseqüentes rasteiam a história de certos componentes da pesquisa de *survey*. Contra o pano de fundo desta breve visão geral, vamos examinar agora o lugar da pesquisa de *survey* no contexto geral da ciência.

Características Científicas da Pesquisa de *Survey*

Como já foi observado, a pesquisa de *survey* é só uma de muitas ferramentas de pesquisa disponíveis para pesquisadores sociais. Vale a pena repetir que métodos de *survey* não são apropriados a muitos tópicos de pesquisa, nem oferecem necessariamente a melhor abordagem a assuntos aos quais podem ser aplicados. Mesmo assim, o *survey* pode ser usado vantajosamente no exame de muitos temas sociais e é particularmente eficaz quando combinado com outros métodos.

Mais importante, acredito que a pesquisa de *survey* oferece o melhor exemplo para ensinar metodologia nas ciências sociais. Se você compreende bem a lógica e as capacidades da pesquisa de *survey*, está muito bem preparado para aprender e usar outros métodos de pesquisa social. Digo isto porque vejo a pesquisa de *survey* como um crustáceo: todos os ossos estão do lado de fora. Nos capítulos seguintes examinaremos cuidadosamente todas as aproximações, compromissos e outras carências da pesquisa de *survey*. Como discutiremos em breve, a pesquisa de *survey* tem uma função pedagógica, porque todas as deficiências ficam mais claras nela do que em outros métodos de pesquisa social, permitindo assim avaliações mais conscientes de suas implicações.

A Pesquisa de *Survey* É Lógica

Pesquisa de *survey* é guiada por todas as restrições lógicas já abordadas nos dois capítulos anteriores. Além do mais, na prática, os dados de *survey* facilitam a aplicação cuidadosa do pensamento lógico. Embora este tópico vá ser explorado com mais detalhes no Capítulo 15, um exemplo será apropriado agora.

Num estudo sobre envolvimento de episcopais na sua igreja,⁹ Charles Glock, Benjamin Ringer e eu descobrimos que mulheres de classe social baixa eram mais atuantes do que mulheres de *status* social mais alto. Quanto mais alta a classe social, menor a participação. Procuramos explicar este fenômeno em termos da *teoria da privação do envolvimento na igreja* — pessoas a quem se nega prestígio e *status* na sociedade secular se envolvem mais provavelmente na vida da igreja como fonte alternativa de gratificação. Sugerimos que a maior intensidade de atividade na igreja de mulheres de classe social baixa reflete a maior privação delas na sociedade secular. Por extensão, sugerimos que, entre as mulheres que gozavam de algum grau de gratificação de *status* secular, classe social

não afetaria envolvimento na igreja. Os dados disponíveis indicavam se as entrevistadas na pesquisa já haviam ocupado algum cargo em instituição secular. Esperávamos, logicamente, que, entre mulheres com cargo na vida secular, classe social não afetaria envolvimento na igreja. Esta expectativa foi testada empiricamente e provou estar correta.

O formato da pesquisa de *survey* muitas vezes permite desenvolvimento e teste rigorosos, passo a passo, de tais explicações lógicas. Além disso, o exame de centenas e até milhares de entrevistas de *survey* permite testar proposições complexas envolvendo diversas variáveis em interação simultânea.

A Pesquisa de *Survey* É Determinística

Sempre que o pesquisador de *survey* procura explicar as razões para e as fontes de eventos, características e correlações observados, a investigação deve assumir uma postura determinística. O fato do formato *survey* permitir elaboração clara e rigorosa de um modelo lógico clarifica o sistema determinístico de causa e efeito.

Além disso, a disponibilidade de numerosos casos e variáveis permite ao analista documentar processos causais mais elaborados. Podemos ir além da observação inicial de uma correlação entre variável independente e dependente, para examinar o papel de diversas variáveis intervenientes. Assim, voltando ao exemplo acima, notamos que classe social não afeta diretamente envolvimento de mulheres episcopais na igreja, tendo efeito através da variável interveniente gratificação secular. Mulheres de classes sociais baixas têm probabilidade menor de serem eleitas para ocupar cargos em organizações seculares. Conseqüentemente, estas mulheres, como grupo, se envolvem mais na vida da igreja, mas, se *recebem* gratificação secular, sua classe social não afeta envolvimento na igreja.

A Pesquisa de *Survey* É Geral

Surveys amostrais quase nunca são realizados para descrever a amostra particular estudada. São realizados para se entender a população maior da qual a amostra foi inicialmente selecionada. Assim, a Gallup pode entrevistar 1.500 eleitores americanos para predizer como dezenas de milhões votarão no dia da eleição. *Surveys* amostrais de consumidores visam conhecer preferências e comportamentos de consumidores em geral.

Do mesmo modo, análises explicativas em pesquisas de *survey* visam desenvolver proposições gerais sobre

comportamento humano. O formato *survey* permite este objetivo científico genérico de duas maneiras. Primeiro, o grande número de casos estudados num *survey* permite que achados possam ser replicados entre vários subconjuntos da amostra do *survey*. Se uma correlação geral é encontrada, por exemplo, entre escolaridade e padrões de compra, os pesquisadores conseguem determinar facilmente se esta relação ocorre igualmente entre homens e mulheres, protestantes e católicos, brancos e negros, pessoas de regiões geográficas diferentes etc. Replicar um achado entre subgrupos diferentes fortalece a certeza de que ele representa um fenômeno geral na sociedade. Segundo, o relato cuidadoso da metodologia de um *survey* facilita réplicas posteriores por parte de outros pesquisadores e/ou entre outras amostras e subgrupos. Deste modo, a generalizabilidade dos achados pode ser testada e retestada.

A Pesquisa de *Survey* É Parcimoniosa

Como pesquisadores de *survey* têm à disposição um grande número de variáveis, estão em posição excelente para examinar cuidadosamente a importância relativa de cada uma. Como todo cientista, pesquisadores de *survey* querem obter o máximo de compreensão com o menor número de variáveis. Não precisam adivinhar as variáveis mais relevantes no desenho inicial do estudo, contudo (ou, pelo menos, não tanto quanto pesquisadores usando outros métodos). Já que o formato *survey* permite obter muitas variáveis que podem ser quantificadas e processadas por computador, os pesquisadores de *survey* podem construir vários modelos explicativos e então selecionar o que melhor servir seus propósitos.

A Pesquisa de *Survey* É Específica

No contexto da característica científica da especificidade, a natureza tipo crustáceo da pesquisa de *survey* é extremamente relevante. Ironicamente, esta característica também expõe os métodos de *survey* a um máximo de críticas.

Suponha que um *survey* conclua que conservadorismo político e preconceito contra mulheres se relacionam positivamente. Tal conclusão se baseou em definições operacionais específicas de conservadorismo e de preconceito contra mulheres. A medição de cada variável foi construída a partir de respostas específicas a itens específicos de questionário codificados e quantificados de forma específica. Visto que todos estes detalhes constam do relatório da pesquisa, o leitor crítico está

apto a descobrir que respondentes descritos como “muito preconceituosos” são os que concordaram com cinco itens específicos do questionário. Os críticos podem objetar que este critério não capturou plenamente o significado de “preconceito” como eles — e, por extensão, outros — entendem o termo. Podem, então, decidir que a conclusão geral não tem fundamento.

De certo modo, tal resposta seria válida. A conceituação e a medição de variáveis estão no âmago da prática científica, e se variáveis não são adequadamente conceituadas e medidas, correlações observadas entre elas podem não fazer sentido. Assim, se observadores independentes discordarem sobre como variáveis foram medidas, podem discordar logicamente da conclusão geral.

Muitas vezes se passa por alto esta facilidade com que leitores críticos podem chegar a, e talvez, documentar seus desacordos. Porque os analistas de *survey* descreveram precisamente como desenvolveram e fizeram suas medidas, o leitor sabe exatamente o que estas medidas representam. A superficialidade e as aproximações que fazem parte de toda pesquisa científica são, simplesmente, mais aparentes num *survey*.

Em contraste, suponha que você estudou a relação entre conservadorismo e preconceito contra mulheres mergulhando como observador participante em vários grupos políticos. Você pode relatar que suas observações sugerem que os *liberais* têm mais preconceito contra mulheres que os conservadores. Em seu relatório, você pode descrever qualitativamente como distinguiu conservadores de liberais e preconceituosos de não-preconceituosos, captando a profundidade de significado que estes termos têm no uso da linguagem comum, evitando assim a aparente superficialidade de itens de questionários e medidas de *survey*.

Infelizmente, o observador independente não tem como saber precisamente como suas descrições foram usadas na prática, nem como julgar até que ponto seu viés, talvez inconsciente, pode ter afetado suas classificações de pessoas em termos de política e preconceito. Isto não quer dizer que suas observações e designações foram tendenciosos ou que sua conclusão estava incorreta. É importante, porém, o fato de que nem você nem o observador independente são capazes de decidir a respeito. (Deve-se observar, é claro, que, na prática, a observação participante varia em rigor e especificidade. O ponto é que o *próprio método* força o pesquisado de *survey* a ser explícito.)

A pesquisa científica visa a conceituações e medidas cada vez mais sofisticadas e úteis, mas a cada passo do caminho

os métodos usados devem ser especificados. Por sua própria natureza, a pesquisa de *survey* se encaixa como uma luva nesta característica.

Conclusão

As demais características da ciência são igualmente relevantes no contexto da pesquisa de *survey*. Onde a ciência deve ser *empiricamente verificável*, a pesquisa de *survey* oferece um método de verificação empírica. Os comentários nas seções anteriores também ilustram como o formato *survey* é adequado à *intersubjetividade*.

Finalmente, métodos de pesquisa de *survey* facilitam a *abertura da ciência*. Já que a pesquisa de *survey* envolve a coleta e *quantificação* de dados, os dados coletados se tornam fonte permanente de informações. Um corpo de dados de *survey* pode ser analisado pouco depois da coleta e confirmar uma determinada teoria de comportamento social. Se a própria teoria sofrer modificações mais tarde, é sempre possível retornar ao conjunto de dados e reanalisá-los sob a nova perspectiva teórica. Esta nova análise não poderia ser realizada tão facilmente no caso de métodos de pesquisa menos rigorosos e menos específicos.

Comparação do Survey com Outros Métodos

As seções anteriores mostraram diversas comparações entre pesquisa de *survey* e outros métodos científicos de pesquisa social. Foi salientado que uma investigação abrangente se beneficiaria com o uso de métodos diferentes focados num só tópico. A esse respeito, três pontos devem ser ressaltados.

Primeiro, a lógica do experimento controlado pode se constituir num guia útil à lógica da análise do *survey*. Onde o experimentador isola a variável experimental pelo uso de grupos emparelhados ou randomizados — grupos experimental e de controle —, o analista do *survey* busca o mesmo fim controlando as variáveis *post-facto*. Por exemplo, o experimentador pode garantir que os grupos experimental e de controle tenham a mesma distribuição de sexos, para evitar uma possível influência desta variável no experimento. O analista do *survey* consegue isto garantindo que subgrupos na amostra tenham a mesma distribuição de sexos ou testando separadamente a relação observada entre homens e mulheres.

Todavia, o objetivo lógico de isolar variáveis relevantes pela exclusão da influência de variáveis estranhas é a mesma para os dois métodos.

Segundo, a codificação das respostas do *survey* é essencialmente uma instância de análise de conteúdo. Frequentemente, o pesquisador de *survey* fará perguntas *abertas* que pedem resposta nas próprias palavras do respondente. Tais respostas, porém, devem sempre ser codificadas em tipos de respostas. Pode, então, ser necessário codificar um conjunto de respostas como apoiando ou opondo-se em geral a uma determinada lei. A esse respeito, os pesquisadores de *survey* têm muito a aprender com as experiências e métodos dos analistas de conteúdo.

Terceiro, as entrevistas de *survey* podem beneficiar-se com a experiência dos observadores participantes. Métodos para obter *rapport*, manter neutralidade e fazer observações precisas são importantes para ambas as atividades.

A Pesquisa de Survey É Realmente Científica?

Em vista das críticas por vezes feitas aos cientistas sociais, aos pesquisadores de *survey* e aos sociólogos praticantes de *survey* em particular, permito-me concluir este capítulo um tanto chauvinistamente. Há vários anos, Allan Mazur publicou um artigo intitulado “A Menor das Ciências”,¹⁰ onde levantava a questão de se a sociologia podia ser chamada de ciência. Como sugere o título do artigo, ele concluiu que essa designação mal se justificava. Não é de se surpreender que o artigo suscitasse respostas imediatas e muitas vezes acaloradas.

Como a maioria dos seus comentaristas, positivos ou negativos, Mazur também concordaria que nenhuma resposta definitiva é possível na questão de se uma disciplina como a sociologia é ou não uma ciência. Como vimos no Capítulo 1, qualquer definição de ciência deve ser arbitrária; portanto, não pode haver respostas absolutas. Ao mesmo tempo, a questão merece ser discutida, mesmo não podendo ser respondida.

Penso que, no futuro, pesquisadores olharão para trás, para essa era, e concluirão que o uso da pesquisa de *survey* por sociólogos e outros cientistas sociais foi um período crítico *no desenvolvimento da ciência em geral*. Esta expectativa hipotética não é um mero vago sonho megalomaniaco de um pesquisador de *survey* que, por coincidência, foi treinado como sociólogo. Os comentários de dois outros pesquisadores ilustram a razão desta expectativa.

O economista Daniel Suits, falando há alguns anos em Honolulu, numa conferência de pesquisadores e planejadores governamentais, abandonou a costumeira terminologia acadêmica e não falou nem das “ciências duras” (como física e química) nem das “ciências macias” (como sociologia, ciência política e pesquisa de mercado), preferindo distinguir entre ciências “duras” e “fáceis”. Seu ponto era que físicos podem realizar pesquisas científicas facilmente, dado o assunto com que trabalham, enquanto cientistas sociais têm um osso muito mais duro para roer.

Martin Trow rasteou algumas das implicações desta situação num contexto diferente.¹¹ Abordando os usos dos métodos de *survey* no campo da educação, Trow notou uma anomalia interessante. Ele observou primeiro que o ambiente educacional é quase ideal para pesquisa de *survey*: os entrevistados são articulados, conhecem questionários, são fáceis de enumerar e amostrar, e os questionários podem ser administrados em condições controladas na sala de aula. Trow indagou, então, por que a maioria dos *surveys* educacionais são tão triviais e não-sofisticados? Sua resposta foi que as condições são *boas demais*. Pesquisadores educacionais nunca tiveram que lidar com condições imperfeitas de pesquisas e como nunca foram obrigados a fazer compromissos e aproximações no desenho e na execução das pesquisas, nunca tiveram dificuldades com a lógica básica da pesquisa científica de *survey*. Nunca tendo de se afastar do ideal óbvio, nunca conseguiram entender plenamente por que ela era considerada ideal.

Penso que esta observação de Trow tem um significado bem mais geral. As próprias adversidades com que as ciências sociais se deparam requerem a criação de um sistema lógico de entendimento mais sofisticado. Além disso, a pesquisa de *survey*, por causa de sua especificidade operacional, pode prover o veículo mais útil para lidar com tais adversidades de forma rigorosa e sistemática. Alguns exemplos esclarecem mais este ponto.

Quando químicos desejam estudar as propriedades do hidrogênio, provavelmente não se preocupam muito com técnicas de amostragem. Já que um átomo de hidrogênio é igual a qualquer outro, qualquer átomo serve para o estudo. Mas cientistas sociais não podem estudar qualquer indivíduo ou grupo de indivíduos que lhes sejam convenientes. Os químicos podem generalizar seus achados para qualquer hidrogênio baseados apenas no estudo de átomos convenientes, mas os cientistas sociais devem desenvolver um entendimento mais sofisticado do conceito lógico de generalizabilidade, bem como dos métodos operacionais para alcançá-la. Evidentemente,

outros cientistas tem problemas de amostragem — geneticistas, por exemplo —, e cientistas sociais têm encontrado orientação nos trabalhos de outras disciplinas. Não obstante, os problemas de amostragem e de generalizabilidade são maiores no estudo do comportamento social. Pesquisa de *survey* é um veículo excelente para o desenvolvimento de métodos úteis e, por extensão, de entendimento mais amplo.

O ato de medir é outro exemplo dos problemas da pesquisa de *survey*. Suponha que você está tentando medir algo como preconceito. Não há conceituação clara ideal a partir da qual trabalhar, nem há operação de medição fácil, disponível para se aproximar desse ideal. Se você usar técnicas de *survey*, o problema é ainda mais básico. Você pode descobrir que um item do questionário produz uma resposta preconceituosa de um respondente, enquanto outro item gera uma resposta não-preconceituosa. Mais ainda, uma pequena variação na redação de um item pode afetar a resposta obtida.

Do mesmo modo, os cientistas sociais reconhecem que a simples presença dos pesquisadores pode afetar os entrevistados. Pedir uma opinião pode cristalizar mais ainda a opinião que existia antes da entrevista; alguns participantes formam opiniões na hora. Seria como se nas ciências físicas uma barra de metal esticasse ou encolhesse de comprimento quando um físico chegasse para medi-la. Claramente, a situação enfrentada pelos cientistas sociais requer uma compreensão mais sofisticada de conceituação e medição. A multiplicidade de variáveis relevantes e a natureza complexa e probabilística de causação no comportamento social pedem um entendimento mais sofisticado do que a ciência realmente é. A lista de exemplos seria interminável.

Neste contexto, observe-se que as “ciências fáceis” não são tão claras e diretas como se imagina. Por exemplo, os físicos sabem que não podem simultaneamente medir o lugar e a velocidade de um objeto; medir um afeta o outro. Cientistas médicos são obrigados a usar *placebos* (pílulas de açúcar, por exemplo) para testar drogas novas porque alguns pacientes parecem melhorar quando acreditam estar recebendo uma poderosa droga nova e/ou porque os próprios pesquisadores médicos podem ver desenvolvimentos otimistas em pacientes que eles acreditam estar recebendo a nova droga. Ao mesmo tempo, físicos nucleares afirmam que partículas se movimentando numa certa direção sem receber impulso adicional *provavelmente* continuarão a se movimentar naquela mesma direção.

Em suma, as ciências sociais são um exemplo ímpar do que o físico Heinz Pagels chama de *as ciências da complexidade*.¹²

Pagels argumenta que o desenvolvimento do computador representa uma inovação na possibilidade de lidar com sistemas complexos, e em nenhuma área isto é mais verdade do que nas ciências sociais. Sugiro até que o computador será para as ciências sociais o que o telescópio foi para a astronomia e o microscópio para a biologia. Neste processo, as ciências sociais expandirão radicalmente nossa visão do que pode ser a *ciência*.

Notas

¹ BOTTOMORE, T. B., RUBEL, Maximilien (Ed.). *Karl Marx: Selected Writings in Sociology and Social Philosophy*. New York: McGraw-Hill, 1956. Ver página 208 para as questões citadas.

² LAZARFELD, Paul F., OBERSCHALL, Anthony R. Max Weber and Empirical Research. *American Sociological Review*, p.185-199, abr. 1965.

³ Ver, por exemplo, HYMAN, Herbert. *Secondary Analysis of Sample Survey*. New York: John Wiley & Sons, 1972.

⁴ Uma ótima revisão da obra de Stouffer pode ser encontrada no livro póstumo de STOUFFER, Samuel A. *Social Research to Test Ideas*. New York: Free Press, 1962.

⁵ Ver STOUFFER, Samuel A. et al. *The American Soldier: Studies in Social Psychology in World War II*. Princeton: Princeton University Press, 1949 e 1950. v.I-V.

⁶ STOUFFER, Samuel A. *Communism, Conformity and Civil Liberties*. Garden City: Doubleday & Company, Inc., 1955.

⁷ LAZARFELD, Paul F. *The People's Choice*. New York: Columbia University Press, 1948.

⁸ Ver KENDAL, Patricia, LAZARFELD, Paul F. Problems of Survey Analysis. In: MERTON, Robert, LAZARFELD, Paul (Ed.). *Continuities in Social Research Studies in the Scope and Method of the American Soldier*. New York: Free Press, 1950.

⁹ GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley, CA: University of California Press, 1967.

¹⁰ MAZUR, Allan. The Littlest Science. *American Sociologist*, p.195-200, ago. 1968.

¹¹ TROW, Martin. Education and Survey Research. In: GLOCK, Charles Y. (Ed.). *Survey Research Methods in the Social Sciences*. New York: Russell Sage Foundation, 1967. p.315-375.

¹² PAGELS, Heinz. *The Dreams of Reason*. New York, NY: Simon & Schuster, 1988.

Leituras Adicionais

BAINBRIDGE, William Sims. *Survey Research: A Computer-Assisted Introduction*. Belmont, CA: Wadsworth, 1989.

DILLMAN, Don A. *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons, 1978.

GLOCK, Charles Y. (Ed.). *Survey Research in the Social Sciences*. New York: Russell Sage Foundation, 1967.

LAZARFELD, Paul F. "Introduction" to Samuel A. Stouffer. In: STOUFFER, Samuel A. (Ed.). *Social Research to Test Ideas*. New York: Free Press, 1962. p.XV-XXXI.

PUBLIC OPINION QUARTERLY, v.51, n.4 (Parte 2). Edição de 50º aniversário (inverno 1987) com vários artigos revendo a história da pesquisa de *survey*.

Desenho de Pesquisa de *Survey*

2

CAPÍTULO 4
TIPOS DE DESENHOS DE PESQUISAS

CAPÍTULO 5
A LÓGICA DA AMOSTRAGEM DO *SURVEY*

CAPÍTULO 6
EXEMPLOS DE DESENHOS DE AMOSTRAGEM

CAPÍTULO 7
CONCEITUAÇÃO E DESENHO DE INSTRUMENTOS

CAPÍTULO 8
CONSTRUÇÃO DE ÍNDICES E ESCALAS

Os cinco capítulos da Parte 2 abordam diversos aspectos do desenho do *survey*. Examinaremos alguns dos diferentes tipos possíveis de *surveys*, a seleção de amostras e a importante questão da medição.

Há uma tendência a considerar a análise de dados de *survey* como mais desafiadora e excitante do que o desenho do *survey* e a coleta de dados. Confesso sentir isto e desconfio que a maioria dos outros pesquisadores de *survey* concordariam. Durante a análise, começa-se a ganhar entendimento do assunto estudado e pode-se compartilhar as descobertas com os colegas.

Mas é errado supor que o desenho do estudo é menos desafiador do que a análise ou requer menos brilho e engenhosidade. Frequentemente, aliás, a engenhosidade se torna necessária na análise do *survey* justamente porque esteve ausente no desenho e na execução da pesquisa. Mesmo podendo achar satisfação na superação de problemas de desenho através de manipulações analíticas inteligentes, deve continuar sendo

motivo de embaraço haver tais problemas; muitas vezes se descobre haver problemas de desenho que nenhuma engenhosidade resolve.

Quando fica mais competente nos diversos aspectos do desenho do *survey*, você compreende que ele requer as mesmas capacidades lógicas de solução de problemas que a análise do *survey*. Neste aspecto, o desenho é igualmente excitante e desafiador.

Diversos desenhos de pesquisas são descritos no Capítulo 4, para familiarizá-lo com o leque de opções e ajudá-lo a escolher quais são os mais adequados a cada finalidade de pesquisa.

Os Capítulos 5 e 6 abordam a amostragem de *survey*: o Capítulo 5 trata da lógica da amostragem e o Capítulo 6 apresenta exemplos para esclarecer a implementação daquela lógica na prática.

Depois de discutir amostragem, daremos atenção ao tema da *medição* — começando no Capítulo 7 com a conceituação e o desenho dos instrumentos. Abordaremos o processo através do qual pesquisadores de *survey* refinam seus conceitos e elaboram itens de questionário para medi-los. O Capítulo 8 continua este assunto, abordando a construção de índices e escalas compostos durante a análise de dados. O Capítulo 8 pode parecer cronologicamente fora de ordem, mas coloquei-o ali porque ele completa o fluxo lógico estabelecido no Capítulo 7.

Capítulo 4



Tipos de Desenhos de Pesquisas

“Pesquisa de *survey*” se refere a um tipo particular de pesquisa social empírica, mas há muitos tipos de *survey*. O termo pode incluir censos demográficos, pesquisas de opinião pública, pesquisas de mercado sobre preferências do consumidor, estudos acadêmicos sobre preconceito, estudos epidemiológicos etc. *Surveys* podem diferir em termos de objetivos, custos, tempo e escopo. Vários desenhos básicos podem ser englobados no termo *survey*.

Este capítulo começa com uma breve discussão dos possíveis objetivos da pesquisa de *survey*. Examina em seguida o conceito de *unidade de análise*. Finalmente, descortina uma visão geral das diferentes estratégias disponíveis para atingir suas finalidades.

Finalidades da Pesquisa de *Survey*


Há provavelmente tantas razões diferentes para se fazer *surveys* quanto há *surveys*. Um político pode encomendar um visando a sua eleição. Uma empresa de *marketing* pode fazer um *survey* visando vender mais sabonetes da marca X. Um governo pode fazer um *survey* para projetar um sistema de trânsito de massa ou para modificar um programa de bem-estar social.

A variedade de tais propósitos seria longa demais para enumerá-las todas aqui, porém três objetivos gerais permeiam todos esses interesses: descrição, explicação e exploração. Um *survey* pode visar (e usualmente visa) atingir mais de um destes objetivos, mas é útil examiná-los separadamente.

Descrição

Surveys são freqüentemente realizados para permitir enunciados descritivos sobre alguma população, isto é, descobrir a distribuição de certos traços e atributos. Nestes, o pesquisador não se preocupa com o porquê da distribuição observada existir, mas com o que ela é.

Distribuições por idade e por sexo relatadas pelo U. S. Bureau of Census são exemplos deste tipo de *survey*. De forma semelhante, o Ministério do Trabalho pode procurar descrever a extensão do desemprego no país num ou em vários pontos do tempo. Também o Gallup pode tentar descrever os percentuais do eleitorado que votarão nos vários candidatos a presidente, a distribuição das atitudes com relação à Guerra do Golfo, a distribuição de atitudes em relação à educação sexual nas escolas ou o uso de flúor na água. O percentual de uma população que provavelmente comprará um novo produto comercial é outro exemplo de pesquisa descritiva.

 O *survey* amostral é um veículo para descobrir estas distribuições. A distribuição de traços numa amostra cuidadosamente selecionada de uma população maior pode ser medida e uma descrição comparável da população maior pode ser inferida a partir da amostra. (O Capítulo 5 discute a lógica destas inferências.)

Além de descrever a amostra total (e inferir para a população total), pesquisadores de *survey* muitas vezes descrevem subamostras e as comparam. Assim, o Gallup pode começar com um relato das intenções de votos de todo o eleitorado e depois descrever separadamente Republicanos e Democratas, homens e mulheres, eleitores de diferentes idades etc. As descrições de diversos subconjuntos podem ser comparadas, mas a finalidade primordial é descrever e não explicar as diferenças. (Tais comparações constituem, contudo, um passo intermediário entre a lógica da descrição e a da explicação, como discutiremos mais detalhadamente no Capítulo 13.)

Explicação

Apesar da maioria dos *surveys* visar, pelo menos em parte, à descrição, muitos têm o objetivo adicional de fazer asserções *explicativas* sobre a população. Ao estudar preferências eleitorais, por exemplo, você pode querer explicar por que alguns eleitores preferem um candidato, enquanto outros optam por outro candidato. Ao estudar desemprego, você pode querer explicar por que parte da força de trabalho está empregada e o restante não.

Explicar quase sempre requer análise *multivariada* — o exame simultâneo de duas ou mais variáveis. Preferências por diferentes candidatos políticos podem ser explicadas em termos de variáveis como filiação partidária, educação, raça, sexo, região do país etc. Ao examinar as relações entre preferências por candidatos e as diversas variáveis explicativas, o pesquisador pode tentar “explicar” por que eleitores escolheram um ou outro candidato. (A Parte 4 deste livro aborda longamente a lógica da explicação.)

Exploração

Métodos de *survey* podem também fornecer um “mecanismo de busca” quando você está começando a investigação de algum tema. Por exemplo, durante o Movimento pelo Discurso Livre, na Universidade da Califórnia em Berkeley, nos anos 60, fiz parte de um grupo de pesquisadores no campus que planejou estudar exaustivamente a natureza, as fontes e as conseqüências do radicalismo estudantil. Embora tivéssemos muitas idéias sobre o assunto, receávamos haver desconsiderado alguns componentes adicionais da situação. Se um grande estudo se baseasse somente em nossas preconcepções, corríamos o risco de não pegar alguns elementos críticos.

O método da pesquisa de *survey* oferece uma técnica para resolver esta dificuldade. Um questionário foi construído e cerca de cinquenta estudantes de diferentes orientações políticas foram entrevistados em profundidade. Não se fez nenhuma tentativa de selecionar uma amostra representativa de alunos, nem os dados foram coletados de forma padronizada. Os entrevistados foram encorajados a falar livremente sobre seus pontos de vista políticos e suas atitudes com relação ao radicalismo estudantil.

As entrevistas resultaram num desenho de pesquisa largamente revisto para o estudo principal. Os entrevistados mencionaram fatores relevantes ao radicalismo estudantil que não havíamos antecipado. Certas orientações políticas que antes pareciam contraditórias agora faziam mais sentido. Estes fatores adicionais foram subseqüentemente levados em consideração no desenho da pesquisa principal.

Mas há coisas que o estudo exploratório não fez. Primeiro, não respondeu às questões básicas que levaram ao planejamento da pesquisa. Também não elucidou adequadamente a natureza do radicalismo estudantil nem abordou questões sobre as fontes e conseqüências desse radicalismo. O modo como os dados foram colhidos claramente impedia estes resultados. Mas o

estudo exploratório suscitou novas possibilidades, mais tarde exploradas no *survey* mais controlado.

São estes os três objetivos básicos da pesquisa de *survey*. Como comentamos, a maioria dos estudos tem mais de um objetivo (às vezes todos os três), mas estes objetivos provêm princípios organizacionais úteis no desenho de *surveys*. Antes de abordar os tipos específicos de desenhos de estudo possíveis para alcançar estes objetivos, vamos introduzir um termo básico, mas às vezes confuso: *unidades de análise*.

Unidades de Análise

Pesquisa de *survey* fornece técnicas para se estudar quase todo mundo. Esse *mundo* estudado num *survey* são as unidades de análise. Tipicamente, a unidade de análise num *survey* é uma pessoa, mas não necessariamente e, aliás, muitas vezes não é. Quaisquer que sejam as unidades de análise, dados são colhidos para se descrever cada unidade individual (por exemplo, uma pessoa). As muitas descrições, então, são agregadas e manipuladas para descrever a amostra estudada e, por extensão, a população representada pela amostra.

Numa pesquisa de mercado sobre as preferências de consumo das marcas X e Y, cada consumidor amostrado e estudado é a unidade de análise. Cada um é descrito em termos da marca que prefere. As várias preferências são, então, agregadas para descrever a população de consumidores em termos de percentagens de predileção pelas duas marcas. Num *survey* sobre desemprego, quem integra a força de trabalho é a unidade de análise. Cada um é descrito como empregado ou desempregado. Estas descrições individuais são então usadas para descrever toda a força de trabalho em termos de taxa de desemprego.

As unidades de análise são tipicamente pessoas, mas podem também ser famílias, cidades, estados, nações, companhias, indústrias, clubes, agências governamentais etc. Em cada caso, as unidades individuais de análise são descritas e estas descrições, agregadas para descrever a população representada pelas unidades. Por exemplo, você pode colher dados descrevendo cidades de um país. Cada cidade pode ser descrita em termos do tamanho da população e todas as cidades do país podem, então, ser descritas em termos da população média. Variáveis adicionais descrevendo as cidades podem ser introduzidas para explicar por que algumas cidades são maiores do que outras.

Evidentemente, um determinado *survey* pode envolver mais de uma unidade de análise. Um *survey* de residências numa certa cidade pode tentar prover a seguinte informação: percentual de estruturas residenciais em condições deterioradas, distribuição racial dos chefes de família, renda familiar média anual, taxa de desemprego e distribuição etária-sexual da população residente. Nestes exemplos, as unidades de análises seriam, respectivamente, estruturas residenciais, residências, famílias, membros da força de trabalho e residentes.

Unidades de análise num *survey* podem ser descritas com base em seus componentes. Por exemplo, cidades podem ser descritas quanto às suas taxas de desemprego ou composições raciais. *Surveys* podem ser feitos para viabilizar estas descrições. Se, porém, o objeto da pesquisa for descrever cidades e agregar as várias descrições para descrever todas as cidades, então a unidade básica de análise para o estudo é a cidade. Ao mesmo tempo, unidades de análise podem ser descritas em termos dos grupos a que pertencem. Assim, indivíduos podem ser descritos em termos do número de pessoas nas famílias ou da condição das suas estruturas residenciais.

A aplicabilidade dos métodos de *survey* a diversas unidades de análise por vezes confunde o pesquisador principiante e resulta na seleção de uma unidade de análise inapropriada para uma linha de investigação particular. Os perigos da falácia ecológica devem ser conscientemente evitados.¹

Suponha que você pretende explorar a possível relação entre raça e crime: são os negros ou os brancos que mais provavelmente terão comportamento criminoso? A unidade apropriada de análise para esta investigação é o indivíduo. Amostras de respondentes negros e brancos serão estudadas e suas respectivas taxas de criminalidade computadas e comparadas.

Entretanto, considerando a disponibilidade de certos dados municipais, você pode ficar tentado a abordar o problema de outra forma. É possível obter facilmente taxas gerais de criminalidade nas cidades mais importantes do país, assim como achar dados sobre a composição racial destas cidades. Analisando estes dados, você pode descobrir que as taxas de criminalidade são mais altas em cidades com maior proporção de população negra. Daí, você pode concluir que negros têm taxas de criminalidade mais altas do que brancos. No entanto, esta linha de investigação está sujeita à falácia ecológica, porque não há garantia de que crimes cometidos em cidades predominantemente negras foram cometidos por negros. É concebível que as taxas mais altas de criminalidade podem

ocorrer entre brancos residindo em áreas predominantemente negras. Tal má interpretação não seria possível se você tivesse utilizado a unidade de análise correta.²

Qualquer que seja a natureza dos dados usados para descrever as unidades de análise, é importante que eles sejam previamente identificados. Do contrário, o desenho da amostra e os métodos de coleta de dados podem impedir a análise mais apropriada ao estudo. Se houver mais de uma unidade de análise no estudo, esta consideração se tornará ainda mais importante. Um exemplo final deve esclarecer.

Suponha que, num estudo da população de residências numa certa cidade, você decida colher dados através de entrevistas com os chefes de uma amostra de famílias, mas queira fazer afirmações sobre estruturas residenciais, residências, famílias e pessoas. Dados sobre as estruturas podem ser coletados perguntando questões relevantes aos entrevistados (qual é a idade da sua casa?) e por meio de observações (quando andares tem a casa?). Dados sobre a residência também podem ser coletados entrevistando (quantas famílias moram nesta casa?) e observando (qual é a raça do chefe da família?). Pode-se pedir dados ao respondente para descrever as famílias residentes (qual é a renda anual?) e os indivíduos residentes (qual o sexo e a idade de cada pessoa?).

Num estudo complexo como este, com várias unidades de análise, deve-se ter cuidado especial na organização dos dados para análise. Se você quiser determinar o percentual de estruturas residenciais em estado deteriorado, cada estrutura deve ser descrita apenas uma vez na computação, independentemente do número de famílias ou pessoas morando lá. Se você quiser determinar o percentual da população residindo em estruturas deterioradas, cada estrutura deve ser lançada na computação tantas vezes quantas forem as pessoas morando nela. Por exemplo, se cinco pessoas residem numa estrutura deteriorada, cada pessoa deve ser descrita como morando numa estrutura deteriorada e todas as pessoas descritas seriam agregadas para descrever a população da cidade.

O caminho mais seguro a seguir num *survey* complexo como este é criar arquivos separados de dados para cada unidade de análise. Neste exemplo, você deve criar um arquivo para estruturas, um para residências, um para famílias e outro para pessoas. Cada arquivo conteria todos os dados relevantes para análise daquela unidade, mesmo que inicialmente os dados possam ter sido colhidos de outra unidade de análise.

Assim, cada arquivo de pessoas pode ter indicação da raça do chefe da família à qual a pessoa pertence. Uma vez determinada a unidade de análise para uma determinada computação, você pode fazê-la facilmente usando o arquivo de dados correspondente.

Desenhos Básicos de *Surveys*

Após especificar os objetivos e as unidades de análise do seu *survey*, você pode escolher entre diversos desenhos diferentes. Nesta seção, discutiremos *surveys interseccionais* [*cross-sectional*] e *surveys longitudinais* e o uso de *surveys* interseccionais como aproximação de *surveys* longitudinais. Na seção seguinte, discutiremos variações destes desenhos básicos.

Surveys Interseccionais

Num *survey* interseccional, dados são colhidos, num certo momento, de uma amostra selecionada para descrever alguma população maior na mesma ocasião. Tal *survey* pode ser usado não só para descrever, mas também para determinar relações entre variáveis na época do estudo.

Uma pesquisa para determinar intenções de votos é um exemplo de pesquisa interseccional. Vale a pena notar que, tipicamente, se indaga aos entrevistados: "Se a eleição fosse hoje, em quem você votaria?" Os resultados são apropriadamente relatados assim: "Se a eleição fosse hoje, o candidato X venceria disparado." Um *survey* para determinar taxa de desemprego descreveria o desemprego da população na época do estudo.

Do mesmo modo, um *survey* interseccional da relação entre religiosidade e preconceito relataria esta relação na época do estudo. Você pode relatar que pessoas religiosas são mais preconceituosas do que pessoas irreligiosas, mas reconhecendo que esta relação pode mudar posteriormente. Um *survey* subsequente poderia encontrar uma relação bem diferente. (Em geral, porém, tais relações tendem a persistir mais tempo do que as descrições.)

Surveys Longitudinais

Alguns desenhos de *survey* — descritivos ou explicativos — permitem análise de dados ao longo do tempo. Dados são coletados em tempos diferentes, e relatam-se mudanças de

descrições e de explicações. Os principais desenhos longitudinais são *estudos de tendências*, *estudos de cortes* e *estudos de painel*.

Estudos de Tendências. Uma população pode ser amostrada e estudada em ocasiões diferentes. Ainda que pessoas diferentes sejam estudadas em cada *survey*, cada amostra representa a mesma população. As pesquisas eleitorais no decorrer de uma campanha política são um bom exemplo de estudos de tendências. Diversas vezes, durante a campanha, amostras de eleitores são selecionados e perguntados em quem votarão. Comparando os resultados das várias pesquisas, os pesquisadores podem determinar mudanças nas intenções de voto.

No estudo constante do preconceito nos EUA, freqüentemente os entrevistados são indagados se acham que crianças negras e brancas deveriam estudar na mesma escola. Com o passar dos anos, os percentuais a favor de escolas racialmente integradas vêm aumentando consistentemente. Tais dados permitem que pesquisadores observem tendências nas atitudes com relação à integração racial.

Deve ser ressaltado que estudos de tendências muitas vezes envolvem longos períodos de coleta de dados. Tipicamente, você não colhe pessoalmente todos os dados usados num estudo de tendências, mas, ao invés, faz uma análise secundária de dados coletados ao longo do tempo por vários outros pesquisadores.

Até agora comentamos apenas estudos descritivos de tendências, mas não há motivos para um pesquisador não examinar tendências nas relações entre variáveis. Um exemplo é a relação entre filiação religiosa e preferência política. Tradicionalmente, católicos e judeus nos EUA tendem a votar no Partido Democrata mais do que os protestantes, mas esta relação pode ser examinada ao longo do tempo.

Estudos de Coortes. Estudos de tendências se baseiam em descrições de uma população *geral* (como eleitores americanos) ao longo do tempo, embora mudem os integrantes daquela população. Pessoas vivas e representadas no primeiro estudo podem ter morrido por ocasião do segundo, e pessoas ainda não nascidas, quando do primeiro estudo, podem estar representadas no segundo. Assim, um estudo de tendências de atitudes entre estudantes numa universidade reflete uma população diferente de estudantes cada vez que um *survey* é feito.

Um estudo de coortes focaliza a mesma população específica cada vez que os dados são coletados, embora as amostras estudadas possam ser diferentes. Por exemplo, podemos selecionar uma amostra de estudantes se formando na universidade

em 1990, para medir suas atitudes em relação ao trabalho. Cinco anos depois, selecionamos e estudamos outra amostra da mesma turma. A amostra é diferente a cada vez, mas continuamos descrevendo a turma de 1990. (Se estudássemos os formandos de 1995 na segunda vez, teríamos um estudo de tendências de turmas de formandos e não um estudo de coorte da turma de 1990.)

O exemplo seguinte ilustra um tipo diferente de estudo de coorte. Num momento, você pode selecionar uma amostra de todos os americanos entre vinte e trinta anos de idade. Dez anos mais tarde, a amostra será de pessoas dos trinta aos quarenta, e assim por diante. Isto é um estudo de coorte de um determinado grupo de idade. O mesmo estudo pode ser feito através de análise secundária de dados previamente colhidos. Numa ocasião, você pode analisar os entrevistados com vinte anos no estudo de 1970, os de trinta num estudo em 1980, os com quarenta num estudo em 1990 e assim por diante.

Estudos de Painel. Tanto os estudos de tendências como os de coorte permitem análise de processo e mudança no tempo, o que é difícil num *survey* interseccional. Mas esses métodos têm severas desvantagens. Um estudo de tendências pode determinar que os eleitores, como um grupo, estão mudando do candidato A para o B, mas não indica *quais* pessoas estão mudando, atrapalhando tentativas de explicar por que está ocorrendo esta mudança.

Estudos de painéis envolvem a coleta de dados, ao longo do tempo, da mesma *amostra* de respondentes, que se chama *painel*. Num estudo político, você pode reentrevistar todos os membros do seu painel em intervalos mensais durante a campanha eleitoral. Cada vez você pergunta em quem votariam; se ocorrer mudança, você saberá quais pessoas mudaram em qual direção. Analisando outras características de quem muda e de quem não muda, talvez você possa explicar as razões da mudança. (Claro que você também perguntaria aos membros do painel por que haviam mudado as intenções de voto.)

Exceto em certos casos limite, estudos de painel têm de ser realizados como parte de um programa de pesquisas. Enquanto estudos de tendências e de coortes podem ser feitos através de análise secundária de dados coletados previamente, estudos de painel não podem. Conseqüentemente, tendem a ser caros e demorados, além de sofrerem de dois outros problemas.

Esgotamento de Painel, o primeiro problema, se refere ao montante de não respostas que ocorrem nas últimas ondas de entrevistas. Algumas pessoas entrevistadas no primeiro

survey podem não querer ou não poder participar posteriormente. O ponto forte dos estudos de painel é a capacidade de examinar os mesmos respondentes em diferentes ocasiões; esta vantagem se perde quando respondentes não participam de diversos *surveys*.

Segundo, a análise dos dados do painel pode ser bem complicada. O principal mecanismo analítico é a *tabela de mudanças*, que faz tabulações cruzadas de alguma característica em mais de uma ocasião. Por exemplo, respondentes que preferem o candidato A no primeiro *survey* são divididos entre os que continuam a preferi-lo no segundo *survey* e os que agora preferem o candidato B. Os eleitores que, da primeira vez, preferiram o candidato B são divididos de forma semelhante. Aumentando o número de *surveys*, de variáveis e a complexidade das variáveis, a análise e a apresentação dos dados podem se tornar inadministráveis (ver o Capítulo 14 sobre as dificuldades de análises tipológicas).

Por todas estas razões, estudos de painel são realizados com menos freqüência do que outros tipos de *survey*. Mas o estudo de painel é o desenho de *survey* mais sofisticado para a maioria dos propósitos explicativos. (É o que mais se aproxima do clássico experimento em laboratório.)

Aproximação de *Surveys* Longitudinais

Sem dúvida, o *survey* interseccional é o desenho de pesquisa usado mais freqüentemente, embora muitas, senão a maioria, das questões que o pesquisador quer responder envolvam alguma noção de mudança no tempo. Alguns mecanismos podem ser empregados num *survey* interseccional para aproximar o estudo de processo ou mudança.

Primeiro, os respondentes devem fornecer dados relevantes para questões que envolvem processo. Por exemplo, pode-se perguntar a eles a renda da família no ano atual e no anterior, e estes dados usados como se tivessem sido colhidos num estudo de painel com duas ondas de entrevistas, com um ano de intervalo. Mas há dois perigos aqui. Primeiro, que respondentes possam não conseguir informar com exatidão. Quanto mais longe tiverem de ir na memória, menos exata, provavelmente, será a informação prestada. Segundo, os pesquisadores não devem se enganar e interpretar os dados do ano anterior como interseção da população daquele ano, visto que a amostra é limitada à população atual.

Uma segunda forma de aproximar-se de um estudo no tempo é fazer comparações de idade ou de coorte num *survey* interseccional. Numa pesquisa, jovens podem ser menos religiosos que os velhos e isso ser interpretado como um declínio na religiosidade da população. (Por outro lado, o fato de que pessoas tendem a se tornar mais religiosas com a idade pode dar conta completamente das diferenças observadas.) Outro estudo pode descobrir que calouros são intelectualmente menos sofisticados que alunos do quarto ano e concluir que a educação universitária aumenta a sofisticação. (Reconhece, contudo, que a população da qual os novos calouros são tirados pode ser menos sofisticada, apontando para uma tendência e não para um processo. Pode ser também que alunos menos sofisticados que foram calouros há quatro anos atrás saíram da universidade antes da realização do *survey*.)

Finalmente, dados interseccionais podem às vezes ser interpretados logicamente, para indicar um processo temporal. Um estudo do uso de drogas entre estudantes, por exemplo, indicou que todos os que declararam uso de maconha também informaram experiência anterior com álcool. Além disto, todos os estudantes que relataram terem usado LSD também relataram o uso de maconha (e álcool). É razoável concluir, desses dados, que o progresso no uso de drogas é do álcool para a maconha para o LSD. Se, por exemplo, alguns alunos tivessem usado maconha antes do álcool, então o *survey* interseccional devia ter descoberto alguns que haviam usado maconha, mas não álcool. Da mesma forma, deviam ter sido descobertos estudantes que tomaram LSD mas não outras drogas, se o LSD tivesse sido usado antes das outras. Como nenhum estudante assim foi encontrado, enquanto foram encontrados muitos que haviam tomado apenas álcool ou todas as drogas estudadas menos LSD, os pesquisadores concluíram, por motivos lógicos, que o processo temporal passava de álcool para maconha para LSD. (*Nota:* esta conclusão não apóia a suposição de que uma droga leva fisiologicamente à outra.)

Variações dos Desenhos Básicos

A seção anterior delineou os desenhos básicos de *survey*; todo *survey* pode ser caracterizado nos termos discutidos. Ao mesmo tempo, estes desenhos básicos podem ser modificados de vários modos para se enquadrarem no que requer

uma determinada pesquisa. Esta seção aborda algumas das modificações mais comuns.

Amostras Paralelas

Um problema de pesquisa pode às vezes ser particularmente relevante para mais de uma população. Por exemplo, se você for um pesquisador educacional, pode querer amostrar as atitudes dos estudantes em relação à proposta de um código de conduta estudantil. Talvez você também se interesse em saber o que os professores e até os administradores escolares acham do código. Neste caso, você pode amostrar cada população separadamente e administrar o mesmo questionário (ou outro ligeiramente modificado) a cada amostra. Os resultados seriam então comparados. Num outro exemplo, você pode querer examinar as crenças religiosas de fiéis da Igreja Metodista e compará-las com as crenças do clero metodista. Novamente, cada população seria amostrada e estudada. Tais estudos são denominados *amostras paralelas*.

Em alguns casos, a amostra de uma população pode ser usada para gerar a amostra da outra. Por exemplo, estudantes universitários podem ser amostrados, e questionários, enviados aos estudantes e seus pais. As respostas dos alunos como um todo podem ser comparadas às respostas dos pais como um todo.

Estudos Contextuais

Como observamos ao discutir unidades de análise, pessoas podem ser descritas em termos dos grupos a que pertencem. Assim como se pode descrever uma família como grande, membros desta família podem ser descritos como pertencendo a uma família grande. Colher dados sobre partes do ambiente ou meio da pessoa e usar tais dados para descrever o indivíduo constitui um *estudo contextual* — um exame do contexto do indivíduo.

Usando o exemplo das amostras paralelas, dados colhidos dos pais podem ser usados para descrever seus filhos. Um aluno pode ser descrito como filho de pai politicamente liberal, de mãe idosa etc., e estes dados serem então usados na análise das atitudes do estudante.

Ao estudar membros de igrejas, você pode coletar dados sobre a igreja a que pertence cada um e, talvez, sobre o ministro da mesma. Um membro da igreja pode ser então descrito

como pertencendo a uma igreja rica e grande no centro da cidade, com um ministro abaixo dos quarenta anos.

Entretanto, tais análises exigem que os dados colhidos sobre os contextos dos respondentes sejam identificados com eles, de forma a serem incluídos nos seus arquivos de dados. Isto não pode ser feito enviando questionários anônimos à amostra de estudantes e pais. Os estudantes têm de ser identificados no questionário (por nome ou número) e os pais serem identificados de alguma forma que os ligasse aos filhos.

Estudos Sociométricos

Tipicamente, *surveys* estudam uma amostra de uma determinada população, coletando dados sobre os indivíduos na amostra, para descrever e explicar a população que representam. Todavia, o formato básico do *survey* pode ser usado para um exame mais abrangente de algum grupo, observando as inter-relações entre seus membros. Um bom exemplo é o *estudo sociométrico*.

Suponha que você queira aprender algo sobre a seleção de amigos íntimos entre crianças na escola. Você pode fazer um *survey* convencional entre uma amostra de alunos e pedir-lhes várias informações sobre os amigos mais próximos. Num desenho sociométrico, você estudaria *todos* os alunos numa sala e pediria a cada um que identificasse os amigos pelo nome. Desta forma, você pode descobrir que Jack escolheu Bill como melhor amigo, mas Bill, por sua vez, escolheu Frank. Você pode também descobrir que cinco alunos da sala escolheram Mary como sua melhor amiga e ninguém escolheu Ruth. Estas espécies de análises podem crescer em complexidade, para possibilitar um exame abrangente de toda uma rede de amizades. Procurando explicar por que certos estudantes foram mais escolhidos do que outros, você teria um corpo de dados sobre os estudantes escolhidos, a partir dos auto-relatos de características, incluindo quem eles escolheram.

Deste modo, você pode examinar vários possíveis fatores que governam a formação de amizades, incluindo variáveis como sexo, raça, status econômico e inteligência. A coleta de dados de todos os membros do grupo poupa ao pesquisador ter que pedir aos participantes descrever detalhadamente suas seleções; além disso, você teria acesso à rede de amizades. Variantes desta técnica podem ser usadas com qualquer grupo e a pesquisa pode ser conduzida ao longo do tempo ou de forma interseccional.

Escolhendo o Desenho Adequado

Dadas as várias opções disponíveis ao pesquisador de *surveys*, a questão é "Qual desenho devo escolher?" Esta pergunta não pode ser respondida em abstrato, porque diferentes problemas de pesquisa requerem desenhos diferentes. Mesmo assim, podem-se dar algumas diretrizes gerais.

Primeiro, se seu objetivo for uma descrição de tempo único, um *survey* interseccional é provavelmente o desenho mais adequado. Você identifica a população relevante, seleciona uma amostra de respondentes e faz o *survey*. A mesma opção se aplica a uma pesquisa na qual interessa descrever subconjuntos. Caso você pretenda documentar diferenças de atitudes políticas entre homens e mulheres, pode consegui-lo com um *survey* interseccional.

Mais comumente, porém, você se interessa em examinar algum tipo de processo dinâmico, com mudanças ao longo do tempo. Por exemplo, ao lidar com as fontes e/ou consequências da religiosidade, a questão da mudança no tempo fica implícita, se não explícita. Implicitamente se supõe que algumas pessoas se tornam religiosas e que ser religioso tem efeitos subsequentes em atitudes e/ou comportamentos. O ideal é selecionar uma amostra de respondentes num ponto da vida antes de desenvolverem uma orientação religiosa e estudá-los ao longo do tempo, cobrindo o período durante o qual alguns deles se tornam religiosos, acompanhando-os durante o período em que a religiosidade tem efeitos sobre outros aspectos de suas vidas.

Um estudo assim leva anos para se completar. Provavelmente, você começa com uma amostra de pré-adolescentes e os segue através da meia-idade até se tornarem idosos. O tempo e custo exigidos por um estudo assim provavelmente o inviabilizariam. Por isto, tais temas são mais freqüentemente trabalhados em *surveys* interseccionais. Em vez de observar os efeitos de várias condições e experiências sociais enquanto ocorrem, você examina seus possíveis efeitos comparando respondentes que passaram por estas condições e experiências no passado com aqueles que não passaram.

Por exemplo, pessoas casadas geralmente são menos religiosas que as não casadas. O estudo de painel permite observar o declínio de religiosidade em alguns participantes após o casamento. Mas o *survey* interseccional permite comparar os níveis de religiosidade entre casados e não casados

num dado momento, observando menos religiosidade em geral nos respondentes casados. Embora você não possa observar o efeito do casamento na religiosidade na época do casamento, pode querer inferir este efeito com base na diferença neste tempo único.³ (A Parte 4 aborda a lógica destas inferências.)

Surveys de painel são mais viáveis quando o fenômeno estudado tem duração relativamente curta. Uma campanha eleitoral é um exemplo. Como uma campanha deve durar menos de um ano, é possível fazer diversas ondas de entrevistas com o painel do *survey* durante a campanha, monitorando mudanças nas intenções de votos neste período e coletando dados relevantes à explicação destas mudanças. Com tão curta duração, devem ser reduzidas as dificuldades de esgotamento do painel, e fica mais fácil localizar os respondentes. (Num período mais longo, muitos respondentes podem se mudar da cidade e fica mais difícil localizá-los.)

Resumindo, sempre que o problema de pesquisa envolver o exame de mudanças individuais no tempo, o *survey* de painel seria teoricamente o desenho mais adequado. Se o processo de mudança ocorrer durante um tempo relativamente curto, o *survey* de painel pode ser viável. Mais comumente, porém, o pesquisador é obrigado a depender de dados interseccionais para inferir o processo de mudança nos indivíduos com o passar do tempo.

Se, por outro lado, o problema de pesquisa tratar pouco de amplas tendências temporais — do ponto de vista descritivo —, as dificuldades serão menores na maioria dos casos. Em muitas situações, você verá que outros pesquisadores já coletaram e relataram todos os dados que você precisa. Por exemplo, se tiver interesse nas mudanças gerais de atitudes com relação à Guerra da Indochina, você vai descobrir que pesquisadores acadêmicos e comerciais vêm colhendo informações relevantes há anos. Sua tarefa, então, pode ser apenas localizar tais estudos, comparar a natureza dos itens de questionário e os desenhos de amostras, para discutir as mudanças observadas. Em outras situações, você pode verificar que só um *survey* colheu dados relevantes ao seu interesse de pesquisa ou que o último estudo realizado já está defasado. Em tal caso, você pode querer fazer um novo *survey* interseccional — comparável em desenho amostral e itens de questionário — a fim de fornecer uma nova medição visando um exame de tendências.

Resumo

São estes então os desenhos básicos e as variações mais comuns disponíveis ao pesquisador de *survey*. A discussão anterior de modo algum exaure as possibilidades de desenho, mas deve oferecer-lhe estímulo e orientação suficientes para construir o desenho de estudo mais apropriado às suas necessidades de pesquisa.

Os melhores estudos muitas vezes são os que combinam mais de um desenho, já que cada desenho dá uma perspectiva diferente ao assunto estudado. Ao mesmo tempo, você deve se precaver contra desenhar um *survey* tão complexo que não haja tempo nem dinheiro para executá-lo. Já enfatizamos no Capítulo 3 que todos os desenhos de *survey* representam compromissos. O bom pesquisador consegue o melhor compromisso possível.

Notas

¹ Para uma discussão mais abrangente deste tópico, vide ROBINSON, W. S. Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, p.351-357, jun. 1950.

² Deve-se reconhecer que o pesquisador às vezes não tem condições de conduzir uma investigação usando as unidades de análise mais apropriadas. Nestes casos, uma "análise ecológica" pode ser a única abordagem viável para o assunto no momento. O famoso estudo de Durkheim sobre suicídio é um excelente exemplo. Para uma discussão de como Durkheim evitou a falácia ecológica, vide SELVIN, Hanan. Durkheim's Suicide and Problems of Empirical Research. *American Journal of Sociology*, p.607-619, maio 1958. Outro exemplo é STOUFFER, Samuel A. Effects of the Depression on the Family. In: STOUFFER, Samuel A. (Ed.). *Social Research to Test Ideas*. New York: Free Press, 1962. p.134-153.

³ Reconhece-se, evidentemente, que esta inferência a partir de dados interseccionais está sempre sujeita a questionamento. Neste caso, pode-se argumentar que pessoas religiosas têm menor probabilidade de se casarem, ou seja, que a *direção causal* é oposta à que foi inferida.

Leituras Adicionais

COOK, Thomas D., CAMPBELL, Donald T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally, 1963.

GLOCK, Charles Y. *Survey Design and Analysis in Sociology*. In: GLOCK, Charles Y. (Ed.). *Survey Research in the Social Sciences*. New York: Russell Sage Foundation, 1967. p.1-62.

HUNT, Morton. *Profiles of Social Research: The Scientific Study of Human Interactions*. New York: Basic Books, 1965.

LAZARFELD, Paul F., PASANELLA, Ann K., ROSENBERG, Morris (Ed.). *Continuities in the Language of Social Research*. New York: Free Press, 1972. seções III-IV.

MILLER, Delbert. *Handbook of Research Design and Social Measurement*. New York: Longman, 1983.

Capítulo 5

A Lógica da Amostragem do *Survey*

Neste livro, como em muitos outros, o termo *survey* tem sido usado com o sentido implícito de “*survey* por amostragem”, por oposição ao estudo de todos os componentes de uma população ou grupo. Tipicamente, métodos de *surveys* são usados para estudar um segmento ou parcela — uma amostra — de uma população, para fazer estimativas sobre a natureza da população total da qual a amostra foi selecionada. Embora a prática de usar amostras seja aceita mais ou menos tacitamente, as razões para amostrar, em geral, não são conhecidas. Consideremos brevemente estas razões, antes de nos voltarmos para a lógica e as técnicas da amostragem.

Por que Fazer Amostragem?

Há duas razões para justificar a amostragem: tempo e custo. Uma entrevista de um *survey* domiciliar abrangente pode exigir de uma a três horas e de \$40 a \$100. A economia de estudar 2.000 pessoas em vez, digamos, 500.000 é óbvia. Portanto, amostrar muitas vezes torna possível um projeto, enquanto a recusa em amostrar acabaria com a pesquisa.

No entanto, a amostragem não deve ser vista como um mal necessário. Em geral não se reconhece que *surveys* por amostragem são muitas vezes mais precisos do que entrevistar todos os componentes de uma população. Este fato aparentemente bizarro resulta de diversas características da logística do *survey* por entrevista.

Primeiro, um grande projeto de entrevistas requer um número enorme de entrevistadores. Pesquisadores caracteristicamente tentam limitar seu pessoal aos melhores entrevistadores disponíveis, e um grande projeto provavelmente exige contratar qualquer pessoa disponível, resultando que a qualidade geral dos entrevistadores é mais baixa do que a habitual. A qualidade dos dados coletados se reduz pela menor qualidade dos entrevistadores. Além disso, uma pesquisa em menor escala permite procedimentos de acompanhamento mais severos, aumentando a taxa de entrevistas completas.

Segundo, já que entrevistar todos os membros de uma população grande requer uma fase de entrevistas longa, fica difícil e mesmo impossível especificar a *época* à qual se referem os dados. Se a pesquisa visar medir o nível de desemprego numa grande cidade, o índice de desemprego fornecido pelos dados do *survey* não se referem à cidade, seja do início, seja do fim das entrevistas. Em vez disso, tem que dar alguma data hipotética ao índice de desemprego, talvez o ponto médio do período de entrevistas. (Pedir aos entrevistados para responder sobre uma data especificada uniforme introduz o problema da recordação imprecisa.) O problema da atribuição do tempo é inerente a qualquer projeto de entrevistas não executado num só momento, e a gravidade do problema cresce com a duração do período de entrevistas. Se entrevistar tomar dez meses — com os índices de desemprego variando neste período —, o índice resultante não terá significado.

Finalmente, os requisitos administrativos de um grande *survey* são muito maiores do que os em geral encontradas pelos pesquisadores por *surveys*. A supervisão, os relatórios, o treinamento e assim por diante são muito mais difíceis num grande *survey*, e a qualidade dos dados coletados pode ser menor do que a obtida num trabalho menor e mais manejável. (Vale a pena observar que o Bureau do Censo nos EUA produz um *survey* por amostragem após o censo decenal para avaliar os dados colhidos na enumeração total.)

Dados por Amostragem São Realmente Exatos?

Apesar da discussão anterior, talvez você ainda sinta algum desconforto quanto à amostragem. Uma vez que é possível uma amostra representar erroneamente a população da qual é obtida, o pesquisador que usa métodos de amostragem enfrenta um perigo inevitável. No entanto, como mostraremos neste capítulo, procedimentos de amostragem bem-estabelecidos

podem tornar aceitável esse perigo. *Surveys* por amostragem podem permitir estimativas muito precisas sobre as populações que retratam. Mas o pesquisador de *surveys* por amostragem deve estar preparado para tolerar uma certa ambigüidade, uma vez que raramente é possível determinar o grau de precisão dos achados de uma amostra.

Pesquisadores de opinião política representam um grupo de pesquisadores por *surveys* a quem é dada a oportunidade de verificar a precisão dos achados das suas amostras. O dia da eleição é o do julgamento final para eles, e suas experiências são instrutivas em relação à questão mais geral da precisão dos *surveys* por amostragem.

A maioria dos críticos dos métodos de *survey* por amostragem lembra a votação do *Literary Digest* de 1936, que previu a vitória de Alfred M. Landon sobre Franklin D. Roosevelt por maioria absoluta. Pesquisando uma amostra de mais de 2 milhões de eleitores pelo correio, o *Digest* previu que Landon derrotaria Roosevelt em quase 15 pontos percentuais. A razão principal deste fracasso foi a *moldura de amostragem* usada pelos pesquisadores (discutiremos este conceito depois). A amostra do *Digest* foi tirada de catálogos telefônicos e listas de registros de automóveis, um procedimento de amostragem que foi suficiente nas eleições de 1920, 1924, 1928 e 1932, mas que em 1936 não forneceu um corte representativo dos eleitores americanos. No rastro da Depressão, e no meio do New Deal, números sem precedentes de americanos pobres apareceram para votar, e estas pessoas pobres não estavam adequadamente representadas pelos catálogos telefônicos e listas de registros de automóveis.¹

Em 1936, George Gallup previu corretamente que Roosevelt ganharia um segundo mandato. Os procedimentos de amostragem de Gallup diferiam dos do *Literary Digest*. O Instituto Americano de Opinião Pública de Gallup foi pioneiro no uso da *amostragem por cotas* (descrita posteriormente neste capítulo), que garantia a todos os tipos de eleitores, ricos e pobres igualmente, adequada representação na amostra. Onde a pesquisa do *Digest* falhou em encontrar e interrogar os eleitores pobres — predominantemente Democratas —, a amostragem por cotas de Gallup o fez.

Doze anos depois, Gallup e a maioria dos outros pesquisadores de opinião política passaram pelo constrangimento de prever a vitória de Thomas Dewey sobre Harry Truman. Como observou Goodman Ace, acidamente: "Todos acreditam nas pesquisas de opinião pública. Todos, desde o homem da rua até o presidente Thomas E. Dewey."² Vários fatores conspiraram para

causar o fiasco da pesquisa de 48. Para começar, a maioria dos pesquisadores terminou sua pesquisa cedo demais, apesar de uma tendência constante em direção a Truman no transcorrer da campanha. Grande número dos eleitores que disseram não saber em quem votariam foram predominantemente para Truman. Mais importante, contudo — o fracasso de 48 apontou uma série de falhas inerentes à amostragem por cotas, o mesmo método que fora um avanço considerável sobre os métodos de amostragem do *Literary Digest*. Em 1948, diversos pesquisadores acadêmicos de *survey* estavam experimentando métodos de *amostragem probabilística*. De forma geral, eles tinham muito mais sucesso do que os amostradores por cotas, e a amostragem probabilística continua sendo o método mais respeitado por pesquisadores de *survey*.

Esta breve discussão conta uma história parcial das fases iniciais do *survey* por amostragem nos EUA, talvez à custa da fé do leitor nos métodos destes *survey*. Para compensar, consideremos o currículo mais recente da precisão da pesquisa de opinião política. Em novembro de 1988, George Bush recebeu 54% dos votos para presidente; Michael Dukakis, 46%. Mas os eleitores não tiveram que esperar os resultados das urnas para saber quem havia vencido. Várias pesquisas de boca de urna, entrevistando amostras de eleitores logo após saírem das seções eleitorais, mostraram as seguintes estimativas dos resultados, apresentados na Tabela 5-1.

TABELA 5-1
Resultados das pesquisas de boca de urna: novembro de 1988

	Bush	Dukakis
CBS News / New York Times	54%	46%
ABC News / Washington Post	54%	46%
NBC News / Wall Street Journal	54%	46%
Los Angeles Times	54%	46%

FONTE - Dados tirados de *Public Opinion*, p.25-26, jan./fev. 1989. Os dados da CBS foram refeitos para excluir as "respostas não".

No lugar dos 2 milhões de eleitores avaliados pelo *Literary Digest* em 1936, estas quatro pesquisas de boca de urna amostraram entre 6.000 e 20.000 eleitores para predizer a votação dos mais de 90 milhões de americanos que foram às urnas. Os resultados destas pesquisas não foram surpreendentes, apenas confirmaram a tendência relatada nas pesquisas feitas na semana anterior à eleição.

O que as grandes pesquisas tinham a dizer, em geral baseadas em amostras de 2.000 ou menos entrevistados, está apresentado na Tabela 5-2.

TABELA 5-2
Pesquisas de opinião política antes da eleição de 1988

	Bush	Dukakis
CBS News / New York Times (4-5/11/88)	55%	45%
Pesquisa Gallup (3-5/11/88)	56%	44%
USA Today / CNN (3-6/11/88)	55%	45%
Pesquisa Harris (2-5/11/88)	53%	47%
ABC News / Washington Post (2-5/11/88)	55%	45%
NBC News / Wall Street Journal (1-5/11/88)	53%	47%

FONTE - Dados tirados de *Public Opinion*, p.39, nov./dez. 1988. Todos os dados foram refeitos para excluir as "respostas não".

Surveys por amostragem podem ser extremamente precisos. Ao mesmo tempo, devemos conceder que, ainda hoje, muitas vezes não o são. O resto do capítulo apresenta as razões e as regras para obter precisão na amostragem.

Dois Tipos de Métodos de Amostragem

É útil distinguir dois grandes tipos de métodos de amostragem: *amostragem probabilística* e *amostragem não probabilística*. O cerne deste capítulo se dedica à amostragem probabilística, por ser hoje o método mais respeitado e útil. Uma parte menor do capítulo considera os vários tipos de métodos de amostragem não probabilística.

Vamos iniciar com uma discussão sobre a lógica da amostragem de probabilidade, seguida de uma breve taxonomia de conceitos e terminologia de amostragem. Depois nos voltaremos para o conceito de distribuição amostral, que é a base para estimar a precisão dos achados do *survey* por amostragem. Estas discussões teóricas serão seguidas por considerações sobre *molduras de populações e de amostragem*, focalizando os problemas práticos de como determinar o grupo alvo do estudo e como começar a selecionar uma amostra. A seguir examinaremos os tipos básicos de desenhos de *surveys*: *amostras aleatórias simples*, *amostras sistemáticas*, *amostras estratificadas* e *amostras por conglomerados [cluster]*. Então

discutiremos e descreveremos brevemente a amostragem não probabilística.

O capítulo termina com breves considerações sobre alguns usos não-*survey* dos métodos de amostragem em áreas como a análise de conteúdo, a observação participante e a análise histórica. Esperamos que você se familiarize o suficiente com a *lógica* do *survey* por amostragem para poder aplicar este conhecimento numa ampla variedade de situações.

Lógica da Amostragem de Probabilidade

A história das pesquisas de opinião políticas mostra que *surveys* por amostragem podem ser bem precisos. Ao mesmo tempo, deve ficar igualmente claro que as amostras devem ser selecionadas cuidadosamente. Consideraremos rapidamente o porquê disso.

As Implicações da Homogeneidade e da Heterogeneidade

Se todos os membros de uma população fossem idênticos uns aos outros em todos os sentidos, não haveria necessidade de procedimentos cuidadosos de amostragem; qualquer amostra seria suficiente. De fato, neste caso extremo de *homogeneidade*, um caso bastaria como amostra para estudar as características de toda a população.

Antes de descartar essa idéia como impossível, lembre-se de que muito da amostragem científica é feita desta forma. Nas ciências físicas, em algumas situações, é seguro fazer tal suposição. O químico que testa certas propriedades do carbono, por exemplo, não precisa passar pela enumeração sofrida de todo o carbono no mundo, para depois selecionar uma amostra probabilística de moléculas de carbono para estudo. Do mesmo modo, o médico ou o clínico que examina o sangue de uma pessoa não precisa retirar todo o sangue dela, para depois selecionar uma amostra probabilística de células sanguíneas. Para fins práticos, qualquer amostra de sangue da pessoa é suficiente.

Mas, se há variação ou heterogeneidade na população estudada, o pesquisador deve usar procedimentos de amostragem mais controlados. Vale a pena assinalar a aplicabilidade maior deste princípio além da pesquisa social. As origens da moderna teoria de amostragem estão na pesquisa agrícola, especialmente no trabalho de R. A. Fisher, cujo nome continua associado a algumas das estatísticas de *survey* mais usadas.

Para nossos fins, é mais importante assinalar a heterogeneidade dos grupos sociais. Como as pessoas diferem de várias formas, cada população humana é composta de indivíduos variados. Uma amostra de indivíduos de uma população deve conter essencialmente a mesma variação existente na população, para permitir descrições úteis dela. A amostragem probabilística é um método eficiente para extrair uma amostra que reflita corretamente a variação existente na população como um todo.

Viés Consciente e Inconsciente de Amostragem

É claro que qualquer um pode selecionar uma amostra de *survey*, mesmo sem treinamento ou cuidado especial. Para selecionar uma amostra de 100 estudantes universitários, uma pessoa pode ir ao campus e entrevistar os alunos que ali encontrar andando. Este tipo de método de amostragem, muitas vezes usado por pesquisadores não treinados, tem problemas muito sérios.

Para começar, há o perigo dos vieses pessoais do pesquisador afetarem a amostra selecionada desta forma; portanto, ela não representaria realmente a população estudantil. Suponha, por exemplo, que você se sente intimidado por estudantes com ar *bippy* e acha que eles ridicularizariam seu trabalho. Portanto, você pode evitar, consciente ou inconscientemente, entrevistar tais pessoas. Por outro lado, você pode sentir que as atitudes de estudantes de aspecto "sério" não seriam relevantes para suas finalidades de pesquisa, evitando entrevistá-los. Mesmo que buscasse entrevistar um grupo "equilibrado" de estudantes, talvez você não soubesse quais as proporções adequadas dos tipos diversos de alunos para tal equilíbrio, ou talvez você não tenha como identificar os diferentes tipos, apenas vendo-os passar.

Mesmo fazendo um esforço de entrevistar cada décimo estudante que entrasse na biblioteca universitária, você não garantiria uma amostra representativa, uma vez que tipos diferentes de alunos visitam a biblioteca com frequências diferentes. Portanto, a amostra sobre-representaria aqueles estudantes que usam a biblioteca com mais frequência.

Representatividade e Probabilidade de Seleção

Amostras de *surveys* devem representar as populações das quais são retiradas, se devem fornecer estimativas úteis quanto às características daquela população. Não necessitam, contudo,

ser representativas em todos os aspectos; representatividade, no sentido que tem para a amostragem, limita-se às características relevantes para os interesses substantivos da pesquisa. (Este ponto ficará mais evidente nas discussões posteriores sobre estratificação.)

Um princípio básico da amostragem probabilística é: *uma amostra será representativa da população da qual foi selecionada se todos os membros da população tiverem oportunidade igual de serem selecionados para a amostra.*³ Amostras com esta qualidade são freqüentemente rotuladas de *amostras MIPSE* (método de igual probabilidade de seleção). Embora discutamos mais tarde variações deste princípio, trata-se do princípio primordial que fundamenta a amostragem probabilística.

Indo além deste princípio básico, devemos entender que amostras — mesmo as amostras MIPSE cuidadosamente selecionadas — raramente, talvez nunca, são *perfeitamente representativas* das populações das quais são retiradas. A despeito dessa limitação, a amostragem probabilística oferece duas vantagens especiais para os pesquisadores.

Primeiro, amostras probabilísticas, embora jamais perfeitamente representativas, são tipicamente *mais representativas* do que outros tipos de amostras, porque os vieses discutidos acima são evitados. Na prática, há maior possibilidade de uma amostra probabilística ser representativa da população da qual é extraída do que uma amostra não probabilística.

Segundo, e mais importante, a teoria da probabilidade permite estimar a precisão ou representatividade da amostra. É concebível que você, por meios totalmente aleatórios, selecione uma amostra que represente quase perfeitamente a população maior. No entanto, é muito mais provável que não o faça, e você não tem como avaliar a probabilidade de ter atingido a representatividade. Por outro lado, a amostra probabilística pode fornecer uma estimativa precisa do seu sucesso ou fracasso. Após apresentar uma breve taxonomia da terminologia da amostragem, examinaremos os meios pelos quais se pode estimar a representatividade da sua amostra.

Conceitos e Terminologia de Amostragem

As seguintes discussões da teoria e prática da amostragem usam vários termos técnicos. Para facilitar a compreensão das discussões, é importante defini-los. Na maior parte, usaremos termos habitualmente usados em manuais sobre amostragem e estatística.

Ao apresentar esta taxonomia sobre terminologia e conceitos de amostragem, quero reconhecer uma dívida para com Leslie Kish e seu excelente manual sobre o *survey* por amostragem.⁴ Embora eu tenha modificado algumas das convenções usadas por Kish, a sua apresentação é a mais importante fonte para a nossa discussão.

Elemento

Um *elemento* é a unidade sobre a qual a informação é coletada, e que serve de base para a análise. Tipicamente, na pesquisa de *survey*, os elementos são pessoas ou certos tipos de pessoas. Entretanto, outros tipos de unidades — famílias, clubes, corporações etc. — podem também constituir os elementos de um *survey*.

Universo

Um *universo* é a agregação teórica e hipotética de todos os elementos definidos num *survey*. Se o americano individual for o elemento do *survey*, então “americanos” serão o universo. Porém, um universo de *survey* é totalmente não especificado quanto a tempo e lugar e é essencialmente um termo inútil.

População

População é a agregação teoricamente especificada de elementos do *survey*. Embora o termo vago “americanos” possa constituir o universo de um *survey*, delinear a população inclui definir o elemento “americano” (por exemplo, cidadania e residência) e o referencial de tempo para a pesquisa (americanos desde quando?) Traduzir o universo “nova-iorquinos adultos” para uma população trabalhável requer especificar a idade, definindo “adulto”, as fronteiras de Nova Iorque etc. Especificar o termo “estudante universitário” inclui levar em conta estudantes de horário integral e parcial, candidatos a diplomas ou não, estudantes dos cursos de graduação e pós-graduação e atributos similares.

Embora se tenha que começar especificando cuidadosamente a população, em geral a licença poética permite que você escreva seu relatório em termos do universo hipotético. Por facilidade de apresentação, mesmo o pesquisador mais consciencioso normalmente fala de “americanos” e não de “cidadãos residentes nos Estados Unidos da América a partir de 12 de novembro de

1989". A orientação principal neste assunto, como na maioria dos outros, é não confundir ou iludir os leitores.

População do *survey*

Uma *população do survey* é a agregação de elementos da qual é de fato extraída a amostra do *survey*. Lembre-se de que uma população é uma especificação teórica do universo. Na prática, raramente você pode garantir que todos os elementos que se enquadram nas definições teóricas estabelecidas têm efetivamente a chance de serem selecionados para a amostra. Mesmo as listas de elementos compiladas especificamente para fins de amostragem costumam ser algo incompletas. Alguns estudantes são sempre inadvertidamente omitidos das listagens de alunos. Há usuários de telefones que exigem que seus nomes e números não sejam incluídos nos catálogos. Portanto, a população de *survey* é a agregação dos elementos dos quais a amostra é selecionada.

Muitas vezes, pesquisadores decidem limitar suas populações de *survey* mais severamente do que o indicado nos exemplos acima. Firmas de pesquisa de opinião podem limitar suas "amostras nacionais" aos quarenta e oito estados adjacentes, omitindo, por razões práticas, o Alasca e o Havaí. Um pesquisador, ao selecionar uma amostra de professores de psicologia, pode limitar a população do *survey* aos professores de psicologia em departamentos de psicologia, omitindo os que pertencem a outros departamentos. (Num certo sentido, podemos dizer que estes pesquisadores redefiniram seus universos e populações, contanto que tenham tornado estas revisões claras para os leitores.)

Unidade de Amostra

Uma *unidade de amostra* é o elemento ou conjunto de elementos considerados para seleção em alguma etapa da amostragem. Numa amostra simples, de estágio único, as unidades de amostra são o mesmo que os elementos. Mas, em amostras mais complexas, pode-se usar níveis diferentes de unidades de amostra. Por exemplo, um pesquisador pode, primeiro, extrair uma amostra de bairros numa cidade, depois uma amostra de residências nesses bairros e, enfim, uma amostra de adultos nas residências selecionadas. As unidades de amostra, nestas três etapas de amostragem são, respectivamente, bairros, residências e adultos, sendo

que só os últimos são os elementos. De forma mais específica, os termos "unidades primárias de amostra," "unidades secundárias de amostra" e "unidades finais de amostra" devem ser usadas para designar as etapas sucessivas.

Moldura de Amostragem

Uma *moldura de amostragem* é a lista de unidades de amostra da qual a amostra, ou alguma etapa dela, é selecionada. Se uma amostra simples de estudantes for selecionada de uma listagem de estudantes, a listagem seria a moldura de amostragem. Se a unidade primária de amostra para uma amostra populacional complexa for o bairro, a lista dos bairros é a moldura de amostragem — na forma de livreto impresso, cartão arquivo da IBM, fita magnética ou arquivo em disquete.

No planejamento de amostra de etapa única, a moldura de amostragem é a lista dos elementos compondo a população de *survey*. Na prática, muitas vezes são as molduras de amostragem disponíveis que definem as populações de *survey*, e não o contrário. Frequentemente, você começa com um universo ou talvez uma população em mente para sua pesquisa e depois procura possíveis molduras de amostragem. As molduras disponíveis são examinadas e avaliadas, e você escolhe a moldura que representa a população de *survey* mais apropriada para seus fins.

A relação entre populações e molduras de amostragem é crítica, não tendo recebido atenção suficiente. Uma seção posterior deste capítulo aprofunda a questão.

Unidade de Observação

Uma *unidade de observação*, ou unidade de coleta de dados, é um elemento ou agregação de elementos de que se coleta informação. A unidade de análise e a unidade de observação muitas vezes são o mesmo — a pessoa individual —, mas isso não é necessário. Por exemplo, entrevistar chefes de família (as unidades de observação) para coletar informação sobre todos os seus membros (as unidades de análise).

A tarefa é mais simples quando a unidade de análise e a unidade de observação são o mesmo. Mas, frequentemente, isso não é possível ou viável, obrigando-o a usar alguma engenhosidade para obter dados relevantes de suas unidades de análise, sem de fato observá-las.

Variável

Uma *variável* é um conjunto de características mutuamente excludentes, como sexo, idade, emprego etc. Podemos descrever os elementos de uma população em termos das suas características individuais numa variável. Tipicamente, *surveys* visam descrever a distribuição das características de uma variável numa população. Assim, você pode descrever a distribuição etária de uma população examinando a frequência relativa das diferentes idades dos seus membros.

Repare que uma variável, por definição, deve ter *variação*; se todos os elementos na população têm a mesma característica, esta característica é uma constante na população e não parte de uma variável.

Parâmetro

Um *parâmetro* é a descrição sumária de uma variável numa população. A renda média de todas as famílias numa cidade e a distribuição etária da população da cidade são parâmetros. Parte importante da pesquisa por *surveys* envolve a estimativa dos parâmetros populacionais baseada em observações amostrais.

Estatística

Uma *estatística* é a descrição sumária de uma variável num *survey* por amostragem. Portanto, a renda média computada a partir de um *survey* por amostragem e a distribuição etária daquela amostra são estatísticas. Usa-se estatísticas de amostras para fazer estimativas sobre parâmetros populacionais.

Erro amostral

O *erro amostral* é discutido com mais detalhes numa seção posterior deste capítulo. Métodos de amostragem probabilística raramente ou nunca fornecem estatísticas exatamente iguais aos parâmetros que buscam estimar. Entretanto, a teoria da probabilidade permite estimar o grau de erro a ser esperado num determinado desenho de amostragem.

Níveis de Confiança e Intervalos de Confiança

Estes termos também são discutidos com mais detalhes numa seção posterior. Computar o erro amostral permite expressar a precisão das suas estatísticas em termos do nível de confiança,

estabelecendo que essas estatísticas estão dentro de um intervalo especificado do parâmetro. Por exemplo, você pode dizer que está "95% confiante" que sua estatística (por exemplo, 50% favorecem o candidato X) está dentro de mais ou menos (+) 5 pontos percentuais do parâmetro populacional. Quando você expande o intervalo de confiança para uma dada estatística, sua "confiança" aumenta; você pode dizer que tem 99,9% de confiança de que sua estatística está dentro de + 7,5 pontos percentuais do parâmetro.

Teoria da Amostragem Probabilística e Distribuição Amostral

Esta seção examina a teoria básica da amostragem probabilística tal como é aplicada ao *survey* por amostragem; consideraremos a lógica da distribuição amostral e erro amostral em relação a uma *variável binomial*, ou seja, uma variável consistindo em duas características.

Teoria de Amostragem Probabilística

A finalidade última do *survey* por amostragem é selecionar um conjunto de elementos de uma população de tal forma que descrições destes elementos (estatísticas) descrevam com precisão a população total da qual foram selecionadas. A amostragem probabilística fornece um método para ampliar a possibilidade de alcançar esta meta, bem como métodos para estimar o grau de sucesso provável.

Seleção aleatória é a chave deste processo. Num processo de seleção aleatória, cada elemento tem chance igual de seleção, independente de qualquer outro evento do processo de seleção. Lançar uma moeda perfeita é o exemplo mais frequentemente citado; tirar uma cara ou uma coroa é independente de seleções anteriores de caras ou coroas. Outro exemplo é jogar um conjunto de dados perfeitos.

Estas imagens de seleção aleatória, contudo, raramente se aplicam diretamente aos métodos de *survey* por amostragem. É bem mais provável você usar tabelas de números aleatórios ou programas de computador que selecionam aleatoriamente as unidades de amostragem. A disponibilidade destes auxílios à pesquisa torna-os um bom começo para nossa discussão sobre a amostragem aleatória.

As razões para usar métodos de seleção aleatória, ou seja, tabelas de números aleatórios ou programas de computador,

Distribuição Binomial da Amostragem

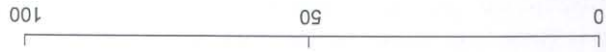
são duas. Primeiro, este procedimento funciona como antídoto aos vieses conscientes ou inconscientes do pesquisador. O pesquisador que seleciona casos de forma intuitiva pode muito bem selecionar casos que apóiam expectativas ou hipóteses de pesquisa preexistentes. Segundo, a seleção aleatória dá acesso à teoria da probabilidade, que é a base para as estimativas sobre parâmetros populacionais e as estimativas de erro. Passemos agora ao exame deste segundo aspecto.

A forma mais clara de discutir o conceito da distribuição amostral é usar um exemplo simples de *survey*. Suponha que queremos estudar a população estudantil de uma universidade para medir a aprovação ou rejeição de um código de conduta estudantil proposto pela administração. A população do *survey* será a agregação de estudantes relacionados numa listagem de alunos (a moldura de amostragem). Os elementos serão os estudantes individuais da universidade. A variável considerada será atitudes quanto ao código, uma variável binomial — aprovar ou rejeitar. Selecionaremos uma amostra aleatória de estudantes para estimar todo o corpo estudantil.

A Figura 5-1 apresenta um eixo de ordenadas representando todos os valores possíveis deste parâmetro na população — de 0 até 100% de aprovação. O ponto médio do eixo — 50% — representa a situação na qual metade dos alunos aprova o código, enquanto a outra metade rejeita-o. Vamos supor que demos um número a cada aluno da listagem, e que selecionamos 100 números de uma tabela de números aleatórios. Os 100 estudantes cujos números forem relacionados ao código estudantil — se aprovam ou rejeitam. Suponhamos que esta operação nos dá 48 alunos que aprovam o código e 52 que rejeitam. Podemos representar esta estatística marcando a ordenada no ponto correspondente a 48%.

Agora suponha que selecionamos outra amostra de 100 alunos exatamente da mesma forma, que medimos sua aprovação/rejeição do código estudantil e que 51 estudantes desta amostra aprovam o código; esse resultado pode ser representado por outra marca no local apropriado da ordenada. Repetindo o processo, podemos descobrir que, numa terceira amostra, 52 alunos aprovam o código.

FIGURA 5-1
Porcentagem de alunos que aprovam o código estudantil



A Figura 5-2 mostra as três diferentes estatísticas representando as porcentagens de alunos em cada uma das três amostras aleatórias que aprovam o código estudantil. A regra básica da amostragem aleatória é que as amostras extraídas de uma população fornecem estimativas do parâmetro relativo à população total. Portanto, cada amostra aleatória dá uma estimativa da porcentagem de estudantes em todo o corpo estudantil que aprovam o código de conduta. No entanto, infelizmente selecionamos três amostras e agora temos três estimativas distintas.

FIGURA 5-2

Porcentagem de alunos que aprovam o código estudantil

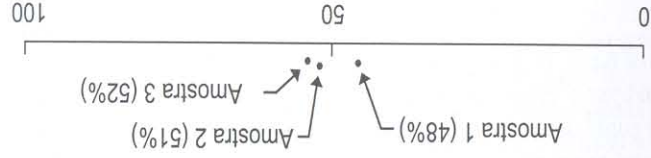
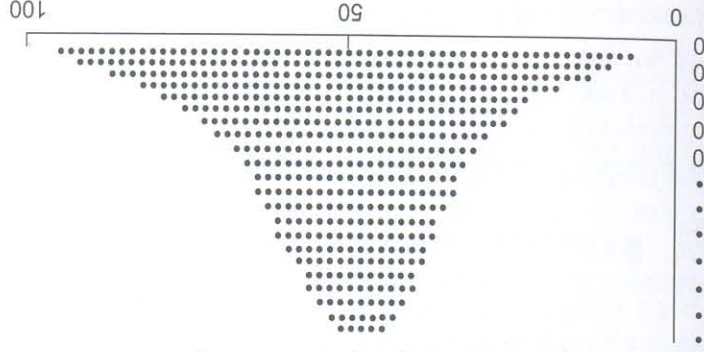


FIGURA 5-3

Porcentagem de alunos que aprovam o código estudantil



Para escapar deste dilema, vamos extrair mais e mais amostras de 100 alunos cada, perguntar a cada uma das amostras sobre sua aprovação/rejeição do código e marcar as novas estatísticas no nosso gráfico de resumo. Ao extrair muitas amostras, vamos descobrir que algumas das novas amostras repetem as estimativas dadas por amostras anteriores. Para explicar esta situação,

acrescentemos um eixo de abcissas à figura, representando o número de amostras que fornecem estimativas. A Figura 5-3 mostra o produto dos nossos novos esforços amostrais.

A distribuição das estatísticas amostrais, mostrada na Figura 5-3, é chamada *distribuição amostral*. Ao aumentar o número de amostras selecionadas e entrevistadas, também aumentamos a extensão das estimativas fornecidas pela operação de amostragem. Num sentido, agravamos nosso dilema de tentar adivinhar o parâmetro da população. No entanto, a teoria da probabilidade oferece certas regras importantes relativas à distribuição de amostras mostrada na Figura 5-3.

Primeiro, se muitas amostras aleatórias independentes são extraídas de uma população, as estatísticas amostrais fornecidas por estas amostras estarão *distribuídas em torno do parâmetro populacional* de uma forma conhecida. Embora haja uma ampla faixa de estimativas, a maioria está mais próxima dos 50% do que de outra parte do gráfico. Portanto, a teoria da probabilidade nos diz que o valor real está próximo aos 50%.

Segundo, a teoria da probabilidade nos dá uma fórmula para estimar *quão próximo* as estatísticas da amostragem estão aglomeradas em torno do valor real. Esta fórmula contém três fatores: o parâmetro, o tamanho da amostra e o *erro padrão* (uma medida do erro da amostragem).

$$\text{Fórmula: } S = \sqrt{\frac{PQ}{n}}$$

Símbolos: P , Q : Os parâmetros populacionais para o binomial; se 60% do corpo estudantil aprovam o código e 40% rejeitam, P e Q são 60% e 40%, ou 0,6 e 0,4. Observe que $Q = 1 - P$ e $P = 1 - Q$.

n : O número de casos em cada amostra.

S : O erro padrão.

Suponha que o parâmetro populacional seja 50% de aprovação e 50% de rejeição. Recorde que estivemos selecionando amostras de 100 casos cada. Quando estes números são postos na fórmula, vemos que o erro padrão é igual a 0,05, ou 5%.

Em termos da teoria da probabilidade, o erro padrão é um dado valioso porque indica em que medida as estimativas amostrais estão distribuídas em torno do parâmetro populacional. Especificamente, a teoria da probabilidade indica que certas proporções das estimativas amostrais estão dentro de incrementos especificados de erros padrão do parâmetro populacional. Aproximadamente 34% (0,3413) das estimativas amostrais estão até um desvio padrão acima do parâmetro da

população, e outros 34% até um desvio padrão abaixo do parâmetro. No nosso exemplo, o erro padrão é 5%, então sabemos que 34% das amostras darão estimativas de aprovação estudantil entre 50% (parâmetro) e 55% (um erro padrão acima); outros 34%, das amostras mostrarão estimativas entre 50% e 45% (um desvio padrão abaixo do parâmetro). Reunindo-as, sabemos que aproximadamente dois terços (68%) das amostras resultarão em estimativas dentro (mais ou menos) de 5% do parâmetro.

Além disso, a teoria da probabilidade diz que cerca de 95% das amostras estarão dentro de (mais ou menos) dois erros padrão do valor verdadeiro, e 99,9%, das amostras dentro de (mais ou menos) três desvios padrão. Neste exemplo, portanto, sabemos que só uma amostra em mil resultaria numa estimativa abaixo de 35% de aprovação ou acima de 65% de aprovação.

A proporção de amostras caindo entre um, dois ou três erros padrão do parâmetro é constante para qualquer procedimento de amostragem aleatória como a que foi descrita, contanto que se selecione grande número de amostras. Contudo, o tamanho do erro padrão é uma função do parâmetro populacional e do tamanho da amostra. Se voltarmos à fórmula, vemos que o erro padrão aumenta como função de um aumento na quantidade PQ . Veja, também, que esta quantidade atinge o máximo quando a população é dividida ao meio. Se $P = 0,5$, $PQ = 0,25$; se $P = 0,6$, $PQ = 0,24$; se $P = 0,8$, $PQ = 0,16$; se $P = 0,99$, $PQ = 0,0099$. Por extensão, se P for 0,0 ou 1,0 (isto é, se 0% ou 100% aprovarem o código estudantil), o erro padrão será 0. Se todos na população têm a mesma atitude (sem variação), então qualquer amostra dará exatamente a mesma estimativa.

O erro padrão é também uma função inversa do tamanho da amostra. À medida que o tamanho da amostra cresce, o erro padrão diminui, e as várias amostras estarão aglomeradas mais próximas ao valor verdadeiro. Outra regra é evidente a partir da fórmula: devido à raiz quadrada na fórmula, o erro padrão fica reduzido pela metade se o tamanho da amostra for *quadruplicado*. Em nosso exemplo, amostras de 100 produzem um erro padrão de 5%; para reduzir o erro padrão a 2,5%, teríamos que aumentar o tamanho da amostra para 400. (O apêndice B traz uma tabela para ajudá-lo a fazer essas estimativas.)

O que discutimos provém da teoria da probabilidade estabelecida, referente à seleção de grandes números de amostras aleatórias. Se o parâmetro populacional for conhecido, e muitas amostras aleatórias forem selecionadas, poderemos prever quantas amostras cairão dentro de intervalos especificados a partir do parâmetro. No entanto, estas condições não são típicas do *survey* por amostragem. Usualmente, o

pesquisador desconhece o parâmetro e conduz um *survey* por amostragem para estimá-lo. Ademais, os pesquisadores de *survey* normalmente não selecionam grandes números de amostras, em geral extraindo só uma. Apesar disso, a discussão acima sobre a teoria da probabilidade fornece uma base para inferências sobre a situação típica dos *surveys*.

Se a teoria da probabilidade diz que 68% das amostras estarão dentro de um erro padrão do parâmetro, o pesquisador infere que uma amostra aleatória tem uma probabilidade de 68% de estar dentro daquela faixa. Neste sentido, falamos de *níveis de confiança*: você tem “68% de confiança” que a estimativa da sua amostra está dentro de um erro padrão do parâmetro, ou tem “95% de confiança” que a estatística da amostra está dentro de dois erros padrão do parâmetro, e assim por diante. Compreensivelmente, a confiança aumenta quanto mais se estende a margem para erro. Você está virtualmente certo (99,9% de confiança) que a estimativa da amostra está dentro de três erros padrão do valor verdadeiro.

Embora confiante (em algum nível) de estar dentro de uma certa faixa do parâmetro, já assinalamos que raramente você conhece o parâmetro. Para resolver este dilema, você substitui o parâmetro na fórmula pela sua estimativa de amostragem; na falta do valor verdadeiro, você o substitui com a melhor estimativa disponível.

O resultado destas inferências e estimativas é que o pesquisador de *survey* torna-se capaz de estimar tanto um parâmetro populacional quanto o grau esperado de erro com base numa amostra extraída da população. Começando com a pergunta: “Qual porcentagem do corpo estudantil aprova o novo código?”, você pode selecionar uma amostra aleatória de 100 alunos e entrevistá-los. Você pode relatar que sua melhor estimativa é que 50% do corpo estudantil aprovam o código, e tem 95% de confiança que entre 40 e 60% (mais ou menos dois erros padrão) aprovam. A faixa de 40 a 60% é chamada de *intervalo de confiança*.

Esta é a lógica básica da amostragem probabilística. A seleção aleatória permite ligar os achados do *survey* por amostragem ao corpo da teoria da probabilidade para estimar a precisão daqueles achados. Todos os enunciados de precisão na amostragem devem especificar tanto um nível de confiança quanto um intervalo de confiança. Você pode informar que está $x\%$ confiante de que o parâmetro populacional está entre dois valores especificados.

A discussão precedente considerou apenas um tipo de estatística — as porcentagens produzidas por uma variável binomial ou dicotômica. A mesma lógica se aplica ao exame de outras estatísticas, tais como renda média. Uma vez que as computações são mais complicadas, nesta introdução consideramos apenas os binomiais.

Advertimos que o uso da teoria da probabilidade, como discutido nesta seção, não está plenamente justificado em termos técnicos. A teoria da distribuição amostral, por exemplo, faz suposições que quase nunca se aplicam às situações de *surveys*. A fórmula usada para avaliar o número de amostras que estariam contidas dentro de incrementos especificados de erros padrão, por exemplo, supõe uma população infinita, um número infinito de amostras e amostragem com substituição. Ademais, na nossa discussão, o salto inferencial da distribuição de várias amostras para as características prováveis de uma amostra foi grosseiramente simplificado.

Essas advertências são feitas por razões de perspectiva. Pesquisadores, às vezes, parecem superestimar a precisão das estimativas produzidas pelo uso da teoria da probabilidade em conexão com a pesquisa por *survey*. Mencionaremos em outras partes deste capítulo e do livro que variações nas técnicas de amostragem e fatores de não-amostragem podem reduzir ainda mais a legitimidade de tais estimativas. Não obstante, os cálculos discutidos nesta seção podem ser de grande valor para compreender e avaliar dados. Embora não forneçam estimativas tão precisas quanto se pode supor, podem ser bastante válidos para fins práticos, e são inquestionavelmente mais válidos do que estimativas derivadas menos rigorosamente, baseadas em métodos de amostragem menos rigorosos. Você deve se familiarizar com a *lógica* básica subjacente aos cálculos, para poder reagir sensatamente tanto aos seus próprios dados quanto aos dados informados por outros.

Populações e Molduras de Amostragem

A seção anterior lidou com um modelo teórico para amostragem de *survey*. Se o consumidor de *surveys*, o estudante ou o pesquisador precisam conhecer essa teoria, devem igualmente saber avaliar as condições imperfeitas existentes no campo. Esta seção discute um aspecto das condições de campo que requer compromisso entre as condições e os supostos teóricos — a congruência ou a disparidade entre populações e molduras de amostragem.

Dito simplesmente, uma moldura de amostragem é a lista, ou fac-símile razoável, de elementos dos quais é selecionada uma amostra probabilística. A próxima seção trata dos métodos de selecionar amostras, mas primeiro precisamos considerar a própria moldura de amostragem. Amostras corretamente selecionadas fornecem informações apropriadas para descrever a população de elementos que compõem a moldura de amostragem — nada mais. É preciso enfatizar este ponto, dada a tendência tão comum entre pesquisadores de selecionar amostra de uma dada moldura de amostragem e então fazer afirmações sobre uma população semelhante, mas não idêntica, à população do *survey* definida pela moldura de amostragem. É o problema mais amplo da generalização social-científica, equivalente a estudar uma pequena Igreja Luterana em Dakota do Norte para descrever a religião nos EUA. No restante da seção examinaremos finalidades diversas para os *surveys* e discutiremos boas e más molduras de amostragem para aquelas finalidades.

Surveys de organizações freqüentemente são o tipo mais simples de *survey*, do ponto de vista da amostragem, porque, tipicamente, organizações possuem listas de membros. Nestes casos, a lista de membros é uma excelente moldura de amostragem. Se uma amostra aleatória é selecionada de uma lista de membros, os dados colhidos desta amostra podem ser considerados como representativos de todos os membros — *caso todos os membros forem incluídos na lista*. Se alguns não tiverem sido incluídos, deve-se fazer um esforço para amostrá-los; do contrário, os achados do *survey* só podem ser considerados como representativos daqueles membros na lista.

Populações que oferecem boas listas organizacionais são escolas primárias e secundárias, estudantes universitários; membros de igrejas; operários de fábricas; membros de fraternidades e grêmios, de clubes sociais, de serviço ou políticos; e membros de associações profissionais.

Estes comentários se aplicam principalmente a organizações locais. Organizações estaduais ou nacionais freqüentemente não têm uma única lista de membros facilmente disponível. Por exemplo, não há uma lista única de todos os membros da Igreja Episcopal. Mas um planejamento de amostragem um pouco mais complexo pode aproveitar as listas de membros de igrejas locais, amostrando primeiro igrejas e depois subamostrando as listas de membros dessas igrejas.

Outras listas de indivíduos podem ser especialmente relevantes às necessidades de pesquisa de *survey*. Por exemplo,

agências governamentais mantêm listas de eleitores registrados caso você queira fazer uma pesquisa pré-eleitoral de opinião ou um exame mais detalhado do comportamento eleitoral, mas você deve certificar-se de que a lista está atualizada. Listas semelhantes contêm nomes de donos de automóveis, de recebedores de benefícios sociais, de contribuintes de impostos, pessoas com licenças comerciais, profissionais licenciados etc. Embora algumas destas listas sejam difíceis de conseguir, podem fornecer excelentes molduras de amostragem. Como os elementos de amostragem num *survey* não precisam ser pessoas, há diversas listas com outros tipos de elementos: universidades, negócios de vários tipos, cidades, revistas acadêmicas, jornais, sindicatos, clubes políticos, associações profissionais etc.

Catálogos telefônicos são muito usados para pesquisas de opinião “rápidas e sujas”. Inegavelmente são fáceis de fazer e de baixo custo, o que sem dúvida responde pela sua popularidade, e se você quiser fazer afirmações sobre assinantes de telefones, o catálogo é uma moldura de amostragem *razoavelmente boa*.⁵ Infelizmente, com excessiva freqüência, os catálogos telefônicos são usados como lista da população de uma cidade ou dos seus eleitores. Há muitos defeitos neste tipo de raciocínio, o principal sendo um viés de classe social. Gente pobre tem menos possibilidade de ter telefones; uma amostra de catálogo telefônico, portanto, provavelmente terá uma predisposição para as classes média e alta. (O Capítulo 10 examina um método especial para amostragem em relação a *surveys* telefônicos: *discagem aleatória de algarismos*.)

Catálogos de ruas e mapas de impostos são freqüentemente usados como fontes fáceis para amostras de residências, mas podem sofrer das mesmas desvantagens dos catálogos telefônicos, tais como serem incompletos e/ou enviesados. Por exemplo, em regiões urbanas com zoneamento rigoroso, unidades residenciais ilegais têm pouca possibilidade de aparecerem em registros oficiais. Resulta que, não tendo estas unidades chance de seleção, os achados das amostras não são representativas dessas unidades, tipicamente mais pobres e superlotadas.

Revisão de Populações e Molduras de Amostragem

Uma vez que a literatura da pesquisa por *surveys* tem sido surpreendentemente descuidada com as questões de populações e molduras de amostragem, dou-lhes atenção especial. Para maior ênfase, listo, em revisão, as principais diretrizes a lembrar.

1. Os achados dos *surveys* por amostragem só podem ser considerados representativos da agregação de elementos que compõem a moldura de amostragem.

2. Muitas vezes, as molduras de amostragem não incluem de fato todos os elementos que seus nomes sugerem. (Diretórios de estudantes não incluem todos os estudantes; catálogos telefônicos não incluem todos os assinantes.) Omissões são quase inevitáveis. Portanto, uma primeira preocupação deve ser avaliar a extensão destas omissões e corrigi-las se possível. (Naturalmente, talvez você possa ignorar, sem prejuízo, um pequeno número de omissões de difícil correção.)

3. Para poder generalizar a população que compõe a moldura de amostragem, é necessário que todos os elementos tenham representação igual nela; cada elemento deve aparecer uma só vez. Do contrário, elementos que aparecem mais de uma vez terão maior probabilidade de seleção, e a amostra, no geral, sobre-representará estes elementos.

Outros assuntos mais práticos, relacionados às populações e molduras de amostragem, são tratados em outros capítulos. Por exemplo, o formato da moldura de amostragem — uma lista numa publicação, um arquivo em fichas 3x5, placas *addressograph*, cartões IBM, fitas magnéticas, disquetes — é muito importante. Tais considerações freqüentemente ganham precedência sobre considerações científicas; uma lista mais “fácil” poderá ser escolhida contra uma lista mais “difícil”, mesmo que esta seja mais apropriada à população alvo. Não devemos ser dogmáticos a este respeito; os pesquisadores devem pesar cuidadosamente as vantagens e as desvantagens relativas destas alternativas. O mais importante é que você esteja ciente — e deve informar seu leitor — das deficiências de qualquer moldura de amostragem que escolher.

Tipos de Desenhos de Amostragem

Introdução

Talvez, a essa altura, você esteja apavorado com a importância e as dificuldades para organizar sua moldura de amostragem; é um sentimento correto e saudável. Após estabelecer a moldura de amostragem, você deve selecionar uma amostra de elementos para estudo.

Até agora, focalizamos a *amostragem aleatória simples*. De fato, as estatísticas usadas em geral pelos pesquisadores de *survey* supõem tal amostra. Mas, como logo veremos, há diversas opções para fazer a amostragem, e raramente se opta pela aleatória simples, por duas razões. Primeiro, a amostragem aleatória simples só é possível com as molduras de amostragem mais simples. Segundo, a amostragem aleatória simples pode não ser o método melhor (o mais preciso). Discutiremos agora a amostragem aleatória simples e outros tipos de desenhos de amostragem.

Amostragem Aleatória Simples

A amostragem aleatória simples [AAS] é o método de amostragem básico suposto pelos cálculos estatísticos do *survey*. A matemática da amostragem aleatória é particularmente complexa, e nos desviaremos dela, preferindo descrever os métodos de campo que empregam este método.

Uma vez estabelecida uma moldura de amostragem de acordo com a discussão anterior, você numera cada elemento da lista, atribuindo um e só um número a cada um, sem saltar nenhum número. A seguir, usa uma tabela de números aleatórios para selecionar os elementos da amostra.

Se sua moldura de amostragem está em formato próprio para computadores, como um disquete, uma amostra aleatória simples pode ser extraída automaticamente, usando um programa de computador relativamente simples. Com efeito, o computador pode numerar os elementos da moldura de amostragem, gerar sua própria série de números aleatórios e imprimir a lista de elementos selecionados.

Amostragem Sistemática

A amostragem aleatória simples raramente é usada na prática. Em geral, não é o método de amostragem mais eficiente, e já vimos que pode ser trabalhoso quando feito manualmente. A AAS habitualmente requer uma lista de elementos; quando esta lista está disponível, geralmente os pesquisadores usam o método de *amostragem sistemática*, em vez da amostragem aleatória simples.

Na amostragem sistemática, cada k° elemento na lista total (por exemplo, cada 100^o) é escolhido para inclusão na amostra. Se a lista tem 10.000 elementos, e você quer uma amostra de 1.000 elementos, seleciona cada décimo elemento para a amostra. Para garantia contra algum viés humano

possível ao usar este método, você seleciona o primeiro elemento aleatoriamente. No exemplo acima, você começa escolhendo um número aleatório entre 1 e 10; o elemento com este número é incluído na amostra, mais cada décimo elemento seguinte. Refere-se tecnicamente a este procedimento como uma “amostra sistemática de início aleatório”.

Dois termos são usados freqüentemente em conexão com a amostragem sistemática. O *intervalo de amostragem* é a distância padrão entre os elementos selecionados na amostra — 10 no exemplo acima. A *razão amostral* é a proporção de elementos selecionados da população — um décimo no exemplo acima. (A razão amostral é igual a $1/\text{intervalo de amostragem}$.)

Na prática, a amostragem sistemática é virtualmente idêntica à amostragem aleatória simples. De fato, se a lista de elementos for “randomizada” antes da amostragem, pode-se argumentar que uma amostragem sistemática retirada daquela lista é, na realidade, uma amostra aleatória simples. Os debates quanto aos méritos relativos da amostragem aleatória simples e da amostragem sistemática foram resolvidos largamente em favor do método mais simples, a sistemática. Empiricamente, os resultados são virtualmente idênticos. Além disso, como veremos, a amostragem sistemática, em algumas situações, é um pouco mais precisa do que a amostragem aleatória simples.

A amostragem sistemática traz um perigo: o arranjo dos elementos na lista pode desaconselhar uma amostra sistemática. Este perigo é em geral referido pelo termo *periodicidade*. Se a lista de elementos estiver organizada num padrão cíclico que coincide com o intervalo de amostragem, resultará uma amostra grosseiramente tendenciosa. Dois exemplos bastam.

Num estudo de soldados durante a II Guerra Mundial, os pesquisadores selecionaram cada décimo soldado das listas das unidades. Mas estas estavam organizadas sob a forma de uma tabela organizacional, listando primeiro os sargentos e depois cabos e praças, pelotão por pelotão, sendo que cada pelotão tinha dez membros. Portanto, cada décima pessoa na lista era um sargento de pelotão, e a amostra sistemática selecionada tinha apenas sargentos. Poderia acontecer, também, que nenhum sargento fosse selecionado.

Em outro exemplo, suponha que queremos selecionar uma amostra de apartamentos num prédio. Se a amostra for extraída de uma lista dos apartamentos organizados em ordem numérica (por exemplo, 101, 102, 103, 104, 201, 202, e assim por diante), há o risco de o intervalo de amostragem coincidir

com o número de apartamentos num andar ou seu múltiplo. Neste caso, a amostra poderia incluir apenas apartamentos do lado noroeste, ou só apartamentos perto dos elevadores. Se estes tipos de apartamentos tiverem alguma característica peculiar em comum (por exemplo, aluguel mais elevado), a amostra seria tendenciosa. Risco semelhante pode surgir numa amostragem sistemática de casas numa subdivisão organizada com o mesmo número de casas num quarteirão.

Ao considerar uma amostra sistemática a partir de uma lista, você deve examinar cuidadosamente a natureza da lista. Se os elementos estão arrançados em alguma ordem particular, certifique-se de que aquela ordem não envies a amostra a ser selecionada e tome medidas para contrabalançar qualquer possível viés (por exemplo, tirar uma amostra aleatória simples de porções cíclicas).

Em resumo, a amostragem sistemática, em geral, é superior à amostragem aleatória simples, pelo menos em conveniência. Os problemas na ordenação dos elementos da moldura de amostragem geralmente podem ser remediados com facilidade.

Amostragem Estratificada

Nas duas seções precedentes, discutimos dois métodos alternativos de seleção de amostras a partir de uma lista. A *amostragem estratificada* não é uma alternativa a estes métodos, mas representa uma possível modificação no seu uso.

As amostragens aleatória simples e sistemática são importantes no sentido de garantirem um grau de representatividade e permitirem uma estimativa de erro. A amostragem estratificada é um método para obter maior grau de representatividade, reduzindo o provável erro amostral. Para compreender por que, voltemos rapidamente à teoria básica da distribuição amostral.

Lembre que o erro amostral é reduzido por dois fatores no desenho da amostra. Primeiro, uma amostra maior produz menor erro amostral do que uma amostra menor. Segundo, uma população homogênea produz amostras com menores erros amostrais do que uma população heterogênea. Se 99% da população concordam com determinado enunciado, é improvável que qualquer amostra probabilística interprete erroneamente a extensão da concordância. Se, por outro lado, a população está dividida meio a meio quanto ao enunciado, então o erro amostral será muito maior.

A amostragem estratificada baseia-se neste segundo fator na teoria de amostragem. Em vez de selecionar a amostra da

população total em geral, você garante que quantidades apropriadas de elementos sejam tirados de subconjuntos homogêneos da população. Por exemplo, estudando estudantes universitários, você pode organizar sua população por turmas, tirando de cada uma quantidades apropriadas. Numa amostra não estratificada, a representação por turma estaria sujeita ao mesmo erro amostral que outras variáveis. Na amostra estratificada por turma, o erro amostral quanto a esta variável reduz-se a zero.

Você pode querer usar um método de estratificação ainda mais complexo. Além de estratificar por turma, pode também estratificar por gênero, por média de notas etc. Assim, talvez você garanta que sua amostra tenha números apropriados de alunos da turma A com nota média 4,0; também números apropriados de alunas dessa turma com média 4,0 etc.

Portanto, a função última da estratificação é organizar a população em subconjuntos homogêneos (com heterogeneidade entre os subconjuntos) e selecionar o número apropriado de elementos de cada subconjunto. Já que os subconjuntos são homogêneos nas variáveis de estratificação, também podem ser homogêneos quanto a outras variáveis. Por exemplo, se idade estiver fortemente relacionada a turma, uma amostragem estratificada por turma também será mais representativa em termos de idade. Como aspirações ocupacionais estão relacionadas a sexo, uma amostra estratificada por sexo será mais representativa em termos das aspirações ocupacionais.

A escolha das variáveis de estratificação depende em geral de quais variáveis estão disponíveis. Sexo pode muitas vezes ser determinado através de uma lista de nomes. As listas universitárias são tipicamente organizadas por turma. As listas de professores podem indicar sua lotação por departamento. Arquivos de organizações governamentais podem estar organizados por região geográfica. Listas de registros de eleitores estão organizadas por seções e zonas eleitorais.

No entanto, ao selecionar as variáveis de estratificação das listas disponíveis, você deve se preocupar sobretudo com as que se relacionam, presumivelmente, com as variáveis que você quer representar com precisão. Como sexo está relacionado a muitas variáveis e muitas vezes está disponível para estratificação, acaba sendo usado freqüentemente. Escolaridade está relacionada a muitas variáveis, mas freqüentemente não está disponível para estratificação. Localização geográfica numa cidade, estado ou nação está relacionada a muitas coisas. Dentro de uma cidade, a estratificação por lugar geográfico, em geral, aumenta a representatividade em áreas como

classe social e grupo étnico. Dentro de uma nação, aumenta a representatividade em uma vasta gama de atitudes, bem como na classe social e na etnia.

Os métodos de estratificação na amostragem variam. Quando se trabalha com uma lista simples de todos os elementos na população, dois métodos predominam. Primeiro, pode-se agrupar os elementos da população em grupos discretos com base nas variáveis de estratificação usadas. Com base na proporção relativa de um grupo na população, seleciona-se — aleatória ou sistematicamente — um número de elementos daquele grupo na mesma proporção do tamanho de amostra desejado. Por exemplo, se alunos da turma A com média 4,0 são 1% da população estudantil e você quer uma amostra de 1.000 estudantes, selecione 10 estudantes dos alunos da turma A com média 4,0.

Como método alternativo, pode-se agrupar estudantes como descrito acima e pôr esses vários grupos juntos numa lista contínua, começando com todos os alunos da turma A com média 4,0 e terminando, digamos, com todas as alunas da turma Z com média 1,0 ou menos. Seleciona-se então uma amostra sistemática, com início aleatório, de toda a lista. Dada a organização de lista, uma amostra sistemática selecionaria números adequados (dentro de uma margem de erro de 1 ou 2) de cada subgrupo. (*Nota:* Uma amostra aleatória simples retirada de tal lista composta cancelaria a estratificação.)

O efeito da estratificação é garantir a representação adequada das variáveis de estratificação para a aumentar a representação de outras variáveis relacionadas a elas. Portanto, tomada como um todo, uma amostra estratificada tem possibilidade de ser mais representativa num certo número de variáveis do que no caso de uma amostra aleatória simples. Embora a amostra aleatória simples ainda seja vista como algo sagrado, deve estar claro que muitas vezes há coisa melhor.

Estratificação Implícita na Amostragem Sistemática

Mencionei antes que a amostragem sistemática pode, em certas condições, ser mais precisa do que a amostragem aleatória simples. É o caso sempre que o arranjo da lista cria uma “estratificação implícita”. Como assinalei, se a lista de estudantes universitários está organizada por turma, uma amostra sistemática produzirá a estratificação por período, ao contrário de uma amostra aleatória simples. Outros arranjos típicos de elementos em listas podem ter o mesmo efeito.

Se uma lista de nomes que compõe uma moldura de amostragem está em ordem alfabética, ela é um pouco estratificada por origem étnica. Por exemplo, todos os McTavishes estão reunidos, bem como os Lees, Wongs, Yamamuras, Schmidts, Whitehalls, Weinstains, Gonzaleses etc. Na medida em que qualquer destes subgrupos forma um subconjunto substancial da população total, estará adequadamente representado numa amostra sistemática retirada de uma lista alfabética.

Num estudo de estudantes da Universidade do Havaí, depois de estratificados por turma, os estudantes foram organizados pelos números de identidade estudantil. No entanto, estes números eram seus números de Seguro Social. Os primeiros três dígitos do número de Seguro Social indicam o estado onde foi emitido. Portanto, dentro de cada turma, os estudantes estavam organizados pelo estado onde havia sido emitido seu número de Seguro Social, fornecendo uma estratificação aproximada por origem geográfica.

Uma lista ordenada de elementos pode ser mais útil do que uma lista não ordenada, aleatória. Enfatizo este ponto para combater a crença infeliz de que listas devem ser randomizadas antes de ocorrer a amostragem sistemática. A lista só deve ser reordenada se o arranjo apresenta o problema da *periodicidade* já discutido.

Amostragem por Conglomerados em Múltiplas Etapas, Geral

As quatro seções precedentes lidaram com procedimentos razoavelmente simples de amostragem a partir de listas de elementos, que representam a situação ideal. Entretanto, infelizmente muita pesquisa social interessante requer selecionar amostras de populações que não podem ser facilmente listadas para fins de amostragem, tais como a população de uma cidade, de um estado ou da nação, todos os estudantes universitários nos Estados Unidos etc. Nestes casos, é necessário criar e executar um desenho de amostra mais complexo. Este desenho tipicamente envolve a amostragem inicial de *grupos* de elementos, conhecidos como *conglomerados* [*clusters*], seguida pela seleção de elementos de cada conglomerado selecionado. As variedades e procedimentos da *amostragem por conglomerados em múltiplas etapas* são dissecadas em detalhe nos exemplos de amostragem no Capítulo 6. Entretanto, vamos delinear aqui.

Pode-se fazer amostragem por conglomerados quando não é possível ou prático compilar uma lista exaustiva dos elementos da população alvo. Todos os membros de igrejas nos Estados Unidos exemplificam uma população assim.

Contudo, muitas vezes os elementos da população já estão agrupados em subpopulações, e uma lista destas subpopulações existe ou pode ser criada de forma prática. Por exemplo, membros de igrejas nos Estados Unidos pertencem a igrejas específicas, e é possível descobrir ou criar uma lista dessas igrejas. Seguindo um formato de amostragem por conglomerado, a lista de igrejas é amostrada numa das maneiras descritas nas seções anteriores (por exemplo, uma amostra estratificada sistemática). Em seguida, obtêm-se listas de membros de cada uma das igrejas selecionadas e amostra-se cada uma das listas, de forma a ter amostras de membros de igreja para estudo.⁶

Outra situação típica é a amostragem em áreas habitadas, como uma cidade. Embora não haja uma lista única da população de uma cidade, os cidadãos residem em quarteirões separados ou zonas de recenseamento. Portanto, é possível começar selecionando uma amostra de quarteirões, criar uma lista de pessoas morando nos quarteirões selecionados e, finalmente, subamostrar pessoas em cada quarteirão.

Num desenho mais complexo, você pode amostrar os quarteirões, listar as residências em cada quarteirão selecionado, amostrá-las, listar as pessoas residindo em cada uma e, finalmente, amostrar as pessoas em cada residência selecionada. Este desenho de amostragem em múltiplas etapas leva à seleção final de uma amostra de indivíduos sem requerer a listagem inicial de todos os indivíduos que compõem a população da cidade.

Portanto, também na amostragem por conglomerados em múltiplas etapas repetem-se dois passos básicos: listar e amostrar. A lista de unidades primárias de amostragem (igrejas ou quarteirões, por exemplo) é compilada e talvez estratificada para amostragem. Uma amostra dessas unidades é selecionada. As unidades primárias de amostragem selecionadas são, então, listadas e talvez estratificadas. A lista de unidades secundárias de amostragem é amostrada, e assim por diante. Os métodos de listagem e amostragem são apresentadas detalhadamente nos exemplos do Capítulo 6.

A amostragem por conglomerados é altamente recomendada pela eficiência, embora à custa da precisão. Enquanto uma amostra aleatória simples extraída de uma lista populacional está sujeita a um único erro amostral, uma amostra por conglomerados em duas etapas está sujeita a dois. Primeiro, a amostra inicial dos conglomerados representa a população de conglomerados somente dentro de uma faixa de erro amostral. Segundo, a amostra de elementos extraída de um conglomerado representa todos os elementos dele apenas dentro de uma faixa de erro amostral. Assim, por exemplo, há o risco de

selecionar uma amostra de bairros desproporcionalmente ricos, bem como uma amostra de residências desproporcionalmente ricas dentro destes bairros. A melhor solução ao problema está no número de conglomerados selecionados no início e no número de elementos selecionados em cada conglomerado.

Em geral, estamos restritos a um tamanho máximo de amostra; por exemplo, podemos ter o limite de 2.000 entrevistas numa cidade. Dada esta limitação maior, temos várias opções para desenhar a amostra por conglomerado. Num extremo, podemos escolher um conglomerado e selecionar 2.000 elementos dentro dele, em outro, escolher 2.000 conglomerados e selecionar um elemento em cada um. Naturalmente, nenhum é aconselhável, mas há muitas opções entre estes extremos. Felizmente, a lógica das distribuições amostrais provê uma diretriz geral a seguir.

Lembre que o erro amostral é reduzido por dois fatores: o aumento do tamanho da amostra e uma maior homogeneidade dos elementos a serem amostrados. Estes fatores operam em cada nível do desenho da amostra em múltiplas etapas. Uma amostra por conglomerados representará melhor todos os conglomerados se um grande número for selecionado e se todos eles forem bastante semelhantes. Uma amostra de elementos representará melhor todos os elementos de um conglomerado se um grande número de elementos for selecionado do conglomerado, e se todos os elementos forem bastante semelhantes.

Entretanto, dado um tamanho máximo da amostra, se o número de conglomerados aumentar, o número de elementos num conglomerado deve diminuir. Neste sentido, a representatividade dos conglomerados aumenta à custa da menor representação dos elementos de cada um, e vice-versa. Felizmente, podemos usar o fator da homogeneidade para aliviar esta dificuldade.

Os elementos de um conglomerado natural dentro de uma população são tipicamente mais homogêneos do que todos os elementos da população total. Os membros de uma igreja são mais semelhantes entre si do que todos os membros de igrejas; da mesma forma, os moradores de um bairro são mais semelhantes entre si do que todos os moradores de uma cidade inteira. Portanto, podem ser necessários relativamente poucos elementos para representar adequadamente um conglomerado natural, enquanto um número maior de conglomerados pode ser necessário para representar adequadamente a diversidade dos conglomerados. Vê-se isto com máxima clareza no caso extremo de conglomerados muito diferentes entre si, compostos de elementos

exatamente idênticos em cada um. Nesta situação, um número grande de conglomerados representaria adequadamente a variedade entre os conglomerados, bastando apenas um elemento de cada um para representar adequadamente todos os outros. Embora esta situação extrema nunca ocorra, ela está mais próxima da verdade mais vezes do que o seu oposto — conglomerados idênticos compostos de elementos muito divergentes.

Portanto, a diretriz geral para o desenho de conglomerados é maximizar o número de conglomerados selecionados, diminuindo o número de elementos de cada um. Mas esta diretriz científica deve ser contrastada com as restrições administrativas. A eficiência da amostragem por conglomerados funda-se na capacidade de minimizar a listagem de elementos da população. Selecionando inicialmente conglomerados, buscamos listar apenas os elementos dos conglomerados selecionados, em vez de todos os elementos da população inteira. Aumentar o número de conglomerados vai diretamente de encontro a este fator de eficiência. Um número menor de conglomerados pode ser listado mais rápida e economicamente do que um número maior. (Lembre-se de que todos os elementos de um conglomerado selecionado devem ser listados, mesmo que poucos vão ser escolhidos na amostra.)

O desenho final da amostra refletirá estas duas restrições. Provavelmente, você vai selecionar o maior número de conglomerados que seu orçamento permitir. Mas, para não deixar esta questão aberta demais, apresentamos uma regra simples. Pesquisadores populacionais convencionalmente selecionam cinco residências por zona ou bairro de recenseamento. Querendo entrevistar um total de 2.000 residências, você seleciona 400 bairros com cinco entrevistas de residências em cada um. Voltaremos a esta regra em exemplos posteriores de desenhos de amostra.

Antes de passarmos a procedimentos mais detalhados da amostragem por conglomerados, cabe lembrar que este método quase inevitavelmente envolve perda de precisão. A maneira como ocorre esta perda de precisão é complexa. Primeiro, um desenho de amostra em etapas múltiplas está sujeito a erros amostrais em cada uma das etapas. Como o tamanho da amostra é necessariamente menor, a cada etapa, do que o tamanho total da amostra, o erro amostral em cada etapa será maior do que no caso de uma amostra aleatória de elementos em etapa única. Segundo, o erro amostral é estimado com base na variância observada entre os elementos da amostra. Quando estes elementos são retirados

de conglomerados relativamente homogêneos, as estimativas do erro de amostragem serão otimistas demais, devendo ser corrigidas à luz do desenho de amostras por conglomerado. (Este problema será discutido em detalhes na consideração posterior sobre análise univariada.)

Amostragem de Aglomerados em Múltiplas Etapas, Estratificação

Até agora, discutimos a amostragem por conglomerados como se fosse selecionada uma amostra aleatória simples a cada etapa do desenho. Na prática, podemos usar as técnicas de estratificação discutidas antes para refinar e melhorar a amostra selecionada.

Exemplos posteriores detalharão métodos possíveis de estratificação; por ora, observemos que nossas opções básicas são essencialmente as mesmas que temos na amostragem de etapa única a partir de uma lista. Por exemplo, ao selecionar uma amostra nacional de igrejas, você pode, no começo, estratificar sua lista de igrejas por denominação, região geográfica, tamanho, localização urbana/rural, e talvez alguma medida da classe social. Informação do censo pode ser usada por pesquisadores de população para estratificar bairros de recenseamento em termos de composição étnica, classe social, valores das propriedades, qualidade de estruturas, natureza da propriedade das construções e tamanho.

Assim que as unidades primárias de amostragem (igrejas ou bairros, nos exemplos) estiverem agrupadas de acordo com as variáveis de estratificação relevantes disponíveis, pode-se usar tanto técnicas de amostragem aleatória simples ou a amostragem sistemática para extrair a amostra. Você pode selecionar um número especificado de unidades de cada grupo ou "estrato", ou arranjar os conglomerados estratificados numa lista contínua, amostrando sistematicamente esta lista. O erro amostral será tão mais reduzido, nesta etapa, quanto mais os conglomerados estiverem combinados em estratos homogêneos. Como antes, o objetivo primordial da estratificação é a homogeneidade.

Não há razão para não se poder estratificar em cada nível de amostragem. Os elementos listados de um aglomerado selecionado podem ser estratificados antes da próxima etapa de amostragem. No entanto, isso não é feito comumente. (Lembre a suposição sobre a homogeneidade relativa dentro dos aglomerados.)

Amostragem de Probabilidade Proporcional ao Tamanho (PPT)

Até agora, falei genericamente sobre a designação de elementos da amostra a conglomerados selecionados em termos de quantos conglomerados devem ser selecionados e quantos elementos deve haver em cada um. Esta seção discute em maior detalhe algumas opções.

Para garantir a seleção de uma amostra representativa de elementos, você deve dar a cada elemento da população total uma oportunidade igual de ser selecionada. A forma mais simples de fazer isso numa amostra por conglomerados é dar a cada um a mesma chance de seleção e selecionar uma determinada proporção de elementos de cada conglomerado selecionado. Com uma população de 100.000 elementos agrupados em 1.000 conglomerados (de tamanhos variados), pretendendo amostrar 1.000 elementos, você seleciona um décimo dos conglomerados (100) com igual probabilidade e subseleciona um décimo dos elementos de cada conglomerado escolhido. Assim, cerca de 1.000 elementos são selecionados e cada elemento na população tem a mesma ($1/10 \times 1/10 = 1/100$) probabilidade de seleção. Esta técnica de seleção de amostras, embora a mais clara e simples, não é a mais eficiente.

A maioria das amostragens por conglomerados envolve conglomerados de tamanhos muito diversos (em número de elementos). O pesquisador de religiões encontra igrejas muito grandes e muito pequenas. O pesquisador populacional acha bairros muito populosos e outros nem tanto. Além disso, os conglomerados pequenos muitas vezes são mais numerosos do que os grandes, embora os grandes possam responder por uma proporção maior da população total. Assim, alguns poucos grandes bairros podem conter uma proporção grande da população de uma cidade, enquanto o número maior de bairros menores de fato contém apenas uma pequena parcela da população.

Selecionar conglomerados com igual probabilidade, retirando uma proporção fixa de elementos dos conglomerados selecionados, resulta no seguinte: (1) Seleciona-se um número relativamente pequeno de conglomerados grandes na primeira etapa da amostragem. (2) Os elementos selecionados para representar todos os elementos dos grandes aglomerados são retirados de muito poucos desses conglomerados. Na situação extrema, toda a população da cidade residindo em dez grandes bairros poderia ser representada pelas pessoas vivendo em apenas um deles.

A seção sobre a amostragem por conglomerados discutiu a maior eficiência inerente à seleção de muitos conglomerados, com poucos elementos sendo retirados de cada um. Este princípio é posto em prática através do método de *amostragem de probabilidade proporcional ao tamanho* (PPT), que permite selecionar maior número de conglomerados, garante a representação dos elementos de conglomerados grandes e dá a cada elemento na população igual chance de seleção.

Na primeira etapa de amostragem, é dada a cada conglomerado oportunidade de seleção proporcional ao seu tamanho (em número de elementos). Grandes conglomerados têm maior chance de seleção do que os menores. Mas, na segunda etapa de amostragem, o mesmo *número* de elementos é escolhido de cada conglomerado selecionado. O efeito destes dois procedimentos é igualar as probabilidades últimas de seleção de todos os elementos, já que elementos em conglomerados maiores têm uma oportunidade de seleção menor *dentro* do seu aglomerado do que elementos em conglomerados menores. Por exemplo, um quarteirão com 100 residências terá dez vezes mais chance de seleção do que um quarteirão com apenas 10. Se ambos os quarteirões forem selecionados e o mesmo número de residências for selecionado de cada quarteirão, as casas no quarteirão maior terão dez vezes menos chance de seleção do que as casas no quarteirão menor. A fórmula abaixo indica a probabilidade de seleção de um elemento num desenho de amostragem PPT.

$$\left[\begin{array}{c} \text{probabilidade} \\ \text{do elemento} \end{array} \right] = \left[\begin{array}{c} \text{número de} \\ \text{aglomerados} \\ \text{selecionados} \end{array} \right] \times \left[\begin{array}{c} \text{parte do aglomerado} \\ \text{tamanho do aglomerado} \end{array} \right] \times \left[\begin{array}{c} \text{elementos} \\ \text{selecionados} \\ \text{por aglomerado} \\ \text{tamanho do} \\ \text{aglomerado} \end{array} \right]$$

Se 100 conglomerados são selecionados e 10 elementos escolhidos de cada um, a partir de uma população total de 100.000, a probabilidade geral de seleção de cada elemento será de 1.000/100.000, ou 1/100. Um conglomerado com 100 elementos tem uma probabilidade de seleção igual a 100 (conglomerados a serem selecionados) vezes 100/100.000 (tamanho do conglomerado/tamanho da população), ou 1/10; cada elemento tem uma chance de 10/100 (elementos por conglomerado/tamanho do conglomerado), ou 1/10, de ser selecionado *dentro* daquele conglomerado; a chance total de seleção do elemento, neste caso, é 1/10 vezes 1/10, ou 1/100. Um conglomerado com 10 elementos tem uma probabilidade de seleção de 100 vezes 10/100.000, ou 1/100, mas a chance de seleção de cada elemento dentro do conglomerado é 10/10, ou 1, fazendo com que a chance total de seleção seja igual a 1/100.

Independente do número de elementos num conglomerado, cada elemento tem, enfim, a mesma probabilidade de seleção. Pode-se ver isto mais claramente na fórmula, notando que o tamanho do conglomerado aparece no numerador e no denominador, cancelando-se mutuamente. Portanto, a probabilidade de seleção se torna o número de conglomerados a serem escolhidos vezes o número de elementos a serem escolhidos de cada conglomerado selecionado, dividido pelo tamanho da população. Isso, é claro, é o tamanho da amostra dividido pelo tamanho da população.

Costuma-se fazer duas modificações neste desenho de amostra PPT. Primeiro, podemos achar necessário representar conglomerados *muito* grandes na amostra; por exemplo, podemos querer garantir que todos os quarteirões (ou igrejas) com mais de, digamos, 1.000 elementos sejam incluídos na amostra. Neste caso, podemos selecionar todos estes conglomerados logo de início (com uma probabilidade de 1,0), devendo-se, então, atribuir aos elementos destes agrupamentos uma probabilidade igual à razão amostral geral. No exemplo anterior, 1/100 dos elementos residindo em cada um dos grandes quarteirões poderiam ser selecionados.

A segunda modificação é relativa a pequenos conglomerados. Se um número padrão de elementos for ser selecionado de cada conglomerado escolhido, conglomerados com menos elementos do que o número padrão representam um problema. Se a meta é dez residências em cada quarteirão selecionado, o que fazer com quarteirões com cinco residências? A solução habitual é combinar pequenos conglomerados, de forma que cada combinação tenha pelo menos o número padrão a ser selecionado. (Se os conglomerados estiverem estratificados, as combinações devem ser feitas dentro de cada estrato. Da mesma forma, pequenos conglomerados podem ser agregados a outros maiores, caso este procedimento garanta maior homogeneidade dentro do conglomerado resultante). O exemplo de uma amostra por conglomerados por área no Capítulo 6 demonstra a necessidade de dar esse passo, para garantir que se levem em conta blocos onde se acredita não haver residências.

A título de resumo, a amostragem por conglomerados é um método de amostragem difícil, mas importante, adequado sempre que for impossível compilar uma lista de todos os elementos da população em estudo. A discussão foi abstrata, mas o exemplo dado no Capítulo 6 oferece um quadro mais claro dos passos dados, de fato, numa seleção complexa de amostras.

Amostragem e Ponderação Desproporcionais

Em última análise, uma amostra probabilística é representativa de uma população se todos os elementos da população tiverem uma chance igual de seleção. Portanto, nas discussões precedentes, assinalamos que os vários procedimentos de amostragem resultam numa chance igual de seleção — mesmo que a probabilidade última de seleção seja o produto de várias probabilidades parciais.

Entretanto, mais genericamente, uma amostra probabilística é aquela em que cada elemento da população tem uma probabilidade *não-zero conhecida* de seleção, mesmo que elementos diferentes tenham probabilidades diferentes. Caso se tenha usado procedimentos probabilísticos controlados de amostragem, qualquer uma dessas amostras será representativa da população da qual é retirada se cada elemento da amostra receber um peso igual ao inverso da sua probabilidade de seleção. Portanto, no caso de todos os elementos da amostra terem a mesma chance de seleção, cada um recebe o mesmo peso: 1. (É chamada de amostra “autoponderada”.)

Às vezes, é correto fazer amostragem e *ponderação* desproporcionais. Por exemplo, você pode amostrar subpopulações desproporcionalmente para garantir número suficiente de casos de cada subpopulação para análise.

Suponha que você está planejando um *survey* por amostragem de 1.000 residências numa cidade e que a cidade tem uma área decadente onde mora um quarto da população total, e você quer analisar detalhadamente essa área. Uma amostra representativa da cidade lhe daria aproximadamente 250 residências nessa região e 750 no resto da cidade. Mas, 250 casos podem não ser suficientes para a análise que você planeja. Uma solução é selecionar o mesmo número de residências (500) da região decadente que o do resto da cidade. Então, as residências da região decadente teriam uma chance desproporcionalmente maior de seleção em relação às de outras partes da cidade.

Contanto que você analise as duas amostras de áreas separadamente ou comparativamente, não precisa se preocupar com a amostragem diferencial. Mas, se quiser combinar as duas amostras para criar um quadro composto da cidade toda, leve a amostragem desproporcional em consideração através de um procedimento de ponderação.

Eis uma solução simples e intuitiva. Como a região decadente tem um quarto das residências da cidade, e a região

não-decadente, os demais três quartos, as 500 residências selecionadas da região não-decadente têm representação na cidade três vezes maior do que as 500 selecionadas da região decadente. Portanto, a solução de ponderação mais simples é contar cada residência selecionada na região não-decadente como *três* residências. Este procedimento dá a impressão de que você teria entrevistado 1.500 moradias na região não-decadente da cidade, que é o número que você teria selecionado e entrevistado se quisesse 500 moradias na região decadente e tivesse recursos para uma amostra total de 2.000.

Às vezes é possível fazer cópias dos arquivos de dados em certos casos — as residências da região não-decadente, neste exemplo. No entanto, alguns programas de análise de dados permitem especificar os pesos a serem atribuídos aos vários casos, e o computador os leva em consideração ao calcular as estatísticas na análise.

A amostragem desproporcional e os procedimentos de ponderação são às vezes usados em situações que envolvem os erros e as aproximações muitas vezes inerentes à amostragem complexa em múltiplas etapas. Isto será discutido em mais detalhes no exemplo de amostra por conglomerados por área no Capítulo 6, mas convém mencionar algumas das condições nas quais a ponderação é freqüentemente necessária.

1. Na amostragem estratificada por conglomerados, um determinado número de conglomerados pode ser selecionado de cada estrato, embora varie o tamanho dos diversos estratos. A ponderação diferencial pode ser usada para compensar aquelas variações.
2. Um certo conglomerado pode ser selecionado num desenho de amostra PPT com base no tamanho esperado ou estimado, mas uma investigação de campo posterior pode indicar que a estimativa inicial estava errada. Portanto, o conglomerado terá recebido uma chance de seleção desproporcionalmente alta ou baixa; a ponderação pode ser usada para compensar este erro.
3. Um desenho de amostra pode pedir a seleção de um décimo dos elementos de um conglomerado, mas este pode conter apenas cinquenta e dois elementos, apenas cinco dos quais sendo selecionados para estudo. Pode-se usar a ponderação para compensar o elemento 0,2 que, logicamente, não tinha como ser selecionado.

4. Dez elementos de um conglomerado podem ter sido selecionados, mas dois não tiveram como ser estudados (por exemplo, recusaram a entrevista). Pressupondo homogeneidade dentro do conglomerado, você pode atribuir peso 1,25 a cada um dos elementos estudados para compensar os dois que não o foram.

Todos estes usos de ponderação são ilustrados no exemplo final de amostragem do Capítulo 6. Entretanto, com a exceção do caso 4 acima, você pode derivar seu próprio procedimento de ponderação, determinando cuidadosamente a probabilidade de seleção, passo a passo, para cada elemento da amostra, e atribuindo um peso igual ao inverso daquela probabilidade. Três outros aspectos da ponderação devem ser discutidos antes de passarmos a outro tema.

Graus de Precisão na Ponderação

Em qualquer desenho complexo de amostra, há diversas opções em relação à ponderação, em conexão com a amostragem intencional e/ou inadvertidamente desproporcional. Podemos computar pesos para cada elemento até várias casas decimais, ou atribuir pesos aproximados para dar conta apenas das instâncias mais grosseiras de amostragem desproporcional. No caso anterior da cidade com a região suburbana sobreamostrada, é improvável que a população daquela região seja exatamente um quarto da população da cidade; pode de fato ser 0,25001, 0,2600 ou 0,2816 da população total. No primeiro caso, provavelmente escolheremos aplicar a ponderação geral aproximada de casos descrita, se não houver outra amostragem desproporcional envolvida noutras etapas do desenho da amostra. Talvez também façamos o mesmo no segundo e terceiro casos. A precisão que buscamos na ponderação deve ser comensurável com a precisão que queremos nos nossos achados. Se nossos objetivos de pesquisa podem tolerar erros de poucos pontos percentuais, provavelmente não perderemos tempo e esforço fazendo ponderações exatas. Além disso, ao decidirmos o grau de precisão requerido, devemos levar em conta o grau de erro a ser esperado da distribuição normal de amostragem mais todos os diversos tipos de erro não-amostrais.

Em última análise, não há uma diretriz firme a seguir quando se determina a precisão a ser buscada na ponderação. Como em vários outros aspectos do desenho de *surveys*, desfrutamos de considerável latitude. Ao mesmo tempo, porém, devemos lembrar-nos da nossa decisão de ponderação ao informarmos

os resultados. Não devemos empregar um procedimento de ponderação aproximada e depois sugerirmos que os achados são precisos dentro de uma faixa minúscula de erro.

Métodos de Ponderação

Tendo delineado as preocupações científicas quanto a determinar o grau de precisão desejado na ponderação, devemos assinalar que a escolha muitas vezes é feita com base nos métodos disponíveis para ponderar. Os três métodos básicos para a ponderação são:

1. Para a ponderação aproximada de amostras extraídas de subpopulações, tabelas com pesos podem ser construídas das tabelas sem pesos, para cada uma das subamostras. No exemplo acima, você pode criar uma tabela simples de distribuições para a amostra suburbana e para a amostra não suburbana separadamente, triplicar o número de casos em cada célula da tabela não suburbana, somar os casos nas duas tabelas e computar as porcentagens para a tabela composta.

2. Para uma ponderação mais extensa e rápida, embora ainda aproximada, os conjuntos de dados para certos casos podem ser clonados. No exemplo anterior, podem ser feitas duas cópias adicionais para cada residência suburbana (totalizando três), o conjunto ampliado de casos não suburbanos podendo então ser combinado com os casos suburbanos, e o conjunto ampliado de dados pode ser analisado como se três vezes mais casas não suburbanas tivessem sido estudadas.

3. Muitos programas de computador para análise de dados são desenhados para atribuir um peso exato para cada caso no arquivo de dados original. Este método é o único apropriado para a ponderação refinada, porque é impossível clonar frações de casos, com algum significado.

Como em outras instâncias, também na ponderação as preocupações científicas em geral submetem-se a razões práticas. Se a análise for conduzida, por exemplo, através de cartões IBM, a ponderação deve, necessariamente, ser aproximada em vez de precisa.

Ponderação e Inferência Estatística

Advirto que os procedimentos de ponderação descritos têm efeitos sérios na maioria dos cálculos relacionados à inferência estatística. Se suas finalidades de pesquisa exigem inferências estatísticas precisas (por exemplo, estimativas de população), feitas com base em dados cuidadosamente ponderados, você deve consultar uma fonte especial sobre este assunto⁷ ou, melhor ainda, consultar um estatístico de amostragem *antes da amostragem ser desenhada*.

Panorama da Amostragem Probabilística

A longa e detalhada discussão precedente foi dedicada ao método chave de amostragem usado na pesquisa de *survey*: amostragem probabilística. Em cada uma das variações examinadas, vimos que os elementos são escolhidos de uma população, com base na seleção aleatória e com probabilidades não-zero conhecidas.

Dependendo da situação de campo, a amostragem probabilística pode ser muito simples ou extremamente difícil, demorada e cara. No entanto, independentemente da situação, continua sendo o método mais eficiente para a seleção de elementos para estudo por duas razões. Primeiro, a amostragem probabilística evita o VIÉS consciente ou inconsciente na seleção de elementos por parte do pesquisador. Se todos os elementos da população têm chance igual (ou desigual, com ponderação posterior) de seleção, há uma boa chance de a amostra assim selecionada representar bastante bem a população. Segundo, a amostragem probabilística permite estimativas de erro amostral. Embora nenhuma amostragem probabilística seja perfeitamente representativa em todos os aspectos, métodos controlados de seleção permitem estimar o grau de erro esperado.

Tendo discutido a amostragem de probabilidade em algum detalhe, examinemos agora, mais sucintamente, alguns métodos comuns de amostragem não-probabilística.

Amostragem Não-Probabilística

Apesar da superioridade dos métodos de amostragem probabilísticos na pesquisa de *survey*, às vezes usam-se métodos não-probabilísticos em seu lugar, em geral em situações em que a amostragem probabilística seria dispendiosa demais

e/ou quando a representatividade exata não é necessária. Os métodos principais da amostragem não-probabilística são os seguintes.

Amostragem Intencional ou por Julgamento

Ocasionalmente, pode-se selecionar a amostra baseado no próprio conhecimento da população e dos seus elementos, e da natureza das metas de pesquisa. Este método de amostragem é às vezes chamado *amostragem intencional* ou *por julgamento*. Principalmente no desenho inicial do questionário, você pode querer selecionar maior variedade de entrevistados para testar a aplicabilidade geral das perguntas. Embora os achados do *survey* não representem qualquer população significativa, o teste pode efetivamente descobrir algum defeito peculiar no seu instrumento de pesquisa. No entanto, deve-se referir a esta situação como um pré-teste e não um *survey* propriamente.

Algumas vezes, você quer estudar um pequeno subconjunto de uma população maior, onde muitos membros do subconjunto são facilmente identificados, mas cuja enumeração total seria quase impossível. Por exemplo, estudar a liderança de um movimento de protesto estudantil; muitos dos líderes são facilmente visíveis, mas não é viável definir e amostrar todos eles. Ao estudar todos ou uma amostra dos líderes mais visíveis, você pode coletar dados suficientes para seus propósitos.

Num desenho de amostra em etapas múltiplas, você pode querer comparar estudantes liberais e conservadores. Uma vez que, provavelmente, não é possível enumerar e amostrar todos os estudantes, você pode optar por amostrar as filiações dos Jovens Democratas e dos Jovens Republicanos. Embora tal desenho de amostra não forneça uma boa descrição dos estudantes liberais ou dos conservadores como um todo, pode ser suficiente para propósitos de comparação.

A *amostragem de zonas eleitorais selecionadas* para pesquisas de opinião política é um processo bem refinado de julgamento. Com base em resultados eleitorais anteriores em alguma área (cidade, estado, nação), você seleciona intencionalmente um grupo de distritos ou zonas eleitorais que, combinados, produzem resultados similares aos de toda a região. Então, em pesquisas posteriores, você só seleciona suas amostras naquelas zonas eleitorais. Naturalmente, a teoria diz que as zonas eleitorais selecionadas oferecem um corte de todo o eleitorado.

Cada vez que há uma eleição, permitindo avaliar a adequação do seu grupo de zonas eleitorais, você deve fazer revisões, adições ou retiradas. A meta é atualizar o grupo de zonas eleitorais, garantindo a boa representação de todas elas. Para ser efetiva, a amostragem de zonas eleitorais requer grande conhecimento político. Deve haver intimidade com a história política e social da área considerada, de modo a basear a seleção de zonas eleitorais num palpite *com conhecimento* quanto à representatividade. Ademais, este sistema de amostragem requer retroalimentação contínua para ser efetiva. Deve-se poder conduzir pesquisas freqüentes e fazer validações eleitorais periódicas.

Amostragem por Cotas

Como disse antes, a amostragem por cotas começa com uma matriz descrevendo as características da população alvo. Por exemplo, você deve saber qual proporção da população é masculina e qual é feminina; para cada sexo, saber qual é a proporção nas várias categorias de idade, e assim por diante. Ao estabelecer uma amostra nacional por cotas, você deve saber qual proporção da população nacional é, digamos, urbana, do leste, masculina, abaixo de vinte e cinco anos, branca, da classe operária, e outros semelhantes, bem como todas as outras permutações nessa matriz.

Após a criação da matriz, com uma proporção relativa atribuída a cada célula, colhem-se dados de pessoas com todas as características de uma dada célula. Todas as pessoas nesta célula, então, recebem um peso apropriado à sua porção na população total. Quando todos os elementos da amostra estiverem ponderados assim, os dados como um todo devem resultar numa representação razoável da população total.

Há vários problemas inerentes à amostragem por cotas. Primeiro, a moldura das cotas (a proporção que as diferentes células representam) deve ser precisa, e muitas vezes é difícil obter informação atualizada para este fim. A falha do Gallup em prever a vitória de Truman em 1948 deveu-se em parte a este problema.

Segundo, pode haver vieses na seleção dos elementos da amostra de uma célula, mesmo que se estime com precisão sua proporção na população. Um entrevistador instruído a entrevistar cinco pessoas que se encaixam num dado conjunto complexo de características pode, ainda assim, evitar pessoas que moram nos andares de cima de um prédio com sete andares de escadas, ou pessoas em habitações muito precárias e/ou com cães ferozes.

Muitas vezes se usa uma variante da amostragem por cotas em conexão aos *grupos focais* na pesquisa de mercado. Nesta técnica, grupos de aproximadamente doze pessoas são, tipicamente, reunidos para discussões focalizadas sobre produtos ou comerciais. Recomenda-se reunir vários grupos de tipos diferentes de pessoas ou misturas específicas de pessoas.

Confiança em Sujeitos Disponíveis

Pesquisadores universitários freqüentemente fazem *surveys* com estudantes matriculados em grandes salas de aula. A facilidade e baixo custo deste método explica sua popularidade, mas estes *surveys* raramente produzem dados com algum valor genérico. Embora sirva à finalidade de pré-testar um questionário, este método de amostragem não deve ser usado para um estudo pretendendo descrever os alunos como um todo.

Parar pessoas numa esquina ou outro lugar quase nunca é um método de amostragem adequado, embora seja empregado com demasiada freqüência, como no caso de pesquisas de mercado em estacionamentos de supermercados. Este *survey* só se justifica se você quiser estudar as características de pessoas passando naquele ponto em particular em horas determinadas.

O termo “entrevista por interceptação em *shoppings*” às vezes é usado em referência a um *survey* no qual consumidores andando em *shoppings* são interceptados e solicitados a participar. Obviamente, é concebível que tais respondentes potenciais seriam apropriados a alguns *surveys* específicos — um estudo de consumidores daquele *shopping*, por exemplo. Entretanto, nestes casos, outro problema de amostragem é que a maior parte daqueles que são abordados se recusa a participar. Mas, num estudo num *shopping* em Chicago, Hornik e Ellis observaram que as taxas de resposta cresciam quando os entrevistadores olhavam e tocavam os entrevistados ao pedir sua cooperação, especialmente se a entrevistadora era mulher.⁸

Mencionei esta pesquisa para indicar que, embora os métodos de amostragem probabilística convencional sejam quase sempre preferíveis na pesquisa de *survey*, há momentos em que outros métodos podem ser preferíveis ou necessários. Mas escolher um modelo de amostragem não-probabilístico jamais deve ser desculpa para a preguiça. Pesquisadores de *survey* devem encontrar formas de providenciar uma amostra representativa da população que querem estudar.

Usos Não-Survey dos Métodos de Amostragem

Nesta discussão sobre a lógica e os métodos de amostragem, focalizamos a pesquisa de *survey* — seleção de elementos de uma população para entrevista ou questionários auto-administrados. No entanto, a lógica básica da amostragem dá à discussão precedente um valor mais geral para o pesquisador social que usa outros métodos de coleta de dados. Fecharemos o capítulo com breves comentários sobre os usos não-*survey* dos métodos de amostragem.

Análise de Conteúdo

Um analista de conteúdo codifica e analisa documentos para fazer afirmações descritivas ou explicativas sobre a literatura composta pelos documentos, seu(s) autor(es) e/ou o meio social do qual fazem parte os documentos. Ele pode analisar os jornais de um país, os trabalhos de ficção de um romancista, a linguagem de projetos legislativos, e assim por diante.

Muitas vezes o volume de documentos para analisar é grande demais para uma cobertura completa. Neste caso, as técnicas de amostragem discutidas acima podem ser facilmente adaptadas. As unidades de amostragem podem ser palavras individuais, frases, parágrafos, artigos, livros etc. As unidades de amostragem podem ser estratificadas de maneira apropriada, e amostras aleatórias, sistemáticas ou mesmo por conglomerados ser selecionadas.

Experimentos de Laboratórios

Sujeitos para experiências de laboratórios muitas vezes são selecionados entre voluntários que respondem a um anúncio. Ocasionalmente, o projeto experimental requer emparelhamento dos sujeitos nos grupos experimental e de controle. Em alguns casos, estabelecem-se cotas para tipos diferentes de sujeitos. Sempre que o número de sujeitos em potencial exceder muito o número requerido pelo experimento, pode-se usar técnicas padrão de amostragem. Métodos de estratificação podem ser usados como auxiliares para satisfazer os requisitos das cotas.

Observação Participante

Diferentemente dos pesquisadores de *survey*, observadores participantes tipicamente buscam mergulhar na totalidade

do fenômeno sob estudo. Podem tentar observar todos os grandes eventos que estão ocorrendo, falar com o maior número de participantes possível, e assim por diante. No entanto, obviamente ninguém consegue observar tudo; alguma seletividade é inevitável. Se tal seletividade for incontrolada, há o risco de reunir um conjunto de observações tendenciosas, tal como um pesquisador de *survey* inepto ao selecionar uma amostra tendenciosa de entrevistados.

Embora não queira sugerir que todos os observadores participantes em todos os lugares em todas as situações de pesquisa devam buscar estabelecer métodos de amostragem rigorosos, estou convencido de que um observador participante bem versado na lógica e, talvez, em alguns dos métodos da amostragem de *survey* tem maior probabilidade de obter um conjunto representativo de observações. Ao avaliar o humor de estudantes reunidos para uma demonstração de protesto, um observador participante deve ter o cuidado de não falar apenas com os líderes da demonstração, devendo considerar a possível diferença entre estudantes que aparentemente comparecem sozinhos e os que vão com amigos, conversar com estudantes em locais diversos da reunião e procurar conversar com os que chegam cedo e com os que chegam atrasados. Embora o observador não tenha condições de estratificar de forma rigorosa a amostra de entrevistados, ele pode ficar mais sensível a reparar diferenças relacionadas às variáveis citadas, refinando as observações de forma correspondente.

Resumo

Este longo capítulo buscou familiarizá-lo com as considerações mais importantes na lógica da amostragem de *survey* e com as técnicas mais comuns de amostragem. Esta discussão não basta para equipá-lo para cada situação de campo que você enfrentará, mas espero que uma compreensão da lógica da amostragem lhe permita chegar a julgamentos razoáveis e seguros por conta própria.

O capítulo seguinte descreve quatro amostras de *survey* com as quais estou intimamente familiarizado. Ao apresentar detalhes específicos destes desenhos de amostras um tanto diferentes e das muitas decisões tomadas na sua execução, espero dar-lhe alguma experiência concreta de amostragem. Repito que estes exemplos não podem ser exaustivos em relação às situações com as quais você se deparará, mas eles foram escolhidos para representar as situações de amostragem mais comuns encontradas na pesquisa de *survey*.

Notas

¹ Para uma discussão recente, ver SQUIRE, Peverill. Why the 1936 *Literary Digest* Poll Failed. *Public Opinion Quarterly*, v.52, p.125-133, primavera 1988.

² Citada na *Newsweek*, p.24, 8 de julho de 1968.

³ Logo veremos que o tamanho da amostra selecionada e as características reais da população maior afetam o grau de representatividade.

⁴ KISH, Leslie. *Survey Sampling*. New York: John Wiley & Sons, 1965.

⁵ Evidentemente, um guia telefônico não inclui novos assinantes nem quem prefere não ter o número listado. A amostragem se complica mais com a inclusão de listagens não residenciais, como lojas e escritórios.

⁶ Veja, por exemplo, GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. Apêndice A.

⁷ Por exemplo, KISH. *Survey Sampling*.

⁸ HORNIK, Jacob, ELLIS, Shmuel. Strategies to Secure Compliance for a Mall Intercept Interview. *Public Opinion Quarterly*, v.52, p.539-551, inverno 1988.

Leituras Adicionais

BABBIE, Earl. *The Practice of Social Research*. Belmont, CA: Wadsworth, 1989. cap.7.

COCHRAN, William G. *Sampling Techniques*. New York: John Wiley & Sons, 1963.

FRANKEL, Martin R., FRANKEL, Lester R. Fifty Years of *Survey Sampling* in the United States. *Public Opinion Quarterly*, v.51, p.S127-S138, inverno 1987.

HANSEN, M. H., HURWITZ, W. N., MADOW, W. G. *Sample Survey Methods and Theory*. New York: John Wiley & Sons, 1953. 2 V.

KISH, Leslie. *Survey Sampling*. New York: John Wiley & Sons, 1965.

SCHEAFFER, Richard L., MENDENHALL, William, OTT, Lyman. *Elementary Survey Sampling*. North Scituate, MA: Duxbury Press, 1979.

Capítulo 6

Exemplos de Desenhos de Amostragem

O Capítulo 5 apresentou a lógica básica da amostragem de *survey* e delineou algumas opções de procedimentos possíveis. Este capítulo apresenta quatro estudos de caso de desenhos de amostras que representam diferentes situações e desenhos de amostragem.

O primeiro exemplo é uma amostragem estratificada sistemática de estudantes da Universidade do Havaí. O segundo é uma amostragem por conglomerados de professores de escolas de medicina em todo o país, atribuindo igual probabilidade de seleção às unidades primárias de amostragem. O terceiro é uma amostra por conglomerados de mulheres de igrejas episcopais no norte da Califórnia, usando um projeto PPT (probabilidade proporcional ao tamanho) para a seleção de unidades primárias de amostragem. O último exemplo é uma amostra complexa de área, projetada para um *survey* de residências em Oakland, Califórnia.

Amostragem de Estudantes Universitários

A finalidade era fazer um *survey*, usando um instrumento auto-administrado, de uma interseção representativa de estudantes do campus principal da Universidade do Havaí. As próximas seções descrevem os passos e as decisões envolvidas na seleção desta amostra.

População de Amostra e Moldura de Amostragem

A moldura de amostragem óbvia era o registro magnético em fita usado pela administração da universidade, com os nomes, os endereços atual e permanente, os números da Previdência Social dos alunos e outras informações como área de estudos, período, idade, sexo e assim por diante.

Mas a fita tinha arquivos de todas as pessoas que, por qualquer definição concebível, poderiam ser chamadas de estudantes. Como vários destes alunos pareciam inadequados para os fins do estudo, foi necessário definir a *população do survey* de forma mais restrita. A definição final incluía os 15.225 alunos de graduação matriculados no campus Manoa da universidade, de todas as faculdades e departamentos, em graduação e pós-graduação, e alunos americanos e estrangeiros. O programa usado para a amostragem limitou o estudo a alunos enquadrados nesta definição.

Estratificação

O programa de amostragem também permitia a estratificação de estudantes antes da seleção da amostra. Decidiu-se que a estratificação por período seria suficiente, embora os estudantes pudessem ter sido ainda mais estratificados dentro de cada período, por sexo, faculdade, área de estudo etc.

Seleção da Amostra

Uma vez agrupados por período (pelo programa de amostragem), extraiu-se uma amostra sistemática a partir da lista reordenada dos estudantes. O tamanho da amostra foi inicialmente definido em 1.100. Para extrair a amostra, o programa foi instruído a usar uma fração amostral de 1/14. Portanto, o programa gerou um número aleatório entre 1 e 14; o estudante com este número foi selecionado, bem como todo décimo quarto estudante daí por diante.

Depois de extrair a amostra desta forma, o computador foi instruído a imprimir o nome e o endereço de cada estudante em seis etiquetas postais auto-adesivas, que foram coladas nos envelopes usados para enviar os questionários.

Modificação da Amostra

A descrição precedente detalha o desenho inicial da amostra para o estudo de estudantes universitários. Antes de enviar os questionários, descobriu-se que custos inesperados na produção dos questionários tornavam impossível cobrir as despesas de

correio para 1.100 estudantes. Portanto, um terço das etiquetas postais foram selecionadas sistematicamente (com início aleatório) para exclusão da amostra. A amostra final foi reduzida para cerca de 770.

Esta modificação da amostra é mencionada para ilustrar a necessidade freqüente de mudar aspectos da pesquisa no meio do caminho. Como uma amostra sistemática de estudantes foi eliminada da amostra sistemática inicial, os 770 restantes ainda podiam ser considerados como representando razoavelmente a população do *survey*. Naturalmente, a redução no tamanho da amostra aumentou a extensão do erro amostral.

Amostragem dos Professores de Escolas de Medicina

Esta seção relata o desenho de amostra usado para selecionar uma amostra de professores de escolas de medicina, para um *survey* nacional estudando os efeitos da orientação científica sobre o cuidado humano de pacientes. O desenho da pesquisa demandava uma amostra de professores de escolas de medicina, nos departamentos de medicina e pediatria.

Em condições ideais, uma lista única de todos os professores nos dois departamentos teria sido obtida ou construída, e uma amostra selecionada a partir dela. Infelizmente, esta lista parecia não existir, de forma que se decidiu selecionar uma amostra por conglomerados em duas etapas. Na primeira, fez-se uma amostra de escolas de medicina; a seguir, professores foram escolhidos de cada uma destas escolas.

Desde o começo, o desenho da amostra foi dificultado pela indisponibilidade de dados. O desenho da pesquisa pedia um exame de professores em tempo integral e em tempo parcial. Embora, na época do estudo, houvesse cerca de 3.700 professores em tempo integral nos dois departamentos, nacionalmente, não havia dados bons disponíveis quanto aos números de professores em tempo parcial. Uma análise dos dados existentes sugeriu que o número total de membros em horário integral e parcial estava em torno de 12.000. Para fins do estudo, decidiu-se que uma amostra de 2.000 seria suficiente (uma fração amostral geral de 1/6).

A Seleção de Escolas de Medicina

Na época da pesquisa, 84 escolas de medicina com cursos de quatro anos de duração pertenciam à Associação Americana de Escolas de Medicina. Estas escolas compuseram a população

do *survey* das escolas, que foram então arranjadas em estratos geográficos e ordenadas por tamanho (número de alunos) dentro dos estratos.

A lista estratificada de escolas foi numerada de 1 a 84 e um número aleatório entre 1 e 6 selecionado (o intervalo amostral). A escola com o número escolhido e cada sexta escola daí por diante foram selecionadas na primeira etapa de amostragem. Foram enviadas cartas aos reitores das escolas selecionadas explicando a finalidade do *survey* e pedindo sua assistência na obtenção de uma lista de professores nos seus departamentos de medicina e pediatria.

Inicialmente, foram selecionadas quatorze escolas. Mas nem todos os reitores quiseram cooperar. À medida que as recusas foram chegando, uma escola alternativa para cada uma foi escolhida da lista, da seguinte maneira: uma escola adjacente à que recusava era escolhida jogando uma moeda.

Seleção de Professores

Logo que reitores concordaram a cooperar com o *survey*, foram compiladas em cada uma listas de todos os professores nos departamentos de medicina e pediatria. Todos estes professores foram incluídos na amostra final, recebendo questionários pelo correio.

Este não foi o melhor desenho de amostra possível. A amostra total de professores foi tirada de um número relativamente pequeno de escolas. Um desenho melhor teria selecionado mais escolas, com menos professores em cada uma. Por exemplo, um terço das escolas e metade dos professores em cada podiam ter sido estudados.

O desenho adotado foi limitado mais por questões administrativas do que científicas. Um estudo-piloto havia mostrado a dificuldade de obter aprovação e cooperação dos reitores. Mesmo quando eles concordavam em cooperar, muitas vezes demoravam a fornecer a lista dos professores. O principal gargalo na amostragem ocorreu neste ponto. Aumentar o número de escolas teria aumentado o tempo e os problemas envolvidos na seleção total da amostra. Apenas por isto, tomou-se a decisão de escolher um sexto das escolas e todo o corpo docente de cada uma.

Amostragem de Mulheres de Igrejas Episcopais

A finalidade deste estudo foi examinar as atitudes de mulheres membros de igrejas na diocese californiana da Igreja Episcopal. Desejava-se uma amostra representativa de todas

as mulheres da igreja na diocese. Como se esperava, não havia uma lista única destas mulheres, de modo que um desenho de amostra em etapas múltiplas foi criado. Na fase inicial de amostragem, as igrejas foram selecionadas; a seguir, mulheres foram selecionadas de cada igreja escolhida. Ao contrário da amostra de escolas de medicina, a de igrejas foi selecionada com *probabilidade proporcional ao tamanho* (PPT).

Seleção das Igrejas

A diocese em questão publica um relatório anual com uma lista das mais ou menos 100 igrejas componentes, com seus respectivos tamanhos em números de membros. A lista constituiu a moldura de amostragem para a primeira etapa.

Desejava-se cerca de 500 respondentes para o estudo, de forma que se decidiu escolher vinte e cinco igrejas com probabilidade proporcional ao tamanho, tirando vinte mulheres de cada igreja. Para isto, a lista de igrejas foi arranjada geograficamente, e foi feita uma tabela semelhante à lista parcial da Tabela 6-1.

TABELA 6-1
Forma usada na lista das igrejas

Igreja	Membros	Membros (acumulados)
Igreja A	3.000	3.000
Igreja B	5.000	8.000
Igreja C	1.000	9.000

Ao lado de cada igreja na tabela, pôs-se seu número de membros, para computar a frequência acumulada ao longo da lista. O total final chegou a cerca de 200.000. O objetivo, neste passo, foi selecionar uma amostra de vinte e cinco igrejas, de tal forma que cada igreja teria uma chance de seleção proporcional ao seu número de membros. Para isto, os totais cumulativos foram usados para criar, para cada igreja, faixas de números iguais aos números de membros daquela igreja. A igreja A na Tabela 6-1 recebeu os números de 1 a 3.000; a igreja B recebeu os números 3.001 a 8.000; a igreja C recebeu 8001 a 9000; e assim por diante.

Selecionando vinte e cinco números entre 1 e 200.000, seria possível selecionar 25 igrejas para a pesquisa. Os 25 números foram selecionados numa amostra sistemática como se segue. O intervalo amostral foi estabelecido em 8.000 (200.000/25), e um início aleatório foi escolhido entre 1 e

8.000. Suponha que o número aleatório escolhido tenha sido 4.538. Como este número estava dentro da faixa de números atribuídos à igreja B (3.001 a 8.000), a igreja B foi selecionada.

Incrementos de 8.000 (o intervalo amostral) foram somados ao início aleatório, e cada igreja em cuja faixa aparecia um dos números resultantes era selecionada para a amostra de igrejas. Deve estar claro que, assim, cada igreja na diocese tinha uma chance de seleção diretamente proporcional ao número de membros. Uma igreja com 4.000 membros tinha duas vezes mais chance de seleção do que uma igreja com 2.000 membros e dez vezes mais do que uma igreja de 400 membros.

Selecionando as Mulheres das Igrejas

Depois de selecionar a amostra de igrejas, obtiveram-se listas das mulheres de cada igreja. Vale notar aqui que, na prática, as listas variavam enormemente em forma e conteúdo. Em vários casos, foram fornecidas listas de todos os membros (homens e mulheres), tornando necessário separar as mulheres antes de amostrar. A forma das listas variava desde listas datilografadas até cartões 3x5 impressos a partir de placas de *addressograph*.

A cada lista que chegava de uma igreja selecionada, computava-se um intervalo amostral para ela, com base no número de mulheres membros e no número desejado (20). Se uma igreja tinha 2.000 mulheres, o intervalo amostral era estabelecido em 100. Um número aleatório era escolhido e incrementado no valor do intervalo amostral, para selecionar a amostra de mulheres daquela igreja. Este procedimento foi repetido para cada igreja.

Observe que este desenho de amostra só dá a cada mulher na diocese uma chance igual de seleção se metade dos membros de cada igreja forem mulheres (ou se uma proporção constante dos membros são mulheres), porque as igrejas tinham chance igual de seleção baseada no número total de membros (não estavam disponíveis números só de mulheres). Dadas as metas do estudo, as leves desigualdades de seleção foram consideradas insignificantes.

Um desenho de amostra mais sofisticado para a segunda etapa teria resolvido este possível problema. Como se deu a cada igreja uma chance de seleção baseada num número presumido de mulheres (supondo-se, por exemplo, 1.000 mulheres numa igreja com 2.000 pessoas), o intervalo amostral podia ter sido computado com base naquela suposição e não no número

real de mulheres listadas. Se tivesse sido suposto, na primeira etapa de amostragem, que uma igreja tinha 1.000 mulheres (de um total de 2.000 membros), o intervalo amostral podia ter sido estabelecido em 50 (1.000/20). Este intervalo podia ter sido usado na seleção de respondentes, independentemente do número real de mulheres listadas naquela igreja. Se, de fato, 1.000 mulheres estavam listadas, a igreja delas teria a chance apropriada de seleção e 20 mulheres seriam selecionadas. Se 1.200 mulheres estavam listadas, a igreja tinha uma chance muito pequena de seleção, mas este desequilíbrio teria sido remediado através da seleção de 24 mulheres pelo uso do intervalo de amostragem preestabelecido. Se, por outro lado, somente 800 mulheres estavam listadas, apenas 16 teriam sido selecionadas.

Amostragem de Casas em Oakland

Este exemplo final representa um dos desenhos de amostra mais complexos da pesquisa de *survey*: uma amostra de área por conglomerados. A finalidade deste estudo, feito em 1966, foi coletar dados relevantes ao estudo da pobreza nas regiões mais pobres de Oakland, Califórnia, usando o restante da cidade para fins de comparação. Uma vez que se ia usar os achados do *survey*, em parte, para apoiar solicitações de recursos federais para Oakland, era essencial que os dados coletados descrevessem precisamente a cidade.

Para fins do estudo, a cidade foi dividida em sete áreas: quatro eram oficialmente designadas áreas pobres, e as três restantes eram vistas tradicionalmente como seções distintas da cidade. O tamanho total da amostra foi estabelecido em 3.500 residências, com base em cálculos cuja complexidade excede o escopo deste livro. Entretanto, note-se que a escolha do tamanho da amostra começou com discussões de políticas referentes à "chance" que as autoridades municipais estavam dispostas a correr de que o *survey*, pelo erro amostral normal, subestimasse os níveis de pobreza e desemprego a ponto de não qualificar a demanda de financiamento federal. Se estes níveis fossem, de fato, altos o bastante para garantir o financiamento, uma amostra perfeitamente precisa demonstraria isto. Mas, com uma amostra pequena, a faixa de erro amostral abria a possibilidade de subestimação. Naturalmente, com o aumento do tamanho da amostra, a chance de subestimação reduzia-se. Portanto, desta maneira estabeleceu-se o tamanho final da amostra em 3.500 residências. (*Nota*: É desta forma que deve ser estabelecido o tamanho das amostras.)

Como o estudo pedia uma comparação das sete regiões da cidade, era importante descrevê-las com igual grau de precisão. Portanto, a amostra foi projetada de forma a selecionar 500 residências de cada região, embora as sete áreas diferissem muito no número total de residências. O restante desta discussão é dedicado aos procedimentos de seleção da amostra usados numa das áreas, já que os procedimentos foram essencialmente idênticos em todas.

Considerações Gerais

Como dissemos, planejamos uma amostragem de área por conglomerados. No primeiro estágio de amostragem, uma amostra estratificada de *quarteirões de recenseamento*¹ devia ser selecionada com probabilidade proporcional ao tamanho. Os quarteirões selecionados seriam listados da seguinte forma: enumeradores visitariam fisicamente cada quarteirão selecionado e preparariam listas de todas as residências ali achadas; então, uma amostra sistemática de cinco residências seria selecionada de cada quarteirão.

Como dissemos, cinco residências por quarteirão é uma convenção comum na amostragem de área por conglomerados. Tanto a teoria de amostragem quanto a experiência de *survey* sugerem que uma amostra de cinco residências representa o ponto de retornos decrescentes na descrição de quarteirões. Cinco residências permitem descrever um quarteirão de forma razoavelmente precisa, dada a homogeneidade tipicamente encontrada entre residências num único quarteirão. Embora, é claro, seis residências dessem uma descrição mais precisa, a redução correspondente na precisão da amostra dos quarteirões, selecionados para descrever a população de todos os quarteirões, anularia esta vantagem. (Supondo que o tamanho total da amostra é fixo.) Num exemplo extremo, você poderia limitar sua amostra a todas as residências encontradas num único quarteirão; ela permitiria uma descrição bem precisa do quarteirão, mas o próprio quarteirão não permitiria descrever muito precisamente todos os quarteirões e, por extensão, todas as residências da cidade.

Selecionar 500 residências de cada região da cidade significava selecionar 100 quarteirões, com cinco residências em cada um. Para aumentar a representatividade da amostra de residências, decidiu-se organizar os quarteirões em estratos relativamente homogêneos (grupos de quarteirões) e selecionar dois quarteirões de cada estrato. Portanto, a tarefa inicial em cada

região foi criar cinquenta estratos; a seguir, dois quarteirões foram selecionados de cada estrato, e, finalmente, cinco casas em cada quarteirão — totalizando 500 residências em cada região.

A Moldura de Amostragem da Primeira Etapa

Foram necessários dois tipos de dados para descrever os quarteirões de recenseamento. O mais importante era saber quantas residências havia aproximadamente em cada um, de forma que se pudesse selecionar uma amostra PPT. Para fins de estratificação, também era preciso conhecer algumas características relevantes dos quarteirões, como sua composição racial e seus níveis socioeconômicos.

As estatísticas dos quarteirões de recenseamento de 1960 dos Estados Unidos ofereciam os dois tipos de dados. Este arquivo indicava o tamanho de todos os quarteirões em 1960 (em número de unidades domiciliares) e também fornecia descrições variadas, como (1) porcentagem de não-brancos, (2) porcentagem de ocupantes pagando aluguel, (3) porcentagem em deterioração, e (4) valor das estruturas (seja em aluguel ou em avaliação). Infelizmente, o estudo foi feito seis anos depois da coleta de dados do censo, de forma que antecipou-se que várias das contagens de casas estariam desatualizadas.

Antes de começar a amostragem, gastaram-se dois meses examinando os mapas de planejamento de cada área. Unidades domiciliares construídas ou demolidas desde o censo de 1960 foram anotadas, e estes dados usados para corrigir o número estimado de residências por quarteirão. Enfim, preparou-se um registro de dados para cada quarteirão, mostrando seu tamanho atual esperado (em número de residências) e as diversas características compiladas no censo de 1960 (que não seriam atualizadas).

As seções seguintes descrevem os procedimentos de seleção da amostra usadas na Área de Pobreza D, que continha 9.938 residências estimadas na época da amostragem. Já que a meta de amostragem para a área era de 500 casas, a fração total de amostragem foi arredondada para 1/20.

Seleção de Grandes Quarteirões

Cada área da cidade tinha algum quarteirão de recenseamento com um grande número de residências. Alguns eram grandes em tamanho geográfico; outros tinham vários prédios grandes de apartamentos. Como a presença destes conjuntos

habitacionais podia prejudicar as técnicas de estratificação planejadas, e pelo fato destes quarteirões serem considerados muito importantes para os propósitos do estudo, decidiu-se que cada quarteirão com 200 ou mais residências seria automaticamente incluído na amostra. Estes quarteirões seriam listados, e um vigésimo (taxa de amostragem total para a região) das residências listadas selecionada para entrevistas.

Na Área de Pobreza D, os quarteirões maiores tinham um total de 702 residências. Portanto, trinta e cinco residências foram selecionadas daqueles quarteirões, deixando 465 para serem selecionadas do restante da área.

Lidando com Quarteirões Pequenos e Quarteirões Zero

Vários quarteirões na Área de Pobreza D (e em outros locais) tinham poucas residências, estimando-se que alguns não tinham nenhuma. Estes quarteirões apresentavam problemas especiais. Primeiro, um quarteirão com tamanho esperado igual a zero não teria chance de seleção. Se as estimativas estivessem incorretas e o quarteirão, de fato, tivesse residências, elas não teriam qualquer chance de seleção. Segundo, uma vez que o desenho de amostra básico pedia a seleção de cinco residências de cada quarteirão selecionado, aqueles com menos de cinco também representaram um problema. Além disso, se as estimativas de tamanho estivessem incorretas, alguns quarteirões que se acreditava ter mais do que cinco casas podiam, de fato, ter menos.

Para resolver estes problemas, cada quarteirão zero e cada quarteirão com um tamanho esperado de menos que dez residências foi "anexado" a um quarteirão adjacente. O número de residências esperado num pequeno quarteirão, se houvesse algum, era somado ao número do quarteirão adjacente, e o par de quarteirões tratado como um quarteirão único para fins de amostragem. O par tinha uma chance de seleção proporcional ao seu tamanho combinado; se o par fosse selecionado na amostra de quarteirões, ambos eram listados e amostrados como se fossem um único quarteirão.

Estratificação

Uma vez que trinta e cinco residências seriam selecionadas de grandes quarteirões, 465 precisavam ser selecionadas do restante da área, com seu total estimado de 9.236 residências. Com cinco residências a serem selecionadas de cada quarteirão da amostra, noventa e três quarteirões seriam necessários. Como dois

quarteirões seriam selecionados de cada estrato, a tarefa neste ponto era criar quarenta e sete (arredondado de 46,5) estratos.

Cada estrato (grupo de quarteirões) tinha que ter duas características. Primeiro, os quarteirões em cada estrato deviam ser tão semelhantes entre si quanto possível, em termos de composição racial, status socioeconômico (SSE) etc., para garantir que todos os tipos de quarteirões fossem selecionados na proporção adequada na amostra total. Segundo, cada estrato devia conter um total de aproximadamente 200 residências. Já que dez casas deviam ser selecionadas de cada estrato (dois quarteirões, cinco residências de cada), um tamanho de estrato de 200 produziria uma fração amostral total de 1/20, a fração estabelecida para toda a área. Naturalmente, todos os quarteirões tinham números especificados de residências, de forma que não foi possível criar estratos com exatamente 200 casas. (Uma discussão posterior tratará da correção estatística para tais variações.)

Ao criar homogeneidade entre os quarteirões agrupados num estrato, foi possível empregar as características de quarteirões fornecidas no censo de 1960. Para evitar confusão nas descrições empíricas seguintes, observe que a criação de homogeneidade foi feita de forma essencialmente *ad hoc* e arbitrária. Embora possa haver a tentação de criar um formato estratificado em bases teóricas desde o início (por exemplo, todos os quarteirões com mais de 75% de residentes não-brancos, mais de 50% de unidades alugadas e menos de \$150 em renda mensal média seriam combinadas para formar um estrato), esta abordagem não seria necessariamente apropriada à natureza da região ou área específica sendo amostrada. Em vez disso, cada grupo de quarteirões (representando uma região da cidade ou subconjunto dela) era examinado para determinar as variações das suas características, e um sistema de estratificação foi desenvolvido para se adequar àquelas características particulares. Portanto, áreas diferentes da cidade foram estratificadas de forma diferente; ademais, deve-se reconhecer que o formato de estratificação específico a uma região era apenas uma de várias, talvez igualmente apropriadas, possibilidades.

Entretanto, como regra geral, as variáveis de estratificação disponíveis foram consideradas numa determinada ordem para cada área da cidade: crescimento recente, composição racial, porcentagem de casas próprias/alugadas, valor das propriedades e deterioração. Sempre que esta ordem de variáveis de estratificação fosse inapropriada para uma dada área, as variáveis inapropriadas eram ignoradas ou consideradas num ponto diferente da estratificação.²

Portanto, em cada uma das sete regiões da cidade, iniciamos a estratificação com base no crescimento recente. Quando os registros de dados atualizados foram criados para cada quarteirão de recenseamento, foi feita uma anotação quanto ao número de unidades residenciais construídas após o censo de 1960. Como parecia razoável supor que quarteirões contendo novas construções difeririam de outros quarteirões, este dado foi usado como primeira variável de estratificação. Por exemplo, todos os quarteirões na Área de Pobreza D contendo alguma nova unidade foram separados para uma estratificação separada. Juntos, estes quarteirões tinham 1.254 residências estimadas. Com 200 casas por estrato, os quarteirões “com crescimento” foram agrupados em seis estratos.

A segunda variável de estratificação usada foi composição racial. Todos os quarteirões com crescimento foram ordenados em termos da porcentagem de residentes não-brancos vivendo em cada um. Alguns quarteirões tinham 80% ou mais de residentes não-brancos. Juntos, estes quarteirões tinham 247 residências. Uma vez que este número estava relativamente próximo da meta de 200 casas por estrato, estes quarteirões foram designados como o primeiro estrato.

Quarteirões contendo entre 30 e 79% de residentes não-brancos tiveram um total de 385 residências; decidiu-se criar dois estratos destes quarteirões. A variável de estratificação seguinte foi a porcentagem de residências alugadas (em contraposição às próprias). Nos estratos aqui considerados, os quarteirões com 36% ou mais de residências alugadas representavam um total de 214; as 171 restantes estavam em quarteirões com menos de 36% de residências alugadas. Estes dois grupos de quarteirões foram designados como o segundo e o terceiro estrato. (*Nota:* Não há nada intrinsecamente significativo na escolha do corte em 36%. A meta era criar estratos de tamanho aproximadamente igual, e o ponto de corte que preenchia esta condição foi usado.)

Quarteirões com menos de 30% de residentes não-brancos tinham 622 residências, pedindo três estratos. Examinando estes quarteirões em termos de todas as variáveis de estratificação, descobriu-se que mais de 10% das residências num grupo de quarteirões foram classificados pelo censo como “em deterioração”. Estes quarteirões tinham 214 residências. Como 10% de deterioração é bastante alto no contexto de todos os quarteirões, este grupo foi designado como sendo o quarto estrato.

Os quarteirões restantes tinham 408 residências, pedindo dois estratos. Quando se examinou a ocupação em termos de aluguéis, descobriu-se que os quarteirões com 86% ou mais de aluguéis tinham 201 residências; estes quarteirões foram

designados como o quinto estrato. Os quarteirões contendo as 207 casas restantes foram designados como o sexto estrato.

Em vista da complexidade deste procedimento, apresentamos na Tabela 6-2 um resumo esquemático da estratificação dos quarteirões com crescimento na Área de Pobreza D. Os grupos de quarteirões designados como estratos são indicados pela anotação S1, S2, e assim por diante. Lembre que a meta da estratificação era criar grupos relativamente homogêneos de quarteirões. Todos os quarteirões no estrato 5 têm em comum: (1) estão localizados na Área de Pobreza D em Oakland; (2) têm menos de 200 residências; (3) experimentaram um crescimento em número de residências desde 1960; (4) têm menos de 30% de residentes não-brancos; (5) têm menos de 10% das suas estruturas em deterioração; e (6) têm mais de 85% de aluguéis. A homogeneidade extrema deste grupo de quarteirões é evidente.

Esta discussão está limitada à criação de seis estratos na Área de Pobreza D. Quarenta e um outros estratos foram criados desta maneira na mesma região, e o processo inteiro foi replicado mais seis vezes para as outras regiões da cidade.

TABELA 6-2
Estratificação de quarteirões com crescimento na Área de Pobreza D

Todos os quarteirões com crescimento (1.254 resid.)	
80% ou mais não-brancos (247 resid.)	S1
30-79% não-brancos (385 resid.)	
36% ou mais alugadas (214 resid.)	S2
35% ou mais alugadas (171 resid.)	S3
29% ou menos não-brancos (622 resid.)	
10% ou mais em deterioração (214 resid.)	S4
9% ou menos em deterioração (408 resid.)	
86% ou mais alugadas (201 resid.)	S5
85% ou menos alugadas (207 resid.)	S6

NOTA - resid. = residências. Usei este termo mais comum em lugar do termo tecnicamente correto *unidade habitacional*. Uma unidade habitacional é um quarto ou conjunto de quartos para uso residencial por uma pessoa ou família; uma residência é a pessoa ou grupo de pessoas residindo numa unidade habitacional.

Seleção de Quarteirões dentro dos Estratos

Dois quarteirões foram selecionados de cada estrato, com probabilidade proporcional ao seu tamanho. Usaremos o estrato 1 do exemplo anterior, que tinha 247 residências, para ilustrar o procedimento usado na seleção de quarteirões.

Para começar, a ordem dos quarteirões no estrato 1 foi randomizada. A seguir, a Tabela 6-3 foi criada a partir das estimativas de residências para cada quarteirão. As primeiras duas colunas na tabela identificam os quarteirões pelo número do trajeto do censo e número do quarteirão (dentro de cada trajeto ou trato). A terceira coluna apresenta os tamanhos estimados dos quarteirões; a quarta coluna dá o total cumulativo nos seis quarteirões. A coluna final apresenta a faixa de números assinalados a cada quarteirão com base no seu tamanho.

Como dois quarteirões deviam ser selecionados, o total cumulativo para o estrato (247) foi dividido por 2 (resultando em 123,5). Um número aleatório entre 1 e 123,5 foi selecionado. O quarteirão 27/5, com uma faixa de 13-42, continha este número aleatório, de forma que foi selecionado na amostra de quarteirões. O número aleatório foi, então, adicionado a 123,5. O quarteirão 27/14, com uma faixa de 149-196, continha este novo número (158,5), de forma que foi o segundo quarteirão selecionado na amostra.

TABELA 6-3
Procedimento de seleção dos quarteirões

Trato (trajeto) do censo	Quarteirão do censo	Número de residências	Total cumulativo	Faixa cumulativa
27	18	12	12	1-12
27	5	30	42	13-42 selecionado
27	23	26	68	43-68
28	4	80	148	69-148
27	14	48	196	149-196 selecionado
28	2	51	247	197-247

NOTA - Total resid./2=123,5; n° aleatório entre 1 e 123,5=35; n° aleatório + (total resid./2)=158,5.

Um programa de computador foi projetado para realizar a maior parte dos passos descritos acima. Os quarteirões foram divididos por estratos. O computador lia os relatórios de dados de um estrato, calculava e imprimia os totais cumulativos para aquele estrato, dividia o total por 2 e imprimia aquele número, e finalmente gerava e imprimia um número aleatório entre 1 e metade do total. Depois determinava quais quarteirões deviam ser incluídos na amostra.

Seleção de Residências dentro de um Quarteirão

O procedimento descrito resultou na seleção de cerca de 700 quarteirões de recenseamento em toda a cidade de

Oakland. Cinco residências (em geral) iam ser entrevistadas em cada quarteirão. Para atingir esta meta, mapas foram preparados de forma a identificar claramente cada um dos quarteirões selecionados, e enumeradores foram mandados para preparar uma lista de todas as residências em cada quarteirão. As listas, preparadas pelo uso de formulários padronizados, assemelhavam-se ao exemplo apresentado na Tabela 6-4.

TABELA 6-4
Amostra de folha de listagem

Número de residências	Nome da rua	Número da rua	Número do apartamento ou outra identificação
01	Walnut St.	2301	
02	Walnut St.	2303A	Duplex
03	Walnut St.	2303B	Duplex
04	Tenth Av.	(102?)	Sem número; casa marrom com sebe
05	Tenth Av.	104	Apart. 101
06	Tenth Av.	104	Apart. 102
07	Tenth Av.	104	Apart. 103
*			
*			
*			
47	Ninth Av.	103	
48	Ninth Av.	101	

Para preparar a listagem, o enumerador ia a uma esquina designada do quarteirão e contornava-o até voltar ao ponto de origem, anotando cada residência no formulário à medida que aparecia. Cada residência, seja uma casa com uma única família, um meio duplex ou um apartamento, era anotada separadamente, recebendo um número em ordem seriada contínua.

O processo de listagem forneceu uma nova estimativa do número de residências em cada quarteirão. (Como os enumeradores podiam cometer erros na listagem, este número ainda assim deve ser considerado uma *estimativa*.) Entretanto, lembre que cada quarteirão era selecionado com base em uma estimativa anterior do seu tamanho. Para levar em conta discrepâncias entre estas duas estimativas, empregamos uma técnica mencionada antes na nossa discussão sobre a amostragem de mulheres da igreja episcopal.

O intervalo amostral usado na seleção de residências de um dado quarteirão foi computado com base na estimativa anterior do seu tamanho e nas cinco residências que se pretendia

selecionar. Se o tamanho do quarteirão tivesse sido estimado em cinquenta residências, o intervalo amostral era estabelecido em dez. Este intervalo foi usado na seleção de residências, mesmo se o número de casas de fato listadas fosse maior ou menor. Se o quarteirão tivesse sessenta residências em vez das cinquenta estimadas, seis eram selecionadas na amostra; se tivesse só quarenta, quatro eram selecionadas. (Uma correção mais precisa também foi empregada, o que discutiremos em breve.)

A seleção de residências foi feita escolhendo um número aleatório entre 1 e o intervalo amostral. O número aleatório foi então somado ao intervalo amostral, e as residências listadas ao lado dos números resultantes na lista foram escolhidas para entrevista.

Quando as estimativas iniciais do tamanho do quarteirão eram muito imprecisas e os procedimentos mencionados produziam apenas uma ou duas residências na amostra (ou mais do que dez), um procedimento diferente foi usado. Nestes casos, um número arbitrário de residências (não mais do que dez) era selecionado, e se fazia uma anotação para que aquelas entrevistas tivessem peso separado durante a análise.

Isso completa nossa discussão sobre os procedimentos usados na seleção de cerca de 3.500 residências na cidade de Oakland, Califórnia, em 1966. O processo todo, incluindo a atualização do tamanho dos quarteirões de recenseamento e a listagem dos quarteirões de amostra, levou cerca de cinco meses, exigindo aproximadamente vinte pessoas no pico do trabalho.

Ponderação das Residências da Amostra

No desenho da amostra aleatória simples, cada elemento na população tem a mesma probabilidade de seleção. Como resultado, a amostra agregada pode ser tomada como representativa da população da qual foi extraída. Se 2.000 respondentes forem selecionados de uma população de 2 milhões, então considera-se que cada entrevistado estaria "representando" mais 999 pessoas não selecionadas. Para estimar o número de pessoas na população com uma determinada característica, multiplicamos por 1.000 o número de pessoas com aquela característica na nossa amostra. Este peso é o inverso da probabilidade dos respondentes serem selecionados na amostra. Quando todos têm a mesma probabilidade de seleção, o peso é irrelevante, exceto para estimar números na população.

Quando os respondentes têm probabilidades diferentes de seleção, a ponderação é mais importante, sendo relevante até mesmo na computação das porcentagens. O estudo

Oakland é um exemplo da necessidade de e dos métodos para ponderar elementos de amostra num desenho complexo de amostra. Entretanto, independentemente da complexidade do projeto, o princípio básico se mantém: o peso do respondente é o inverso da sua probabilidade de seleção na amostra.

Ao computar a probabilidade geral de seleção de um respondente, devemos lembrar-nos de que se temos várias etapas de amostragem, com probabilidades diferentes de seleção em cada uma, estas probabilidades devem ser multiplicadas para determinar a probabilidade geral. Se os respondentes pertencem a um grupo (igreja, quarteirão) com a chance de seleção de um décimo, e eles têm uma chance de seleção de um décimo dentro daquele grupo, sua probabilidade geral de seleção é de 1/100.

No cálculo dos pesos das residências na amostra de Oakland, devemos considerar duas probabilidades separadas de seleção: a probabilidade de um quarteirão ser selecionado e a de uma residência ser selecionada dentro daquele quarteirão. Estas probabilidades devem ser calculadas como se segue.

Probabilidade do Quarteirão. Cada quarteirão tem a probabilidade de seleção igual ao seu tamanho estimado (TEQ), dividido pelo tamanho do estrato (TE), vezes 2 (dois quarteirões selecionados por estrato). Escrevemos esta equação como $2TEQ/TE$. Note que esta fórmula considera tanto a amostragem PPT quanto a variação dos tamanhos dos estratos. No nosso exemplo anterior, o quarteirão 27/5 tinha um tamanho estimado de 30 residências num estrato contendo 247; portanto, sua probabilidade de seleção era $(30 \times 2)/247 = 0,2429$.

Probabilidade da Residência. Dentro de cada quarteirão, cada residência selecionada tinha uma probabilidade de seleção igual ao número selecionado no quarteirão (n) dividido pelo número de fato (NFQ) listado para todo o quarteirão, escrito como n/NFQ . Caso se observasse na listagem que o quarteirão 27/5 tinha 34 residências (em vez de 30), e fossem selecionadas cinco delas, cada uma teria a probabilidade de seleção igual a $5/34 = 0,1471$.

Probabilidade Geral. Multiplicando as fórmulas separadas para obter a probabilidade geral, temos o seguinte: $(2TEQ)(n)/(TE)(NFQ)$. No exemplo acima, a probabilidade geral de seleção é $(2 \times 30 \times 5)/(247 \times 34) = 0,0357$, ou cerca de 1/28. Observe que esta probabilidade é menor do que a fração de amostragem visada de 1/20 na Área de Pobreza D. A razão para esta discrepância é que o estrato 1 tinha 247 residências em vez da meta de 200, e o quarteirão 27/5 tinha 34 residências em vez das 30 estimadas. Portanto, cada uma das cinco residências

selecionadas no quarteirão 27/5 tinha menor probabilidade de seleção do que na intenção original.

No entanto, note que, se o tamanho estimado do quarteirão e o tamanho real fossem idênticos, a fórmula de probabilidade seria reduzida de $(2TEQ)(n)/(TE)(NFQ)$ para $2n/(TE)$, ou duas vezes o número de residências selecionadas dividido pelo número de residências no estrato. Se o tamanho do estrato (TE) fosse 200, a fração de amostragem visada de 1/20 seria atingida, uma vez que cinco residências teriam sido escolhidas no quarteirão.

Ponderação das Residências. Todas as diferenças em probabilidades de seleção foram levadas em consideração na alocação de pesos igual ao inverso da probabilidade geral de seleção de uma residência. No primeiro exemplo acima, cada residência entrevistada no quarteirão 27/5 teria recebido um peso de 28. O suposto é que cada uma daquelas residências representaria a si própria e outras 27 na cidade de Oakland.

Ponderação Adicional por Não-Resposta. Um comentário final. *Surveys* deste tipo nunca conseguem entrevistar todas as residências inicialmente selecionadas na amostra. Algumas pessoas na amostra se recusam a ser entrevistadas e outras não estão disponíveis.

Neste estudo, foi atribuído peso adicional a residências para dar conta da não-resposta. Supôs-se que as residências em que não se conseguiu fazer entrevistas eram mais como as casas entrevistadas no mesmo quarteirão do que como qualquer outra estimativa possível. Portanto, cada entrevista completada recebia um peso igual ao número de residências selecionadas naquele quarteirão dividido pelo número de entrevistas de fato completadas. Se quatro entrevistas em cinco eram completadas, cada entrevista completa recebia um peso adicional de $5/4$ — 1,25. Naturalmente, se todas as residências selecionadas eram entrevistadas com sucesso, o peso adicional era de 1, podendo ser ignorado.

Resumo

Os quatro exemplos de desenhos de amostra foram apresentados de forma a dar um quadro mais realista das situações de amostragem com as quais você pode se deparar na prática. Embora estes exemplos não esgotem as possibilidades de variação em condições de campo, objetivos do estudo e técnicas de amostragem, eles ilustram as situações mais típicas. Espero

que eles tenham ilustrado a lógica básica da amostragem de *survey*, ajudando-o a improvisar sabiamente quando se deparar com um problema novo.

Notas

¹ Em sua maior parte, um quarteirão de recenseamento é o mesmo que um quarteirão de cidade — uma área mais ou menos retangular, limitada por quatro ruas, com casas alinhadas ao longo das ruas. Entretanto, em alguns lugares as ruas são irregulares, e blocos de formato estranho recebem o nome de quarteirões.

² Por exemplo, na época do estudo, uma região da cidade de Oakland era quase totalmente branca. Portanto, fazia pouco sentido estratificar os quarteirões naquela área em termos da composição racial, embora fosse possível agrupar os quarteirões com *quaisquer* residentes não-brancos. Por outro lado, se um pequeno grupo de quarteirões na região tivesse, digamos, 20% ou mais de residentes não-brancos, estes poderiam ter sido colocados em um estrato separado.

Leituras Adicionais

- ALMOND, Gabriel, VERBA, Sidney. *The Civic Culture*. Princeton, NJ: Princeton University Press, 1963.
- BABBIE, Earl R. *Science and Morality in Medicine*. Berkeley: University of California Press, 1970.
- GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967.
- GLOCK, Charles Y., STARK, Rodney. *Christian Beliefs and Anti-Semitism*. New York: Harper & Row, 1966.
- STOUFFER, Samuel A. *Communism, Conformity, and Civil Liberties*. New York: John Wiley & Sons, 1966.

Capítulo 7

Conceituação e Desenho de Instrumentos

A pesquisa científica tem duas metas principais: *descrição* e *explicação*. Pesquisadores medem as distribuições empíricas de valores nas variáveis (descrição) e as associações entre as variáveis para explicar essas distribuições de valores. O Capítulo 13 discute as interconexões lógicas entre descrição e explicação. A finalidade deste capítulo é fundamentar estes processos, examinando a lógica e as possibilidades da medição descritiva.

Seja trabalhando a partir de uma teoria rigorosamente deduzida, seja a partir de um conjunto de suspeitas e palpites, em algum ponto você se depara com conceitos abstratos, não especificados, que acha que o ajudarão a compreender o mundo que o cerca. Na pesquisa de *survey*, estes conceitos devem ser convertidos em perguntas num questionário, de forma a permitir a coleta de dados empíricos relevantes para a análise.

O capítulo começa com notas gerais relativas à lógica da conceituação e da operacionalização. A seguir, considera os diferentes tipos de dados que se pode levantar num *survey*. Finalmente, se debruça sobre algumas técnicas para construir boas perguntas.

Lógica da Conceituação

Muitas vezes se quer estudar conceitos como classe social, alienação, preconceito, sofisticação intelectual, e assim por diante. Talvez você suspeite que a alienação decresce com a mobilidade social ascendente, ou que a sofisticação intelectual

reduz o preconceito. No entanto, antes de embarcar na pesquisa empírica, estes conceitos podem ser apenas idéias gerais na sua cabeça, sem definição precisa do que significam para você.

Por exemplo, o conceito "status social" é freqüentemente usado na pesquisa social, mas de forma alguma está claro seu significado último. Definições diferentes de classe social incluem: renda, prestígio ocupacional, educação, riqueza, poder, status familiar tradicional e avaliação moral. Provavelmente, nenhuma combinação específica destes elementos produz uma definição de status social plenamente satisfatória para algum pesquisador, muito menos para todos os pesquisadores.

Muitas vezes se diz que tais conceitos não especificados têm "riqueza de significado", no sentido de combinarem uma variedade de elementos, sumariando um fenômeno complexo. Embora o termo "status social" evoque diferentes imagens para pesquisadores diferentes, estas imagens provavelmente parecem a todos eles importantes e significativas. Tais conceitos gerais devem ser especificados para a pesquisa empírica rigorosa, ou seja, reduzidos a indicadores empíricos específicos. A operacionalização de conceitos é inevitavelmente insatisfatória, tanto para os pesquisadores quanto para suas audiências. Em última análise, conceitos ricos em significado devem ser reduzidos a indicadores empíricos simplificados e superficiais.

Penso que a maioria dos conceitos interessantes para os pesquisadores sociais *não tem significado real*, nem definição última. Status social é um bom exemplo. O conceito geral de status social seguramente é antigo. Embora não o tenham verbalizado em termos abstratos, os primeiros humanos reconheceram que alguns deles tinham status social maior do que outros. Alguns eram mais poderosos, alguns eram mais respeitados, e outros recebiam maior autoridade sobre seus pares. O reconhecimento destas diferenças persistiu no decorrer da história humana. Max Weber extraiu muitas implicações dessas diferenças. Karl Marx derivou uma teoria da sociedade com base nelas. W. Lloyd Warner e outros sociólogos americanos dos anos 30 e 40 observaram as mesmas diferenças entre moradores de pequenas cidades, e descobriram que membros de uma comunidade podiam identificar os vizinhos com maior ou menor status social.

A existência de diferenças de status entre membros de uma sociedade é clara. Ademais, tais diferenças parecem importantes para compreender outros aspectos da sociedade. Em vista disso, o termo "status social" tem sido aceito como uma anotação resumida do fenômeno. Mas o que realmente significa o termo "status social"? Uma vez que é apenas um termo para denotar um conceito geral, não tem significado

último. Por si, status social só existe como uma assinalação conveniente para uma diversidade de observações empíricas. Deste ponto de vista, nenhum pesquisador pode pesquisar status social correta ou incorretamente; pode apenas fazer medições mais ou menos úteis.

Neste sentido, os cientistas nunca coletam dados, *eles criam dados*. Esta idéia é essencialmente o mesmo sentimento expresso pelo filósofo da ciência Alfred North Whitehead, quando escreveu o seguinte:

A natureza recebe crédito pelo que deveria estar reservado para nós; a rosa pelo seu perfume; o rouxinol pelo seu canto; e o Sol pela sua radiação. Os poetas estão totalmente enganados. Deviam endereçar sua lírica para si próprios, fazendo dela odes de autocongratulação pela excelência da mente humana. A natureza é assunto sem graça, sem som, sem cor; meramente o apressar da matéria, de forma interminável e sem significado.¹

Quando faz várias perguntas e combina as respostas num índice que chama de classe social, você *criou* uma medida de classe social; você realmente *criou* um agrupamento e uma hierarquia de classe social entre seus respondentes. Você não descobriu, apenas, um agrupamento e uma hierarquia que já existia, num sentido real.

Talvez as medidas criadas tenham utilidade, ajudando-o a entender os dados que tem, ou ajudando-o a desenvolver teorias sobre classes sociais. Entretanto, não tem sentido perguntar se você *realmente* mediu classe social, num sentido válido em última análise, porque o conceito só existe em nossas mentes.

A noção de que pesquisadores criam dados e não os coletam é até mais fundamental do que sugere o exemplo do status social. Nos primeiros dias da II Guerra Mundial, Hadley Cantril fez dois *surveys* nacionais do povo americano.² O desenho das amostras e a checagem cruzada das suas características indicavam que ambos forneciam boas estimativas da população americana. Num dos *surveys*, foi perguntado aos respondentes: "Você acha que os Estados Unidos conseguirão ficar fora da guerra?" A maioria (55% dos que opinaram) disse sim. No outro *survey*, a pergunta foi: "Você acha que os Estados Unidos entrarão na guerra antes que ela termine?" A maioria (59% dos que opinaram) disse sim.

Estes itens são muitas vezes citados como exemplo dos efeitos de perguntas "tendenciosas". Entretanto, sempre que é mostrado a estudantes, eles perguntam: "Qual pergunta forneceu a resposta *correta* sobre as expectativas americanas

quanto a entrar na guerra?” Esta pergunta não pode ser respondida. Podemos apenas concluir que, em 1939, não havia algo como “a atitude americana sobre a probabilidade de entrarmos na guerra”. Nenhuma porcentagem representava a proporção que *realmente* pensava que entraríamos na guerra. Portanto, não se podia *coletar* tais dados; podia-se apenas criá-los fazendo perguntas, e a forma pela qual as perguntas eram feitas tinha um impacto sobre as respostas recebidas.

Nunca conseguimos fazer medidas *precisas*, apenas medidas *úteis*. Mas esta suposição não deve ser tomada como justificativa para a anarquia científica. Pesquisa rigorosa ainda é possível; simplesmente é mais difícil do que se imagina. O resto deste capítulo ocupa-se da lógica e das habilidades requeridas na feitura de medidas úteis.

Um Quadro de Referências de Operacionalização

Como sugere a discussão anterior, conceitos são codificações gerais da experiência e das observações. Observamos pessoas vivendo em tipos diferentes de estruturas residenciais e desenvolvemos o conceito de unidade residencial. Observamos diferenças em posição social e desenvolvemos o conceito de status social. Notamos diferenças no grau de compromisso religioso das pessoas e desenvolvemos o conceito de religiosidade. Entretanto, é imperativo reconhecer que todos estes conceitos são anotações resumidas da experiência e das observações.

Na ciência, tais conceitos muitas vezes assumem a forma de *variáveis* reunindo uma coleção de *atributos* relacionados. Assim, o conceito de “gênero” resume dois atributos distintos: masculino e feminino. “Período escolar” é uma variável composta pelos atributos calouro e veterano. “Filiação religiosa” pode incluir atributos como protestante, católico, judeu, budista e assim por diante.

Muitos dos conceitos mais interessantes na ciência social representam *variáveis ordinais* feitas de valores dispostos ao longo de uma *dimensão*. Por exemplo, o conceito de status social implica uma hierarquia de valores como “status elevado”, “status médio”, “status inferior” etc. Religiosidade, preconceito, alienação, sofisticação intelectual, liberalismo e conceitos similares implicam também um conjunto de valores hierarquizados.

A *operacionalização* é o processo pelo qual pesquisadores especificam observações empíricas que podem ser tomadas como indicadores dos atributos contidos em algum conceito. Se o conceito for religiosidade, a operacionalização é o processo

para especificar medições empíricas que indicarão se os respondentes são altamente religiosos, moderadamente religiosos, não religiosos etc. Em geral, diversos destes indicadores são especificados e combinados durante a análise de dados para fornecer uma medida composta (índice ou escala) representando o conceito.

Uma vez que virtualmente todos os conceitos são resumos *ad hoc* da experiência e das observações, não têm significados reais, últimos. Portanto, você não faz medidas corretas ou incorretas, apenas pode determinar até que ponto as medições contribuem para compreender os dados empíricos e desenvolver teorias do comportamento social. Desta perspectiva, é possível dar diretrizes para o processo de operacionalização, para ampliar sua utilidade nas atividades de pesquisa.

Suponha que você queira estudar religiosidade. Talvez queira aprender por que algumas pessoas são mais religiosas do que outras, ou saber as conseqüências de ser mais ou menos religioso. Embora possa iniciar com uma noção geral do que quer dizer com religiosidade, é igualmente provável que você não tenha qualquer indicador em mente. Para iniciar o processo da operacionalização, você deve enumerar todas as diversas subdimensões da variável. Ao fazer isto, deve prestar atenção a pesquisas prévias sobre o tema, bem como às concepções de senso comum sobre religiosidade.

Charles Y. Glock estudou, teórica e empiricamente, as diversas subdimensões da religiosidade; seus estudos propiciam um começo excelente para a abordagem do tema.³ Ele discute *envolvimento ritual* enquanto participação em atividades como serviços religiosos semanais, comunhão (para os cristãos), orações antes das refeições etc. *Envolvimento ideológico* concerne a aceitação de crenças religiosas tradicionais. *Envolvimento intelectual* refere-se à extensão do conhecimento sobre religião. *Envolvimento experiencial* se refere à extensão das experiências religiosas dos respondentes, como ouvir a voz de Deus, possessões religiosas etc. Finalmente, Glock examina o *envolvimento consequencial*, isto é, o quanto o comportamento social é motivado por preocupações religiosas, de acordo com ensinamentos religiosos.

Embora estas cinco subdimensões da religiosidade ajudem a organizar muitos indicadores possíveis, outros indicadores talvez não se encaixem facilmente nelas. Dar dinheiro para a igreja, ou para atividades relacionadas à igreja, pode ser considerado uma indicação de religiosidade, bem como a participação nas atividades sociais da igreja. A lista poderia prosseguir. Portanto, a primeira tarefa é compilar uma lista, tão exaustiva quanto possível, de todos os diferentes indicadores que poderiam ser incluídos no conceito geral de religiosidade.

Ao decidir o que a religiosidade pode *ser*, é preciso considerar também o que ela *não é*. Deve-se tomar muito cuidado para não construir itens de questionário que indiquem ao mesmo tempo a religiosidade e as variáveis que se irá relacionar a ela, na análise. Se você quiser, por exemplo, determinar a relação entre religiosidade e atitudes em relação à guerra, itens medindo o compromisso com a doutrina cristã da “paz na Terra” não seriam boas medidas de religiosidade. Como as respostas refletiriam as atitudes dos respondentes frente à guerra — além de preocupações religiosas —, tais itens seriam “contaminados”. Mesmo que as respostas certamente se relacionem com atitudes relativas à participação na guerra, os achados não contribuiriam em nada para entender a relação geral entre religiosidade e atitudes em relação à guerra.

Deve-se prestar atenção especial ao *oposto* da variável que se busca medir. Se a meta é medir religiosidade, devemos estar sensíveis à variável anti-religiosidade. Algumas pessoas são muito religiosas e outras não religiosas; outras são anti-religiosas. Como anti-religiosidade pode variar em intensidade, você deve determinar a amplitude de variação conceitual da variável. Talvez optando por uma medida de religiosidade variando de baixa a alta, agrupando os anti-religiosos e os que apenas não são positivamente religiosos. Ou talvez estendendo a variável de muito religioso a muito anti-religioso. Em qualquer caso, os itens usados na medida devem basear-se na sua decisão sobre a amplitude. Considere o seguinte exemplo de um item de questionário:

Por favor, diga se você concorda ou discorda da seguinte declaração: “A religião organizada causa mais mal do que bem.”

Concordância indicaria uma orientação anti-religiosa. Por outro lado, a discordância não indicaria necessariamente religiosidade. Embora as respostas certamente estejam fortemente correlacionadas com medidas de religiosidade, o item propriamente dito não mede religiosidade, apenas anti-religiosidade.

Este exemplo ilustra um problema comum na medida de dimensões. Ao invés de medir uma variável de baixo a alto, muitas vezes medimos dois pólos opostos. Por exemplo, raramente medimos graus de conservadorismo político, medindo ao invés variações entre liberalismo e conservadorismo. Num estudo sobre ciência e medicina, tentei medir graus diversos de compromisso com perspectivas científicas entre médicos, mas me vi constantemente envolvido em questões sobre perspectivas anticientíficas.⁴

Este problema não tem solução simples. Os interesses de pesquisa variam demais para permitir a elucidação de regras adequadas de procedimentos. O único conselho, agora, é realizar o exercício demorado de listar todas as possíveis subdimensões da variável, anotando as dimensões a serem excluídas do conceito e especificando os pontos conceituais extremos da dimensão que descreve o conceito. Cada item do questionário examinado deve ser avaliado segundo estas decisões. Você deve se perguntar quais seriam as implicações de cada possível resposta: como esta resposta reflete o conceito básico? Apenas através deste processo você pode gerar dados relevantes para uma análise posterior significativa.

Tipos de Dados

O formato de pesquisa de *survey* gera vários tipos de dados úteis para a pesquisa social, dos quais esta seção oferece uma visão geral.

Embora o capítulo tenha começado afirmando que os dados só existem através do processo científico de gerá-los, mesmo assim tem sentido ver alguns tipos de dados como “fatos”. Por “fatos”, queremos dizer itens de informação que o respondente acredita representarem a *verdade* e que você, em geral, aceita como sendo verdade. As características demográficas dos respondentes se encaixam nesta categoria. Quando pede aos entrevistados para indicarem seu sexo como masculino ou feminino, eles acreditam que as respostas representam fatos indiscutíveis, e você os aceita como tal. O mesmo pode ser dito das informações sobre idade, raça, região de origem etc.

Às vezes, pede-se aos respondentes para darem informações que eles aceitam como declarações verdadeiras, mas que você não aceita necessariamente como tal. Por exemplo, pode-se perguntar se Deus existe. Respondendo sim ou não, eles dizem o que acreditam ser a verdade sobre o tema. Já você vê estas respostas apenas como descrições dos respondentes, e não como respostas à questão da existência de Deus. Solicitados a concordar ou discordar da declaração “As mulheres são mais emotivas do que os homens,” os respondentes dirão o que acreditam ser verdade e você toma as respostas apenas como descrições dos respondentes.

Em outros casos, você pede ao respondente para dar informações reconhecidas por vocês dois como atitudes subjetivas. Você pode perguntar: “Você acha que o presidente está realizando um bom ou um mau trabalho?” Tanto ele quanto você entendem

que se trata de uma opinião e não de um fato. A distinção entre crenças e atitudes, neste sentido, muitas vezes não é clara, mas, em certos problemas de pesquisa, pode vir a ser importante.

Muita pesquisa social interessante envolve medir orientações, como o preconceito, muitas vezes não reconhecidas pelos próprios respondentes. Num *survey*, você pode fazer diversas perguntas que, combinadas, permitem descrever os respondentes como mais ou menos preconceituosos. Em muitos casos, eles não entendem a finalidade latente das perguntas e podem muito bem discordar das suas caracterizações na análise.

A pesquisa de *survey* não permite medida direta do comportamento, embora freqüentemente o comportamento social seja a referência última da pesquisa social. Entretanto, a pesquisa de *survey* permite medida *indireta* do comportamento, muitas vezes de maneiras úteis.

Pode-se pedir aos respondentes para informar seu comportamento passado. Qual sabonete compraram recentemente? Foram à igreja na semana passada? Em quem votaram na última eleição? É claro que estas perguntas estão sujeitas aos problemas de lembrar e ser honesto. Por exemplo, eles podem não lembrar em quem votaram, digamos, na eleição presidencial de 1960. Podem permitir que a memória os engane, particularmente se o comportamento em questão é visto hoje como bom ou mau. (Pesquisas de opinião pós-eleitorais muitas vezes mostram maior porcentagem do eleitorado votando no vencedor do que o que realmente aconteceu.) No entanto, apesar destas desvantagens, relatos de comportamentos passados podem ser muito úteis.

A pesquisa de *survey* também pode examinar comportamento prospectivo, real ou hipotético. Assim, pesquisas de opinião perguntam "Em quem você votará?" Às vezes, é útil criar situações hipotéticas e perguntar aos respondentes como agiriam: "Se seu partido apresentasse uma mulher para presidente, você acha que votaria nela?"

Na maioria das vezes, medidas de comportamento prospectivo são menos confiáveis do que as de comportamento passado. De qualquer forma, tenha em mente essas limitações, quando estiver gerando tais dados. Ao perguntar aos respondentes se votariam numa mulher indicada para presidente pelo seu partido, provavelmente você aprenderá pouca coisa de valor para uma futura candidata feminina. Os dados descritivos produzidos por respostas a uma pergunta destas são provavelmente pouco confiáveis. Entretanto, ao mesmo tempo, os dados gerados pelo item podem ser úteis para descrever os respondentes como mais ou menos antifeministas.

Níveis de Medição

Vimos na seção precedente que o pesquisador de *survey* pode medir diversas variáveis sociais. Nesta seção, olharemos para estas variáveis de outra perspectiva, considerando quatro níveis de medição: nominal, ordinal, intervalo e razão.⁵

Medição Nominal

A medição *nominal* apenas distingue as categorias que compõem uma variável. Sexo, por exemplo, é uma variável composta das categorias masculino e feminino. Outros exemplos são filiação religiosa ou política, região do país e tipo de curso superior.

As categorias compondo uma variável nominal são mutuamente excludentes, não tendo qualquer outra relação entre si. Os níveis restantes de medição, discutidos abaixo, refletem relações adicionais entre as categorias.

Medição Ordinal

A medição *ordinal* reflete uma ordenação entre as categorias que compõem uma variável. Classe social é um exemplo de variável ordinal, composta, talvez, das categorias classe baixa, classe média e classe alta. Outros exemplos são religiosidade, alienação e anti-semitismo.

A medição ordinal é muito usada na pesquisa social científica. Embora tais medidas sejam muitas vezes representadas por números num índice ou escala, estes números não têm outro significado além da indicação da ordem. Assim, uma pessoa com nota 5 numa escala de alienação é *mais alienada* do que uma pessoa com nota 4 neste índice, embora esta medida ordinal não dê qualquer indicação sobre *quão mais alienada* ela é.

Medição de Intervalo

A medição de *intervalo* também usa números para descrever condições, mas eles têm mais sentido do que na medição ordinal, porque as distâncias entre os pontos têm significado real. O exemplo mais comum é a escala de temperatura Fahrenheit. A diferença entre 80 e 90 graus é a mesma que a diferença entre 60 e 70 graus. (Veja que a diferença entre as notas 4 e 5 num índice de alienação não é necessariamente a mesma que a diferença entre 3 e 4.)

Medição de Razão

A medição de *razão* tem as mesmas características da de intervalo, com a característica adicional de ter um *zero real*. Em comparação com a escala Fahrenheit (uma medida de intervalo), a escala de temperatura Kelvin (baseada em graus de temperatura de zero absoluto) é uma medida de razão: enquanto 40 graus Fahrenheit não são necessariamente duas vezes mais quentes do que 20 graus Fahrenheit, 200 Kelvins são duas vezes mais quentes do que 100 Kelvins.

No contexto da pesquisa social científica, a idade é um exemplo de medida de razão. Uma pessoa com vinte anos tem o dobro da idade de alguém com dez anos. Altura, peso e tempo de moradia numa cidade são outros exemplos.

Implicações dos Níveis de Medição

Na discussão sobre a análise de *survey* na Parte 3 do livro, veremos que diferentes técnicas analíticas podem requerer níveis específicos de medição. Por exemplo, se você está analisando a relação entre duas variáveis nominais, algumas técnicas analíticas são inadequadas.

Ao mesmo tempo, é importante compreender que uma variável pode ser tratada de forma diferente em termos dos níveis de medição. Lembre, por exemplo, que idade é uma medida de razão. Portanto, você pode calcular uma equação de regressão associando idade e altura (outra medida de razão). Mas, suponha que você está estudando uma amostra de estudantes variando de quatorze a vinte anos. Dentro desta faixa limitada, você pode aproveitar os intervalos iguais entre as idades, mas não o caráter de razão da variável. Em outro tipo de estudo, você pode escolher juntar idades nas seguintes categorias: menos de quarenta anos = *jovem*; quarenta a sessenta anos = *meia idade*; acima de sessenta = *idoso*. Finalmente, idade pode ser convertida em medida nominal para certos fins. Respondentes podem ser categorizados como membros do “*baby boom*” pós II Guerra Mundial ou não; podem ser categorizados como tendo nascido durante a Depressão dos anos 30 ou não. (Outra medida nominal — baseada na data do nascimento e não na idade — é a classificação por signos astrológicos.)

Ao desenhar o questionário, considere o tipo de análise que fará após a coleta dos dados. Você deve determinar se e como pode medir cada variável, de forma a permitir a análise

requerida. Se a análise requer dados na forma de medidas de razão, você não deve construir seus questionários de forma a criar apenas variáveis nominais.

Guias para a Elaboração de Questões

Ao construir itens do questionário, há várias opções. A experiência passada dos pesquisadores de *survey* fornece muitas diretrizes que podem ajudá-lo a gerar dados úteis para análise. Esta seção lida com os dois temas.

Questões e Declarações

A pesquisa de *survey* é habitualmente vista como fazer perguntas, mas o exame de um *survey* típico provavelmente mostrará tantas declarações quanto perguntas. Isto tem suas razões. Muitas vezes você quer determinar o quanto os respondentes apóiam determinada atitude ou perspectiva. Se for capaz de resumir a atitude numa declaração curta, você pode apresentá-la aos respondentes e perguntar-lhes se concordam ou discordam dela. Rensis Likert formalizou este procedimento com a *escala Likert*, um formato no qual solicita-se aos entrevistados “concordar fortemente”, “concordar”, “discordar”, “discordar fortemente” ou “aprovar fortemente”, “aprovar” etc.

Questões e declarações podem ser usadas vantajosamente numa pesquisa de *survey*. O uso de ambas dá mais flexibilidade ao desenho dos itens e pode tornar o questionário mais interessante.

Perguntas Abertas e Fechadas

Há duas opções para fazer perguntas. Você pode fazer perguntas *abertas*, solicitando aos respondentes darem suas próprias respostas. Por exemplo, pode perguntar “Qual tema você considera mais importante para o país hoje?”, dando espaço para escrever a resposta (ou solicitando resposta verbal ao entrevistador).

Nas perguntas *fechadas*, pede-se para respondentes escolherem uma alternativa numa lista apresentada. Perguntas fechadas são muito populares, porque dão maior uniformidade de respostas e são mais facilmente processadas. Respostas abertas devem ser codificadas antes de entrar com os dados, e há o perigo de alguns respondentes darem respostas irrelevantes para a intenção do pesquisador. Por outro lado, muitas vezes pode-se entrar com as

respostas fechadas diretamente a partir do questionário e, em alguns casos, podem ser marcadas diretamente em folhas para sensores óticos pelos próprios respondentes, para entrada automática dos dados.

A principal desvantagem das perguntas fechadas está na estruturação das respostas. Quando as respostas relevantes a uma questão são relativamente claras, estruturar as respostas pode não apresentar problema. Entretanto, em outros casos, a estruturação das respostas pode deixar passar algumas que são importantes. Por exemplo, ao perguntar sobre “as questões mais importantes com as quais o país se depara”, você pode apresentar uma lista delas, mas deixar passar questões que os respondentes consideram importantes. (Lembre nossa afirmação anterior de que os dados são criados e não coletados.)

Duas diretrizes devem sempre ser seguidas ao construir questões fechadas. Primeiro, as categorias de respostas propostas devem ser *exaustivas*, ou seja, devem incluir todas as respostas possíveis que se pode esperar. Por isso, os pesquisadores muitas vezes acrescentam a categoria “Outros (especifique por favor)”. Ao fazê-lo, lembre que os respondentes tentarão encaixar suas respostas pessoais numa das categorias fornecidas, mesmo que o encaixe não seja perfeito.

Segundo, as categorias de respostas devem ser *mutuamente excludentes*, ou seja, os respondentes não devem se sentir forçados a escolher mais de uma resposta. (Às vezes, o pesquisador pode querer respostas múltiplas, mas estas respostas criarão dificuldades no processamento.) Você pode garantir respostas mutuamente excludentes considerando cuidadosamente cada combinação de respostas e indagando se uma pessoa poderia, razoavelmente, escolher ambas. Muitas vezes, você pede ao respondente para “selecionar a *melhor* resposta”, mas esta técnica não deve ser usada para compensar um conjunto de respostas malpensadas.

Tornando os Itens Claros

Os itens de questionários devem ser claros e não ambíguos, e a proliferação de perguntas pouco claras e ambíguas encontradas em muitos *surveys* torna necessário enfatizar este ponto. Muitas vezes, você fica tão envolvido com o tema estudado que opiniões e perspectivas ficam claras para você, mas não para seus respondentes, muitos dos quais terão dado pouca ou nenhuma atenção ao tema.

Por outro lado, você pode ter uma compreensão superficial do tema, deixando de especificar suficientemente a intenção da pergunta. A questão “O que você pensa sobre o sistema proposto de mísseis antibalísticos?” pode evocar no respondente a contrapergunta: “Qual sistema proposto de mísseis antibalísticos?” Itens de questionário devem ser precisos, de forma que o respondente saiba exatamente qual pergunta espera-se que ele responda.

Evitando Questões Duplas

Com freqüência, os pesquisadores pedem uma resposta única para uma combinação de perguntas. Esta situação parece ocorrer com mais freqüência quando o pesquisador se identifica pessoalmente com uma posição complexa. Por exemplo, você pode pedir aos respondentes para concordarem ou discordarem da declaração “Os Estados Unidos devem abandonar seu programa espacial e gastar dinheiro com programas domésticos”. Embora muitas pessoas concordem e outras discordem de forma inequívoca desta declaração, outras não conseguem responder. Alguns podem querer abandonar o programa espacial e devolver o dinheiro aos contribuintes. Outros podem querer continuar com o programa espacial e também aplicar mais dinheiro em programas domésticos. Estes últimos não podem concordar nem discordar da declaração.

Como regra geral, sempre que a palavra *e* aparece numa pergunta ou declaração, você deve checar se não está fazendo uma pergunta dupla.

Garantindo a Competência dos Entrevistados em Responder

Ao pedir informações, você deve sempre se perguntar se os respondentes são capazes de fazê-lo de forma confiável. Num estudo sobre a criação de crianças, você pode pedir aos respondentes para informarem a idade em que as crianças primeiro “responderam” aos pais. Além do problema de definir o que é “responder aos pais”, é duvidoso que a maioria se lembre disso com precisão.

Em outro exemplo, líderes estudantis governamentais ocasionalmente pedem aos seus constituintes para indicar como as mensalidades dos alunos devem ser gastas. Tipicamente, pede-se aos respondentes para indicar qual porcentagem dos fundos disponíveis deve ser gasta numa longa lista de atividades. Sem um bom conhecimento da natureza destas atividades e dos custos de cada uma, eles não têm como dar respostas com

sentido. ("Custos administrativos" receberão pouco apoio, embora esta categoria seja essencial para o programa como um todo.)

Um grupo de pesquisadores que pesquisava a experiência ao volante de adolescentes insistia em fazer uma pergunta aberta quanto ao número de quilômetros dirigidos desde o momento da habilitação. Embora os consultores advertissem que poucos motoristas estimariam esta informação com alguma precisão, a pergunta foi feita assim mesmo. Alguns adolescentes afirmaram haver dirigido centenas de milhares de quilômetros.

Fazendo Perguntas Relevantes

Da mesma forma, as perguntas devem ser relevantes para a maioria dos respondentes. Quando se pede atitudes sobre um tema sobre o qual poucos pensam ou se importam, os resultados têm pouca possibilidade de serem úteis. Ademais, os respondentes podem expressar atitudes mesmo nunca tendo pensado sobre o assunto, e você corre o risco de se equivocar.

Este ponto é evidenciado quando pesquisadores pedem respostas sobre pessoas e temas fictícios. Numa pesquisa de opinião que fiz para um candidato, perguntei aos respondentes se conheciam quinze figuras políticas da comunidade. Como exercício metodológico, inventei um nome: Tom Sakumoto. Nove por cento dos respondentes disseram que o conheciam. Destes, metade relatou tê-lo visto na televisão e lido sobre ele nos jornais.

Quando se obtêm respostas a temas fictícios, você pode desconsiderá-las. Mas quando o tema é real, você pode não ter como diferenciar respostas que refletem atitudes genuínas de respostas sem sentido para uma questão irrelevante.

O grau em que os respondentes inventam respostas é até certo ponto controlável pelo pesquisador. Um estudo por Bishop, Tuchfarber e Oldendick o demonstra.⁶ Perguntou-se aos respondentes uma de três questões sobre se o "Ato sobre Questões Públicas de 1975" devia ser revogado. Ao primeiro grupo, a pergunta foi "Você concorda ou discorda desta idéia, ou você não pensou muito sobre este assunto?" Apenas 3,7% deram uma opinião sobre a revogação desta lei fictícia. O segundo grupo recebeu a mesma pergunta, mas não lhes foi dada a opção de dizer que não tinham tido chance para pensar sobre isso; 24,3% expressaram uma opinião. No grupo final, os respondentes não tiveram a possibilidade de fugir da resposta, e os que disseram não ter opinião foram pressionados pelos entrevistadores a dar uma resposta; 31,5% expressaram uma opinião.

Usando Itens Curtos

Querendo ser não ambíguo, preciso e sublinhar a relevância de um tema, muitas vezes você é levado a formular itens longos e complicados. Isto deve ser evitado. É irrelevante, para respondentes que não estudaram o item profundamente, que a intenção de um item seja clara, se for estudada com cuidado. O respondente deve poder ler um item rapidamente, entender sua intenção, e escolher ou dar uma resposta sem dificuldade. Em geral, você deve supor que os entrevistados *irão* ler os itens rapidamente e dar respostas rápidas; portanto, você deve fornecer itens claros e curtos que não serão mal interpretados sob tais condições.

Evitando Itens Negativos

Uma negação num item de questionário abre caminho para uma fácil má interpretação. Solicitados a concordarem ou discordarem da declaração "Os Estados Unidos não devem reduzir seus arsenais de armas nucleares", grande parte dos respondentes passará por cima da palavra *não*. Assim, alguns concordarão com a declaração, mesmo sendo a favor da redução de armas nucleares, enquanto outros concordarão por opor-se a ela. E você talvez jamais saiba quais respostas refletem esta confusão.

Num estudo sobre liberdades civis, perguntou-se aos respondentes se sentiam que "os seguintes tipos de pessoas devem ser proibidos de lecionar em escolas públicas", de uma lista incluindo comunistas, membros da Ku Klux Klan, e assim por diante. As categorias de resposta "sim" e "não" foram colocadas ao lado de cada item. Uma comparação das respostas a este item com respostas a outros itens refletindo apoio às liberdades civis sugeriu enfaticamente que muitos respondentes davam a resposta *sim* para indicar que tal pessoa podia dar aulas em vez de indicar que ela devia ser "proibida". (Um estudo subsequente que dava como categorias de respostas "permitir" e "proibir" produziu resultados muito mais claros.)

Evitando Itens e Termos "Tendenciosos"

Tenho repetido que dados de *survey* são criados e não apenas coletados. Portanto, a maneira como os dados são procurados determina a natureza dos dados recebidos. Você deve estar sempre atento ao efeito da redação das perguntas sobre os resultados que obterá.

A maioria dos pesquisadores reconheceria o efeito provável de uma pergunta começando com “Você não concorda com o presidente dos Estados Unidos na crença de que...” e nenhum pesquisador de reputação usaria tal item. Infelizmente, o efeito tendencioso de itens e termos é muito mais sutil do que sugere este exemplo.

A mera identificação de uma atitude ou posição com uma pessoa ou agência de prestígio pode enviesar as respostas. O item “Você concorda ou discorda da proposta do presidente de...” teria este efeito. “Você concorda ou discorda com a recente decisão da Suprema Corte de que...” teria um efeito similar. Isso não quer dizer que esta redação produziria necessariamente um consenso ou mesmo uma maioria em apoio à posição identificada com a pessoa ou agência de prestígio, mas apenas que o apoio provavelmente seria maior do que o que seria obtido sem esta identificação.

Itens de questionários podem ser negativa ou positivamente tendenciosos. “Você concorda ou discorda da posição de Adolf Hitler quando declarou que...” é um exemplo de viés negativo. Muitas vezes o uso de termos como liberal, conservador, comunista, ateu e outros introduz um viés não intencional, embora tais termos às vezes sejam essenciais e apropriados para a intenção da pergunta (por exemplo, “Como você se descreveria politicamente: liberal, conservador, de centro ou alguma outra coisa?”) Da mesma forma, Tom Smith⁷ demonstrou que as pesquisas de opinião pública informam apoio muito maior para assistência governamental a “pessoas pobres” do que para “pessoas recebendo benefícios”. O termo benefícios [*welfare*] tem uma carga muito negativa na opinião pública americana.

Como em outros exemplos, você deve examinar cuidadosamente o objetivo da sua investigação e construir itens que serão mais úteis para ele. Você jamais deve ser levado a pensar que existem, em última análise, formas “certas” ou “erradas” de fazer as perguntas.

Qualidade das Medições

Antes de nos voltarmos para as técnicas específicas de medição, consideremos alguns critérios gerais sobre a qualidade das medições. Para começar, pode-se fazer medições com graus variados de *precisão*, um termo que representa a qualidade das distinções feitas entre os atributos de uma variável. A descrição de uma mulher como “tendo quarenta e três anos” é mais

precisa do que “nos seus quarenta”. Dizer “11,75 polegadas de comprimento” é uma descrição mais precisa do que dizer “mais ou menos um pé de comprimento”.

Como regra geral, medidas precisas são superiores a medidas imprecisas, como dita o senso comum. Não há condições em que medições imprecisas sejam superiores a medições precisas. Mas a precisão nem sempre é necessária ou desejável. Se para o objetivo de pesquisa basta saber que uma mulher está nos seus quarenta, qualquer esforço adicional para descobrir a idade precisa é inútil. Portanto, a operacionalização dos conceitos deve ser guiada parcialmente pela compreensão do grau de precisão requerido. Se suas necessidades não são claras, seja mais e não menos preciso.

Entretanto, não confunda a precisão com a *exatidão*. Descrever alguém como “nascido em Stowe, Vermont” é mais preciso do que “nascido na Nova Inglaterra”, mas suponha que a pessoa em questão de fato tenha nascido em Dover, New Hampshire. Neste caso, a descrição menos precisa seria mais exata, ou seja, refletiria melhor o mundo real.

Precisão e exatidão são qualidades obviamente importantes nas medições de pesquisa e provavelmente não precisam de mais explicações. Mas, quando cientistas sociais constroem e avaliam medições, prestam atenção especial a duas considerações técnicas: confiabilidade e validade.

Confiabilidade

Confiabilidade é o problema de uma determinada técnica, ao ser aplicada repetidamente a um mesmo objeto, produzir, a cada vez, os mesmos resultados. Suponha que eu lhe pedisse para avaliar meu peso. Você me olharia com cuidado e adivinharia que tenho 80 quilos. Agora, suponha que eu lhe peça para avaliar o peso de trinta ou quarenta outras pessoas, e enquanto você estivesse fazendo isso, eu entrasse na fila usando um disfarce e, quando chegasse a minha vez, você adivinhasse 90 quilos. Te peguei! Este pequeno exercício teria demonstrado que fazer estimar o peso das pessoas não é uma técnica confiável.

Entretanto, suponha que eu lhe emprestasse minha balança de banheiro. Não importa o meu disfarce, você presumivelmente anunciaria o mesmo peso para mim nas duas vezes, mostrando que o uso da balança fornece uma medida mais confiável de peso do que a adivinhação. Mas, tal como a precisão, a confiabilidade não garante a exatidão. Suponha que eu calibrasse

a balança para tirar 3 quilos do meu peso apenas para me sentir melhor. Embora você (confiavelmente) informasse o mesmo peso cada vez, sempre estaria errado.

Suponha que estejamos interessados em saber se as pessoas preferem a cerveja A ou a cerveja B. Podemos, é claro, simplesmente perguntar-lhes qual delas preferem — provavelmente a melhor solução. Mas suponha que desejássemos dados mais “objetivos”. Podemos pedir aos respondentes para dizer quantas As e quantas Bs beberam até hoje. Se você pensar bem, verá que esta é uma técnica pouco confiável. Dificilmente alguém saberá dar esta informação. Da mesma forma, não é provável que as pessoas saibam quantas vezes já foram à igreja, quantos quilômetros já dirigiram etc.

Fazer perguntas que as pessoas acham também irrelevantes provavelmente levará a respostas não confiáveis. Por exemplo, a maioria das pessoas desconhece qual a política externa da China em relação à Albânia, e nem se importam.

Problemas de confiabilidade surgem em vários pontos da pesquisa de *survey*. Como veremos no Capítulo 10, entrevistadores diferentes às vezes obtêm respostas diferentes dos entrevistados como resultado das suas próprias atitudes e aparência. Igualmente, codificadores diferentes podem codificar as mesmas respostas abertas de forma diferente. Por exemplo, se formos codificar algumas centenas de ocupações específicas em termos de algum esquema padronizado de codificação, digamos, um conjunto de categorias criado pelo Ministério do Trabalho ou pelo Bureau de Recenseamento, você e eu não codificaríamos todas aquelas ocupações nas mesmas categorias.

Como se criam medições confiáveis? Há diversas técnicas para medir a confiabilidade de itens de questionários, mas os métodos para *maximizar* a confiabilidade são bastante diretos. Faça apenas as perguntas cujas respostas as pessoas provavelmente sabem, pergunte coisas relevantes para elas, e seja claro no que está perguntando. O perigo é que as pessoas *darão* respostas — confiáveis ou não. As pessoas lhe contarão o que acham da relação entre a China e a Albânia mesmo não tendo a menor idéia de qual seja ela.

Validade

No uso convencional, a *validade* se refere ao grau com que uma medida empírica reflete adequadamente o *significado real* do conceito considerado. A primeira consideração

é sobre a *validade aparente*. Certas medidas empíricas podem ou não coincidir com nossas convenções e imagens mentais associadas a um conceito. Podemos discutir quanto à adequação de medir religiosidade através de comparecimento à igreja, mas concordaríamos que ir à igreja tem pelo menos alguma relevância para a religiosidade.

Segundo, os pesquisadores chegaram a acordos concretos sobre a melhor forma de medir alguns conceitos. O Bureau do Recenseamento, por exemplo, criou definições operacionais de conceitos como família, residência e status empregatício que parecem ter validade operacional na maioria dos estudos que usam aqueles conceitos.

Edward Carmines e Richard Zeller⁸ discutem três tipos adicionais de validade: validade relacionada a critério, validade de conteúdo e validade de construção. A *validade relacionada a critério* às vezes é chamada de *validade preditiva* e baseia-se em algum critério externo. Por exemplo, a validade do conselho de faculdade é evidenciada por sua capacidade de prever o sucesso dos alunos. A validade de um teste escrito para motoristas é evidenciada pela relação entre as notas que as pessoas obtêm no teste e sua competência para dirigir. Nestes exemplos, sucesso escolar e competência ao volante são os *critérios*. Em geral, o comportamento pode servir como padrão da validade de critério para muitas medidas de atitude em pesquisa social (por exemplo, pessoas “preconceituosas” de fato discriminam as minorias?), embora a relação entre atitudes e comportamento seja também assunto importante de estudo por si própria.

A *validade de conteúdo* refere-se ao grau com que uma medição cobre a amplitude de significados incluídos no conceito. Por exemplo, um teste de capacidade matemática não pode se limitar apenas à adição, também precisando cobrir a subtração, a multiplicação, a divisão etc. Em outro exemplo, se dizemos que estamos medindo preconceito em geral, nossas medidas refletem o preconceito contra grupos raciais e étnicos, minorias religiosas, mulheres, idosos etc.?

Finalmente, a *validade de construção* baseia-se no modo como uma medida se relaciona a outras variáveis num sistema de relações teóricas. Por exemplo, suponha que você quer estudar as fontes e conseqüências da “satisfação conjugal”, desenvolva uma medida de satisfação conjugal e deseje avaliar sua validade. Além da medida, você terá desenvolvido também certas expectativas teóricas sobre como a variável satisfação conjugal se relaciona a outras. Por exemplo, você pode ter concluído

que maridos satisfeitos têm menos probabilidade de espancaram suas mulheres do que maridos não satisfeitos. Se sua medida de satisfação conjugal se relaciona ao espancamento de mulheres da forma esperada, este resultado constitui evidência da validade de construção da sua medida. Mas, se maridos “satisfeitos” e “insatisfeitos” tiverem igual probabilidade de espancar suas mulheres, este resultado desafiará a validade da sua medida.

Tensão entre Confiabilidade e Validade

Com rodapé destas discussões, deve ser assinalado que muitas vezes há uma certa tensão entre os critérios de confiabilidade e validade. Muitas vezes parece que fazemos uma troca entre as duas.

A maioria dos conceitos realmente interessantes que queremos estudar tem muitas nuances sutis, e é difícil especificar precisamente o que queremos dizer com eles. Os pesquisadores muitas vezes falam desses conceitos como tendo “riqueza de significado”. Se você duvida, tente criar definições satisfatórias e medidas para conceitos como preconceito, satisfação, alienação, religiosidade e liberalismo.

Entretanto, a ciência precisa ser específica para poder gerar medições confiáveis. Portanto, com muita frequência, a especificação de definições operacionais e medições confiáveis parece roubar de tais conceitos a riqueza de significado. Religiosidade é mais do que freqüentar a igreja; preconceito é mais do que a expressão de uma atitude preconceituosa num questionário. Esta situação é um dilema persistente e inevitável para o pesquisador, e você estará bem preparado para isso se estiver de sobreaviso. Esteja preparado e lide com isso. Se você não chegar a um acordo claro sobre como medir um conceito, meça-o de várias formas diferentes. Se o conceito tiver várias dimensões diferentes, meça todas. Acima de tudo, saiba que o conceito não tem qualquer significado além daquele que você lhe atribuiu, e que a única justificativa para dar um significado particular a um conceito é a utilidade. A meta é medir conceitos de forma a nos ajudar a entender o mundo que nos cerca.

Formato Geral dos Questionários

O formato de um questionário pode ser tão importante quanto a natureza e a redação das perguntas. Um questionário malformatado pode levar os respondentes a saltar perguntas,

confundi-las quanto à natureza dos dados pedidos e, no limite, levá-los a jogá-lo fora. Vamos sugerir diretrizes tanto gerais quanto específicas.

Como regra geral, o questionário deve estar bem distribuído e não amontoado. Você deve maximizar o “espaço em branco”. Pesquisadores inexperientes tendem a temer que seus questionários possam parecer muito longos e, por isso, apertam várias perguntas na mesma linha, abreviam perguntas, e tentam usar o menor número de páginas possível. Tudo isso é desaconselhável e mesmo perigoso. Mais de uma pergunta por linha levará as pessoas a saltarem a segunda. Abreviar perguntas resulta em más interpretações. Em geral, se os respondentes observarem que terminaram rapidamente as primeiras páginas de um questionário que, a princípio, parecia longo, se sentirão menos desmoralizados do que se tiverem gasto um longo tempo na primeira página do que parecia um questionário curto. Ademais, cometerão menos erros e não serão forçados a reler perguntas confusas e abreviadas, nem terão sido forçados a escrever uma resposta longa em espaço pequeno.

Não há como exagerar a valia de espalhar as perguntas num questionário. Questionários espremidos são desastrosos, seja em *surveys* por correio ou por entrevistas.

Formatos para Respostas

Há vários métodos para apresentar uma série de categorias para o respondente marcar em resposta a uma pergunta. Minha experiência diz que *caixas* espaçadas adequadamente são a melhor escolha. Se o questionário for composto em tipografia ou formatado num bom sistema de processamento de textos, as caixas podem ser geradas de forma fácil e limpa. Também é possível fazer caixas numa máquina de escrever.

Se o questionário for datilografado em máquina de escrever com chaves, caixas excelentes podem ser produzidas com um colchete esquerdo, um espaço e um colchete direito: []. Se não houver colchetes, parênteses funcionam razoavelmente bem no mesmo modo: (). Mas você não deve usar barras e sublinhamentos. Primeiro, porque requer mais esforço de datilografia; segundo, porque o resultado não é muito limpo, especialmente se as categorias de resposta tiverem de ser em espaçamento único. A Figura 7-1 mostra uma comparação dos diversos métodos.

O pior método é dar espaços em branco para se assinalar, porque os respondentes muitas vezes farão marcas muito

grandes, tornando impossível ver qual resposta foi dada. Além disso, como regra geral, recomenda-se espaçamento duplo entre as categorias, para evitar estas marcas ambíguas.

Pode-se também considerar um método muito diferente. Em vez de fornecer caixas a serem marcadas, você pode pôr números código ao lado de cada resposta, pedindo ao respondente para fazer um *círculo* em torno do número apropriado. Este método tem a vantagem adicional de especificar o número a ser entrado posteriormente na etapa de processamento de dados. Se usar este método, você deve dar instruções claras e visíveis, uma vez que muitos respondentes ficarão tentados a pôr um "X" no número apropriado, dificultando a entrada de dados. (Nota: A técnica pode ser usada com segurança em *surveys* com entrevistas, já que os entrevistadores podem ser instruídos e testados.)

FIGURA 7-1
Formatos de respostas

<input type="checkbox"/> Sim	<input type="checkbox"/> Sim	1. Sim
<input type="checkbox"/> Não	<input type="checkbox"/> Não	2. Não
<input type="checkbox"/> Não sei	<input type="checkbox"/> Não sei	3. Não sei

Questões Contingentes

Muitas vezes, na pesquisa de *survey*, certas perguntas serão claramente relevantes apenas para um subconjunto de respondentes. Você pode perguntar-lhes se estão familiarizados com certo assunto, e depois perguntar aos que responderem sim o que pensam a respeito. Você pode querer saber o status de recrutamento de jovens, sabendo que apenas homens terão este status. Ou você pode perguntar se o respondente pertence a uma certa organização e, se a resposta for positiva, se já tiveram alguma função.

A segunda pergunta em cada exemplo é denominada *pergunta contingente*, significando que a segunda pergunta é contingente quanto à resposta à primeira. Portanto, a pergunta sobre atitude quanto a determinada questão é contingente à familiaridade do respondente com ela.

O uso apropriado de perguntas contingentes pode facilitar a tarefa de responder um questionário e também melhorar a qualidade dos dados produzidos. A alternativa às perguntas contingentes deve ser evitada pelas mesmas razões. Não pergunte "Se alguma vez pertenceu à APP, você já teve algum cargo nela?" Este formato forçará todos os respondentes a lerem a pergunta, embora seja irrelevante para muitos. Aqueles para quem a pergunta é irrelevante serão forçados a decidir se saltam a pergunta, escrever "não se aplica", responder não, ou jogar fora o questionário. (Esta última escolha é bastante comum.)

Há vários formatos de perguntas contingentes. Acho que o melhor é aquele em que as perguntas contingentes são recuadas no questionário, postas em caixas e conectadas à pergunta base por setas a partir das respostas apropriadas. A Figura 7-2 ilustra uma destas perguntas. Usados adequadamente, conjuntos complexos de perguntas contingentes podem ser construídos sem confundir o respondente, como ilustrado na Figura 7-3.

Embora os exemplos acima se refiram primariamente a questionários auto-administrados, a apresentação apropriada de perguntas contingentes é ainda mais importante em questionários por entrevistas. O respondente num questionário pelo correio pode, mesmo com certa insatisfação, reler a pergunta, mas se o entrevistador ficar confuso e fizer perguntas impróprias, toda a entrevista ficará prejudicada.

FIGURA 7-2

23. Você já fez parte da associação local da APP?

- Sim
 Não

Em caso afirmativo:
Você já teve algum cargo na APP local?

Sim
 Não

FIGURA 7-3

14. Você já ouviu falar do Programa Cidade Modelo?

- Sim
 Não

Em caso afirmativo:

a. Você em geral aprova ou desaprova o programa?

Aprova
 Desaprova
 Sem opinião

b. Você já compareceu a alguma reunião da residente Cidade Modelo?

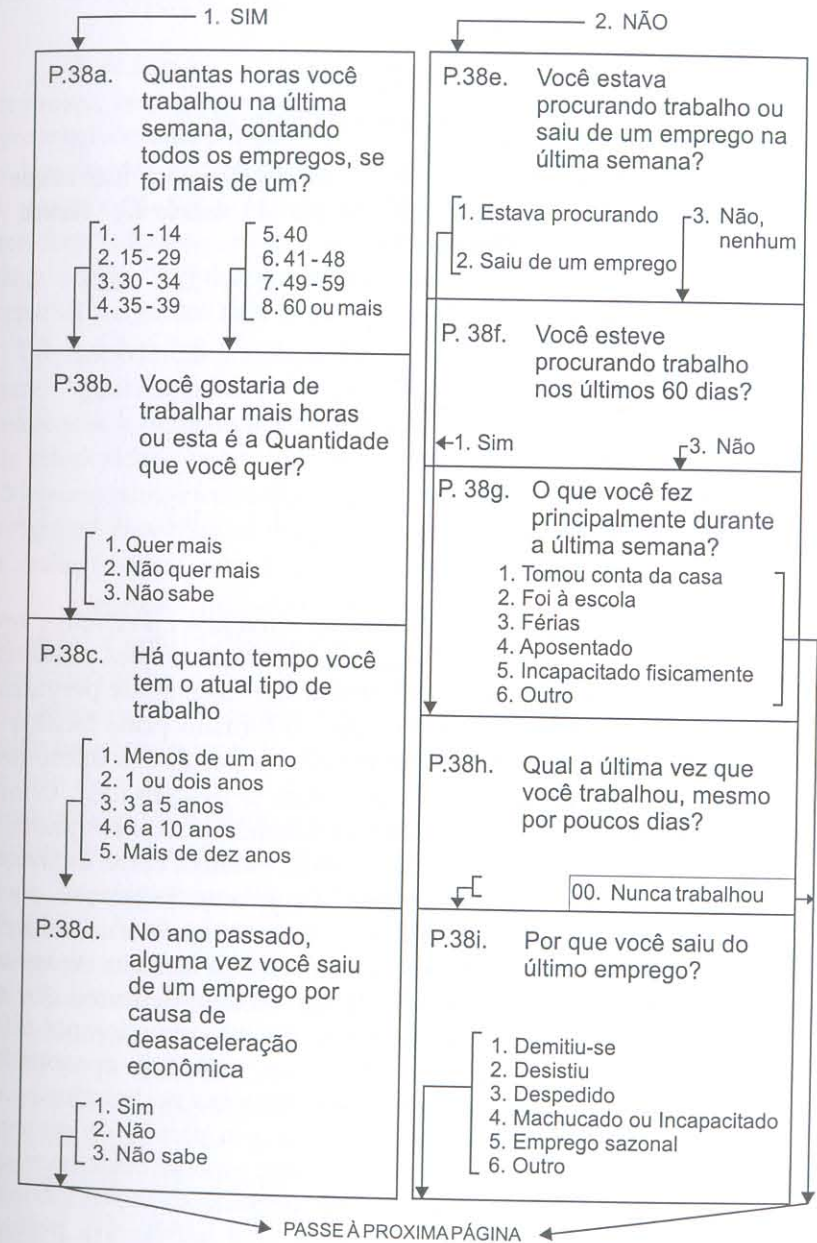
Sim
 Não

Em caso afirmativo:
 Qual foi a última vez que você compareceu a uma reunião?

A Figura 7-4 mostra parte de um questionário de entrevista visando determinar o atual status ocupacional da pessoa. Ler as questões, dar respostas diferentes e seguir as setas apropriadas até a próxima pergunta o orienta sobre a intenção do pesquisador, permitindo que você veja a facilidade que o entrevistador teve para obter a informação relevante de todos os entrevistados.

FIGURA 7-4

P. 38. Você trabalhou alguma vez durante a semana passada? (Inclua trabalho parcial, mas não trabalho em casa.)



Questões Matriciais

Com frequência, você fará várias perguntas com o mesmo conjunto de categorias de respostas. Este é o caso típico sempre que categorias de respostas do tipo Likert forem usadas. Em tais casos, muitas vezes é possível construir uma *questão matricial*, como ilustrado na Figura 7-5.

FIGURA 7-5
Formato de questão matricial

17. Ao lado de cada um dos enunciados apresentados abaixo, por favor indique se você concorda fortemente (CF), concorda (C), discorda (D), discorda fortemente (DF) ou está indeciso (I).

	CF	C	D	DF	I
a. O que este país precisa é de mais lei e ordem.....	[]	[]	[]	[]	[]
b. A polícia deve ser desarmada neste país.	[]	[]	[]	[]	[]
c. Durante revoltas, deve-se atirar no ato em aqueadores..... etc.	[]	[]	[]	[]	[]

Este formato tem numerosas vantagens. Primeiro, usa espaço de forma eficiente. Segundo, os respondentes provavelmente acharão mais rápido completar um conjunto de perguntas apresentadas desta forma. Terceiro, o formato pode facilitar a comparabilidade das respostas dadas a perguntas diferentes tanto para o entrevistado como para o pesquisador. Como os respondentes podem rever rapidamente suas respostas a itens anteriores neste conjunto, podem escolher entre, digamos, “concorda fortemente” e “concorda” com uma declaração, comparando a força da concordância com as respostas anteriores.

Mas há alguns perigos inerentes a este formato. As vantagens podem encorajá-lo a estruturar um item de forma que as respostas se encaixem numa matriz, quando um conjunto diferente, mais idiossincrático, de respostas seria mais apropriado. Além disso, o formato de questões matriciais pode induzir um *padrão-de-resposta* nos respondentes, um padrão de concordância com todas as declarações. Isso seria especialmente provável se o conjunto começasse com várias declarações indicando uma orientação particular (por exemplo, uma perspectiva política liberal), com apenas algumas declarações posteriores indicando a orientação oposta. Os respondentes podem supor que todas as declarações representam a mesma orientação e, lendo rapidamente,

ler errado algumas delas, dando respostas erradas. Você pode reduzir um pouco este problema alternando declarações representando orientações diferentes e fazendo-as curtas e claras.

Ordenando as Questões no Questionário

A ordem na qual são feitas as perguntas pode afetar a resposta, bem como toda a coleta de dados. Por exemplo, a presença de uma pergunta pode afetar as respostas a perguntas subseqüentes. Por exemplo, se são feitas algumas perguntas sobre o perigo da poluição, e uma pergunta subseqüente (aberta) pede aos respondentes para sugerir o que acreditam representar perigo para o país, “poluição” será mais citada do que seria o caso se aquelas perguntas iniciais não tivessem sido feitas.

Se pedirmos aos respondentes para avaliar sua religiosidade geral (“Qual a importância da sua religião para você, em geral?”), as respostas a perguntas posteriores sobre aspectos específicos da religiosidade buscarão consistência com a avaliação anterior. O inverso também é verdade. Se os respondentes ouvem primeiro perguntas específicas sobre aspectos diversos de sua religiosidade, a avaliação geral subseqüente refletirá as respostas anteriores.

Alguns pesquisadores tentam superar este efeito “aleatorizando” a ordem das perguntas — em geral um esforço fútil. Um conjunto “aleatorizado” de perguntas provavelmente parecerá aos respondentes caótico e sem valor. Ademais, terão dificuldade em responder se forem obrigados a mudar continuamente o foco de atenção de um tema para outro. Finalmente, mesmo uma ordenação aleatorizada de perguntas afeta as respostas do modo discutido acima, e você não terá controle sobre este efeito.

A solução mais segura é ter sensibilidade para o problema. Embora você não possa evitar o efeito do ordenamento de perguntas, deve poder estimar este efeito e interpretar os resultados significativamente. Se a ordem das perguntas parece ser uma questão especialmente importante num estudo, você pode fazer mais de uma versão do questionário, com diferentes ordenamentos de perguntas, criando condições para determinar o efeito dele. No mínimo, você deve pré-testar seu questionário usando as formas diferentes.

A ordem desejada das perguntas difere nos questionários auto-administrados e por entrevistas. No primeiro, em geral é melhor começar o questionário com as perguntas mais interessantes. O respondente potencial que corre o olhar casualmente

sobre as primeiras perguntas deve querer respondê-las. Talvez estas questões perguntem coisas sobre as quais os entrevistados estão ansiosos por se expressar. Ao mesmo tempo, contudo, as perguntas iniciais não devem ser ameaçadoras. (Pode ser má idéia iniciar com perguntas sobre comportamento sexual ou uso de drogas.) Em geral, deve-se pedir dados demográficos mais insípidos (idade, sexo etc.) apenas no final dos questionários auto-administrados. Pedi-los no começo, como muitos pesquisadores inexperientes são tentados a fazer, dá ao questionário a aparência inicial de um formulário rotineiro, e quem o recebe pode não encontrar motivação suficiente para terminá-lo.

O oposto em geral é verdade para os *surveys* por entrevistas. Quando o entrevistado em potencial abre a porta ou atende o telefone, o entrevistador deve começar estabelecendo rapidamente a comunicação com o entrevistado. Após breve introdução à pesquisa, o entrevistador deve começar enumerando os membros da residência, pedindo os dados demográficos de cada um. Estas questões são facilmente respondidas e, em geral, pouco ameaçadoras. Após estabelecer a comunicação inicial, o entrevistador pode, então, entrar na área das atitudes e assuntos mais sensíveis. Um *survey* por entrevista começando com a pergunta "Você acredita em Deus?" provavelmente teria um fim rápido.

Instruções

Todo questionário, seja auto-administrado ou aplicado por um entrevistador, deve conter instruções claras e comentários introdutórios onde apropriado.

Instruções Gerais

Todo questionário auto-administrado deve começar com instruções básicas sobre o preenchimento. Embora muitos respondentes estejam familiarizados com técnicas de *survey* normais, é útil começar instruindo-os a indicar as respostas a certas perguntas com uma marca ou um "X" na caixa ao lado da resposta apropriada, ou escrevendo as respostas quando solicitados. Se muitas questões abertas forem usadas, os respondentes devem ser orientados sobre o tamanho esperado das respostas. Se você encorajar respostas escritas, permitindo elaborar respostas a questões fechadas, deve assinalar esta possibilidade.

Se um questionário está ordenado em subseções por conteúdos — por exemplo, atitudes políticas, atitudes religiosas, dados básicos — é útil introduzir cada seção com uma declaração curta sobre seu conteúdo e finalidade. Eis um exemplo de introdução possível: "Nesta seção, gostaríamos de saber quais problemas comunitários as pessoas por aqui consideram os mais importantes." Itens demográficos no fim de um questionário auto-administrado podem ser introduzidos assim: "Finalmente, gostaríamos de saber um pouco sobre você para podermos ver o que diferentes tipos de pessoas acham sobre os temas que estivemos examinando."

Introduções curtas como estas ajudam o respondente a dar sentido ao questionário, e o tornam menos caótico, particularmente quando ele aborda uma variedade de dados. Também ajudam a pôr o respondente no contexto mental apropriado para responder as perguntas.

Instruções Específicas

Algumas perguntas podem requerer instruções especiais para facilitar a resposta adequada, especialmente se uma pergunta difere das instruções gerais ao questionário como um todo. Alguns casos específicos ilustram esta situação.

Apesar de ser desejável ter categorias de respostas mutuamente excludentes em questões fechadas, muitas vezes mais de uma resposta se aplicará a alguns respondentes. Se você quiser uma resposta única, deve deixá-lo bem claro na pergunta. Por exemplo, "Na lista abaixo, marque por favor a razão *principal* da sua decisão de freqüentar a faculdade". Muitas vezes a pergunta principal pode ser seguida de uma nota entre parênteses, por exemplo, "Por favor, marque *a melhor* resposta". Se, por outro lado, você quer que o entrevistado marque tantas respostas quanto for o caso, também deve deixar isto claro.

Quando um conjunto de categorias de respostas tiver de ser ordenado pelo respondente, as instruções devem deixá-lo claro, e um tipo diferente de formato de respostas deve ser usado (por exemplo, espaços vazios ao invés de caixas). As instruções devem indicar quantas respostas devem ser ordenadas (por exemplo, todas, a primeira e a segunda, a primeira e a última, a mais e a menos importante) e qual a ordem de hierarquia (por exemplo, "Coloque um 1 ao lado da mais importante, um 2 ao lado da segunda mais importante, e assim por diante".)

Ordenar uma lista de itens muitas vezes é difícil para os respondentes, porque talvez tenham de ler a lista várias vezes. Além disso, em geral se consegue os mesmos resultados pedindo para *valorar* os itens (por exemplo, "Muito importante", "Importante" etc.) Como assinalam Krosnick e Alwin,⁹ problemas só ocorrem quando os respondentes dão os mesmos valores à maioria dos itens. Naturalmente, como sugerem os autores, dar o mesmo valor a vários itens pode indicar o que eles realmente sentem quanto àqueles itens, e um ordenamento forçado pode produzir respostas não verdadeiras. Mas, se você achar essencial que os respondentes ordenem uma lista de itens, você deve fazer a lista tão curta quanto possível. Como regra geral, não peça para ordenar mais do que cinco itens.

Em questões matriciais de múltiplas partes, você deve dar instruções especiais, salvo se o mesmo formato for usado em todo o questionário. Às vezes, espera-se que os respondentes marquem uma resposta em cada coluna da matriz, em outras vezes que marquem uma resposta em cada linha. Sempre que o questionário contiver os dois tipos de respostas, deve-se dar uma instrução esclarecendo o tipo esperado em cada caso.

Instruções ao Entrevistador

Um questionário confuso auto-administrado pode desmoralizar o respondente, mas um questionário por entrevista confuso pode desmoralizar tanto o entrevistado quanto o entrevistador, pondo em perigo a eficiência deste último. Portanto, instruções suplementares claras aos entrevistadores, fornecidas onde apropriado, são de particular importância.

As instruções que o entrevistador deve ler para os entrevistados e as que não devem ser lidas devem ter formatos diferentes. Por exemplo, as que não devem ser lidas podem estar entre parênteses ou em letras maiúsculas. Uma entrevista pode ser destruída se o entrevistador ler alto algo do tipo "Se o entrevistado é quase analfabeto, então..."

É igualmente importante que um questionário por entrevista tenha um *roteiro verbatim* para o entrevistador ler na entrevista.¹⁰ Em circunstâncias ideais, um entrevistador deve poder conduzir toda a entrevista da introdução inicial ("Olá, meu nome é...") até os comentários finais ("Isso completa a entrevista. Queremos agradecer-lo por...") sem tomar liberdade com uma única palavra. Todas as declarações de transição ao longo do questionário devem ser feitas ("Agora gostaríamos de

passar dos problemas comunitários para os problemas nacionais...") de forma que o roteiro, palavra por palavra, soe natural como uma conversa. O mesmo se aplica à enumeração demográfica dos componentes da residência. Ao invés de instruir o entrevistador a obter a idade de cada membro da família, você deve apresentar uma pergunta padronizada para cada ("Qual era a idade dele no dia primeiro de janeiro de 1990?") O Capítulo 10 enfatiza a importância de se seguir à risca o texto do questionário, mas isso só será possível se o questionário estiver construído adequadamente.

Reprodução do Questionário

Tendo construído o questionário, é preciso fazer cópias suficientes para a coleta de dados. O método para reproduzir os questionários é importante para o sucesso da pesquisa; um instrumento reproduzido com capricho encorajará uma taxa maior de respostas, fornecendo melhores dados.

Há vários métodos para reproduzir questionários, e sua escolha dependerá dos fundos disponíveis, das facilidades locais e do tempo. Reproduções mimeografadas em geral são mais baratas e rápidas, mas com menor qualidade gráfica. Máquinas de fotocopiar variam em velocidade, qualidade e preço, mas pode-se produzir questionários excelentes, especialmente usando um computador e uma impressora laser para o original.

O foto-*offset* de cópia do questionário resulta em melhor qualidade e, além de um certo número de cópias, pode ser mais econômico. Uma matriz em foto-*offset* permite um número incontável de cópias, enquanto uma matriz de mimeógrafo precisa ser refeita depois de algumas centenas de cópias, e a fotocópia mantém o mesmo custo unitário mesmo para muitas cópias.

O melhor método de reprodução do ponto de vista da qualidade profissional é a tipografia. Mas é o método mais caro e pode não ser viável para alguns projetos. Além disso, a tipografia em geral toma mais tempo do que outros métodos. De qualquer forma, você deve investigar as possibilidades locais, equilibrando os valores relativos de tempo, dinheiro e qualidade.

Atualmente, o melhor sistema de reprodução consiste num microcomputador com um processador de textos ou programa de editoração profissional e uma impressora laser para criar o original, em conjunção com o foto-*offset* das cópias.

O questionário pode ser construído de diversas formas. Em alguns casos, é melhor imprimir o questionário numa folha

grande única, dobrável para formar um livreto de painéis desdobráveis. Caso haja muitas páginas no questionário, pode-se prendê-las com um grampeador. O questionário longo de melhor aspecto profissional é um livreto impresso grampeado. Este é o método mais caro.

Uma última preocupação da reprodução do questionário é quantos devem ser pedidos. Para chegar a esta decisão, você deve considerar o tamanho da amostra, o número de cartas de acompanhamento, se houver, num *survey* pelo correio, e a possível necessidade de cópias de discussão, amostras para outros pesquisadores, cópias para inclusão como apêndices em relatórios de pesquisas ou livros código, e assim por diante. Como regra aproximada, você deve estimar o número necessário para a coleta de dados e multiplicar este valor por 1,5 a 2,0, para determinar o número de cópias. Lembre que cópias adicionais produzidas na primeira edição dos questionários serão bem mais econômicas do que as produzidas em edições posteriores.

Resumo

Este capítulo cobriu uma variedade de assuntos, do teórico e filosófico ao técnico e até o mundano. Começou com questões gerais de conceitualização e operacionalização no desenho dos itens do questionário apropriados. Foram dadas algumas diretrizes gerais e específicas para a redação das perguntas e sua apresentação no instrumento de pesquisa. O capítulo concluiu com uma breve discussão sobre o lado técnico da produção de questionários.

A combinação destes temas num único capítulo pode parecer bizarra, mas há um propósito nela. O *valor científico* último de um *survey* pode depender tanto da maneira como o questionário é reproduzido quanto da imaginação na operacionalização de conceitos. Os resultados da pesquisa de *survey* são o produto de muitos passos, indo do teórico ao mundano, e uma fraqueza em qualquer um deles ameaça o todo.

Notas

¹ WHITEHEAD, Alfred North. *Science and the Modern World*. New York: The Macmillan Company, 1925. p.56.

² Citado em SELTZ, Claire et al. *Research Methods in Social Relations*. New York: Holt, Rinehart & Winston, 1959. p.564.

³ GLOCK, Charles Y., STARK, Rodney. *Religion and Society in Tension*. Chicago: Rand McNally, 1965. p.18-38.

⁴ BABBIE, Earl. *Science and Morality in Medicine*. Berkeley: University of California Press, 1970.

⁵ Encontra-se um tratamento excelente deste tema em DAVIS, James A. *Elementary Survey Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1971. p.9 et seq. A maioria dos tratamentos destes níveis de medição usam o termo *escala* (ex., *escala nominal*), seguindo o exemplo de STEVENS, S. S. On the Theory of Scales of Measurement. *Science*, v.103, p.677-680, 1946. Como este livro usa o termo *escala* em sentido diferente (ver Capítulo 8), evitamos seu uso neste contexto.

⁶ BISHOP, George F., TUCHFARBER, Alfred J., OLDENDICK, Robert W. Opinions on Fictitious Issues: The Pressure to Answer *Survey* Questions. *Public Opinion Quarterly*, v.50, p.240-250, verão 1986.

⁷ SMITH, Tom. That Which We Call Welfare by Any Other Name Would Smell Sweeter. *Public Opinion Quarterly*, v.51, p.75-83, primavera 1987.

⁸ CARMINES, Edward G., ZELLER, Richard A. *Reliability and Validity Assessment*. Beverly Hills, CA: Sage, 1979.

⁹ KROSNICK, Jon A., ALWIN, Duane F. A Test of the Form-Resistant Correlation Hypothesis: Ratings, Rankings, and the Measurement of Values. *Public Opinion Quarterly*, v.52, p.526-538, inverno 1988.

¹⁰ Um roteiro não seria apropriado para uma entrevista não estruturada, num estudo exploratório. Neste caso, o entrevistador deve receber um esboço dos temas a cobrir, e talvez um conjunto de questões possíveis.

Leituras Adicionais

CARMINES, Edward G., ZELLER, Richard A. *Reliability and Validity Assessment*. Beverly Hills, CA: Sage, 1979.

GOULD, Julius, KOLB, William. *A Dictionary of the Social Sciences*. New York: Free Press, 1964.

KAPLAN, Abraham. *The Conduct of Inquiry*. San Francisco: Chandler Publishing Co., 1964.

MILLER, Delbert. *Handbook of Research Design and Social Measurement*. New York: Longman, 1983.

OPPENHEIM, A. N. *Questionnaire Design and Attitude Measurement*. New York: Basic Books, 1966.

PAYNE, Stanley. *The Art of Asking Questions*. Princeton, NJ: Princeton University Press, 1965.

SCHUMAN, Howard, PRESSER, Stanley. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context*. New York: Academic Press, 1981.

SMITH, Tom W. The Art of Asking Questions. *Public Opinion Quarterly*, v.51, p.S95-S108, inverno 1987.

Capítulo 8

Construção de Índices e Escalas

Como notamos antes, muitas, se não a maioria, das pesquisas sociais visam determinar as associações entre variáveis. Geralmente, queremos estabelecer que X está relacionado a (ou causa) Y. Mas notamos também que a medição de variáveis, freqüentemente, é tarefa difícil. Normalmente, é impossível chegar a uma medida totalmente inequívoca e completamente aceitável de qualquer variável. Mesmo assim, os pesquisadores não desistem de tentar criar medidas cada vez melhores e mais úteis.

Este capítulo aborda a questão da medição. Especificamente, discute a construção de índices e escalas como medidas de variáveis. Estas medidas cumulativas são freqüentemente usadas na pesquisa social, por vários motivos. Primeiro, apesar do cuidado ao construir questionários, raramente é possível chegar a uma única pergunta que represente adequadamente uma variável complexa. Qualquer item único pode desprestatar alguns respondentes no estudo. Para medir religiosidade, talvez apenas freqüência à igreja não seja um indicador suficiente. Alguns respondentes que freqüentam regularmente a igreja podem ser considerados irreligiosos em outros termos, e alguns que nunca vão à igreja podem ser considerados religiosos. Índices e escalas combinam vários itens de questionário, evitando assim os vieses inerentes aos itens únicos.

Segundo, você pode querer usar uma medida ordinal de uma variável bem refinada — por exemplo, arranjando os respondentes em várias categorias ordinais, da mais baixa à mais alta da variável. Um único item pode não apresentar um número

suficiente de categorias de respostas para permitir esta amplitude de variação. Um índice ou escala, formados a partir de vários itens, pode.

Finalmente, índices e escalas são dispositivos *eficientes* para a análise de dados. Se um único item de questionário nos dá somente uma localização rudimentar de um respondente numa variável, vários itens podem dar uma medida mais exata e abrangente. Mas, normalmente, é impraticável considerar simultaneamente todas as respostas particulares dadas por um respondente. Índices e escalas (especialmente escalas) são dispositivos de *redução de dados*; as várias respostas de um respondente podem ser resumidas num único escore, e mesmo assim os detalhes específicos daquelas respostas serem mantidos quase que na totalidade.

Índices Versus Escalas

Os termos *índice* e *escala* são usados habitualmente de modo impreciso e intercambiável na literatura da pesquisa social. As definições técnicas originais destes termos perderam atualmente seus significados. Antes de considerarmos as distinções entre índices e escalas, vejamos o que eles têm em comum.

Ambos são, tipicamente, medidas *ordinais* de variáveis. São construídos de maneira a ordenar os respondentes do *survey* (ou outras unidades de análise) em termos de variáveis específicas, como religiosidade, alienação, situação socioeconômica, preconceito, sofisticação intelectual etc. O escore de um respondente numa escala ou índice de autoritarismo, por exemplo, dá uma indicação do seu autoritarismo relativo, quando comparado a outros respondentes.

No nosso uso dos termos, escalas e índices são medidas compostas de variáveis, ou seja, medições baseadas em respostas a mais de um item do questionário. Portanto, os escores dos respondentes num índice ou escala de religiosidade são determinadas pelas respostas específicas a vários itens do questionário, cada um dos quais fornecendo alguma indicação da religiosidade deles.

Neste livro, distinguiremos entre índices e escalas pela maneira como os escores são atribuídos aos respondentes. Um *índice* é construído pelo somatório simples de escores atribuídos a respostas específicas aos itens individuais que formam o índice. Uma *escala* é construída pela atribuição de escores a *padrões de respostas* entre os vários itens que formam a escala. Uma escala difere de um índice por tirar vantagem de qualquer possível

estrutura de intensidade entre os itens individuais. Um exemplo simples deve elucidar esta distinção.

Suponha que queiramos medir o apoio dos americanos aos direitos civis de comunistas declarados. Podemos perguntar se um comunista devia ter o direito de ter as seguintes ocupações: (1) advogado, (2) médico, (3) bibliotecário, (4) engenheiro, (5) repórter, (6) professor. Alguns respondentes estão dispostos a permitir que comunistas exerçam todas as ocupações listadas. Alguns não quererem permitir comunistas em nenhuma delas. Outros acham que algumas ocupações podem e outras não. Respondentes que dão respostas mistas presumivelmente indicam que consideram algumas ocupações mais importantes que outras. Contudo, as prioridades relativas das diferentes ocupações variam de respondente para respondente; nenhuma ordenação absoluta é inerente às ocupações em si.

Pelas respostas, você pode construir um *índice* dos compromissos relativos dos respondentes para com as liberdades civis dos comunistas, baseado no número de ocupações que eles lhes permitem exercer. Quem permitir aos comunistas exercer todas as ocupações apóia claramente um grau maior de liberdade civil que quem lhes fechar todas as portas. Além disso, você supõe que quem permitir a um comunista exercer três das ocupações apóia mais as liberdades civis dele do que quem permitir apenas uma ou duas, independente de quais forem as ocupações. Tal índice pode fornecer uma medida ordinal útil e precisa do libertarianismo civil. A parte A da Figura 8-1 ilustra a lógica simples da construção de índices.

Entretanto, suponha que as ocupações do exemplo anterior fossem: (1) escavador de valas, (2) professor de escola secundária, (3) presidente dos Estados Unidos. Neste caso, há toda razão para acreditar que esses três itens têm uma *estrutura de intensidade*. Quem permitir a um comunista ser presidente certamente o deixará ser escavador de valas ou professor secundário. Quem permitir a um comunista cavar valas pode ou não permitir-lhe exercer as outras duas ocupações. Provavelmente, sabendo *a quantidade* de ocupações que um respondente permite a comunistas, saberíamos *quais* profissões permitiria. Nessa situação, uma medida composta formada pelos três itens constituiria uma escala, tal como usamos o termo. A parte B da Figura 8-1 ilustra a construção de uma escala.

Deve estar claro que escalas geralmente são superiores a índices, se por nenhuma outra razão, porque escores de escalas dão mais informação do que escores de índices. Você deve ter cuidado com o uso errado do termo *escala*; chamar uma medida

qualquer de escala ao invés de índice não a torna melhor. Deve também estar alerta em relação a duas outras más interpretações da natureza do escalonamento. Primeiro, se a combinação de vários itens do questionário vai resultar ou não numa escala quase sempre depende da amostra específica de respondentes. Certos itens podem formar uma escala entre os respondentes de uma amostra, mas não de uma outra; não se deve presumir que um certo conjunto de itens formam uma escala só porque a formaram numa determinada amostra. Segundo, usar técnicas de escalonamento (discutidas depois neste capítulo) não assegura a criação de uma escala mais do que usar itens que previamente formaram escalas.

Um exame da literatura substantiva baseada em dados de *survey* mostra que índices são usados muito mais freqüentemente do que escalas. Mas, ironicamente, a literatura metodológica contém muito pouca, se alguma, discussão sobre construção de índices, apesar de abundarem as discussões sobre construção de escalas. Há duas razões aparentes para isto. Primeiro, índices são usados mais freqüentemente porque, em geral, é difícil ou impossível construir escalas a partir dos dados que se têm. Segundo, métodos de construção de índices não são discutidos porque parecem óbvios e diretos.

FIGURA 8-1
Índices versus escalas

Parte A. Lógica da Construção de Índices

Aqui estão várias ocupações que os americanos podem consentir ou não que um comunista declarado exerça. No geral, as diferentes ocupações provavelmente seriam consideradas como tendo mais ou menos o mesmo grau de "importância" pelos respondentes.

Para criar um índice de apoio às liberdades civis, podemos atribuir aos respondentes um ponto para cada ocupação que consentirem a um comunista exercer.

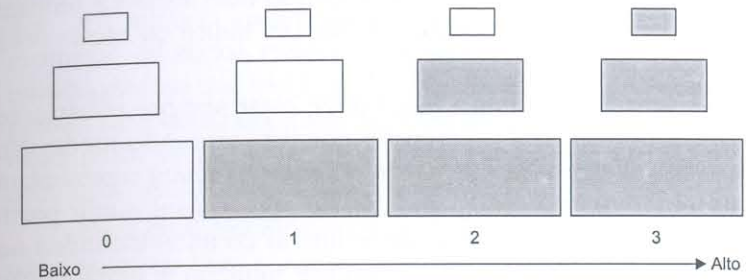
Advogado	Físico	Bibliotecário	Engenheiro	Professor	Repórter
----------	--------	---------------	------------	-----------	----------

Parte B. Lógica da Construção de Escalas

A seguir mostramos três ocupações que os respondentes provavelmente veriam como diferindo em grau de importância com relação ao seu exercício por comunistas. O tamanho das caixas representa os diferentes percentuais de aceitação com relação a um comunista exercendo tal cargo.



A seguir temos os quatro tipos de escalas que esperaríamos obter de uma análise das respostas. Uma caixa sombreada indica a aceitação da ocupação a ser exercida por comunistas.



Mas não é fácil construir índices. Acho que o fracasso generalizado em desenvolver técnicas de construção de índices resultou na criação de muitos índices ruins em pesquisa social. Pensando nisso, concentrei a maior parte deste capítulo nos métodos de construção de índices. Aprendendo a lógica desta construção, você estará melhor preparado para tentar construir escalas. Inclusive, o índice cuidadosamente construído pode acabar se tornando uma escala.

Construção de Índices

Seleção de Itens

Um índice composto é criado para medir uma variável. O primeiro critério para selecionar os itens a serem incluídos no índice é a *validade aparente* (validade lógica). Se, por exemplo, você quiser medir conservadorismo político, cada item deve aparecer para indicar o conservadorismo do respondente (ou seu oposto, o liberalismo). Filiação a partido político seria um desses itens. Se pedirmos aos respondentes para aprovarem ou desaprovarem as idéias de uma conhecida figura pública conservadora, as respostas podem logicamente fornecer outra indicação de conservadorismo. Um pesquisador construindo um índice de religiosidade pode considerar itens investigando a freqüência do respondente à igreja, aceitação de certas crenças religiosas, freqüência de orações, e assim por diante. Cada item listado parece oferecer alguma indicação da religiosidade dele.

Em geral, a literatura metodológica sobre conceitualização e medição enfatiza a necessidade de *unidimensionalidade* na construção de índices e escalas; isto é, uma medida composta

deve representar apenas uma dimensão. Assim, itens que refletem religiosidade não devem ser incluídos numa medida de conservadorismo político, mesmo que as duas variáveis tenham uma relação empírica. Neste sentido, um índice ou escala deve ser unidimensional.

Ao mesmo tempo, você deve estar sempre consciente das nuances sutis da dimensão que está tentando medir. No caso da religiosidade, os itens mencionados acima representam diferentes *tipos* de religiosidade. Se você quiser medir participação ritual em religião, deve limitar os itens incluídos no índice àqueles que especificamente indicam a participação ritual: frequência à igreja, comunhão, confissão, e semelhantes. Mas, se quiser medir religiosidade de modo mais geral, deve incluir um conjunto equilibrado de itens representando cada tipo diferente de religiosidade. Ao final, a natureza dos itens incluídos determinará quão específica ou genérica é a medida da sua variável.

Ao selecionar itens para incluir no índice, você deve considerar também a quantidade de *variância* fornecida pelos itens. Se um item fornece uma indicação de conservadorismo político, por exemplo, você deve anotar quantos conservadores são indicados pelo item. No limite, um item que indicar que ninguém é conservador ou que todos são conservadores não será muito útil para criar um índice. Se só 1% de uma amostra aprovar um político radical de direita, o item que medir isso provavelmente não será muito útil na construção de um índice.

Com relação à variância, temos duas opções. Primeiro, você pode selecionar vários itens que dividam os respondentes mais ou menos ao meio, em termos da variável. Assim, você pode selecionar vários itens, cada um indicando cerca de metade como conservadores e metade como liberais. Apesar de nenhum dos itens, por si só, justificar caracterizar um respondente como “muito conservador”, uma pessoa que parecer conservadora em todos os itens pode ser caracterizada assim. A segunda opção é selecionar itens que difiram em variância. Um item pode indicar que cerca de metade dos respondentes é conservadora, e outro pode indicar que poucos o são. (*Nota:* esta segunda opção é necessária para o escalonamento, mas é também um método razoável para construir índices).

Relações Bivariadas entre Itens

A segunda etapa na construção de índices é examinar as relações bivariadas entre os itens. Se cada item realmente der uma indicação da variável — em termos de validade aparente —,

então os itens devem estar empiricamente relacionados entre si. Por exemplo, se vários itens refletem o conservadorismo ou liberalismo de um respondente, quem aparecer como conservador num item deve também aparecer assim nos outros itens. Mas, tais itens raramente, ou nunca, estão perfeitamente relacionados entre si. (É essa disparidade que leva à necessidade de construir medidas compostas). De todo modo, quem aparecer conservador no item A tem maior probabilidade de aparecer como conservador no item B do que quem figurar como liberal no item A.

Você deve examinar todas as possíveis relações bivariadas entre os vários itens considerados para inclusão no índice, para determinar as forças relativas das relações entre os vários pares de itens. Para tal, pode usar tabelas de porcentagem, coeficientes de correlação, ou ambos. O critério primário para avaliar essas relações é a força delas. Contudo, deve-se usar este critério com sutileza, ficando alerta com itens não relacionados empiricamente entre si. É improvável que itens não relacionados entre si meçam a mesma variável. Um item não relacionado a vários outros provavelmente deve ser abandonado. Ao mesmo tempo, outro sinal de perigo é uma forte correlação entre dois itens. No limite, se dois itens têm uma correlação perfeita, só precisamos de um deles no índice, porque um carrega todas as indicações fornecidas pelo outro. (Esta questão ficará mais clara na próxima seção.)

Um exemplo da literatura substantiva sobre pesquisa de *survey* pode ajudar a ilustrar os passos na construção de índices.¹ Um *survey* com professores de uma escola de medicina estudou o efeito da “perspectiva científica” na qualidade dos cuidados que os médicos dispensam aos pacientes. O objetivo primário foi determinar se médicos com mais inclinação para a ciência eram mais impessoais no tratamento dos pacientes.

O questionário do *survey* ofereceu vários indicadores possíveis das perspectivas científicas dos respondentes. Destes, três itens pareceram — em termos de validade aparente — indicar especialmente se os médicos eram ou não cientificamente orientados. Os três itens foram:

1. Como professor da escola, em qual capacidade acha que pode dar maior contribuição acadêmica: como médico clínico ou como médico pesquisador?
2. Enquanto continua a aprimorar seus conhecimentos médicos, você diria que seus interesses médicos finais estão mais no tratamento completo do paciente ou na compreensão dos mecanismos básicos?

3. No campo da pesquisa terapêutica, em geral você está mais interessado em artigos relatando avaliações da efetividade de vários tratamentos, ou em artigos explorando os argumentos básicos subjacentes ao tratamento?

Em cada item, a segunda resposta indica maior orientação científica do que a primeira. Nas repostas a cada item, quem optar pela segunda alternativa é mais cientificamente orientado do que quem optar pela primeira. Esta conclusão *comparativa* é razoável, mas não devemos pensar que quem opta pela segunda alternativa é “cientista” em algum sentido absoluto. Ele é apenas *mais científico* do que quem opta pela primeira alternativa. Este ponto importante ficará mais claro quando examinarmos a distribuição de respostas que cada item produz.

Quanto ao primeiro item — o melhor papel acadêmico —, só um terço dos respondentes pareceu ser cientificamente orientado, dizendo que sua melhor contribuição acadêmica era como pesquisador. Mas este resultado não significa que só um terço da amostra é de “cientistas”, porque os outros dois itens sugeriram conclusões bem diferentes. Em resposta ao segundo item, o maior interesse médico, cerca de dois terços escolheram a resposta científica, indicando estarem mais interessados em aprender sobre mecanismos básicos do que sobre cuidado completo ao paciente. Em resposta ao terceiro item, preferências de leitura, aproximadamente 80% preferiram a resposta científica.

Estes três itens de questionário não podem nos dizer quantos “cientistas” há na amostra, porque nenhum deles está relacionado a um conjunto de critérios descrevendo o que constitui ser um cientista em sentido absoluto. Usar os itens para isso nos deixaria com o problema de três estimativas bem diferentes de quantos cientistas há na amostra.

Estes três itens de questionário nos dão três indicadores independentes das inclinações relativas dos respondentes quanto à ciência na medicina. Cada item separa os respondentes entre mais e menos científicos. Tendo em vista as diferentes distribuições de respostas produzidas pelos três itens, é claro que cada grupo resultante de respondentes mais ou menos científicos terá integrantes um pouco diferentes. Quem parece científico num item não o parece em outro. Mesmo assim, como cada item mede a mesma dimensão geral, devemos encontrar alguma correspondência entre os diferentes grupos. Quem parecer científico num item deve ter maior probabilidade de parecer científico na resposta a outro item do que quem parecer não-científico na resposta ao primeiro. Devemos encontrar associação ou correlação entre as respostas dadas a quaisquer dois itens.

As tabelas da Figura 8-2 permitem examinar as associações entre as respostas aos três itens.² Três tabelas bivariadas (duas variáveis) são apresentadas, mostrando a distribuição conjunta das respostas a cada par de itens. Apesar de cada item único produzir um agrupamento diferente de respondentes “científicos” e “não-científicos”, a Figura 8-2 mostra que as respostas dadas a cada item correspondem, em certo grau, às respostas dadas a cada outro item.

Um exame das três relações bivariadas da Figura 8-2 apóia a crença de que os três itens medem a mesma variável: orientação científica. Começamos observando a primeira relação bivariada da tabela. As escolhas dos professores de seus melhores papéis acadêmicos e de seus principais interesses médicos nos dão indicações de orientação científica. Quem escolheu “pesquisador”, no primeiro caso, indica maior inclinação científica do que quem respondeu “clínico”. Quem respondeu “mecanismos básicos” parece mais inclinado cientificamente do que quem respondeu “tratamento completo ao paciente”, na pergunta sobre o principal interesse. Se estes itens realmente medirem a mesma coisa, quem parece científico no primeiro item (“pesquisador”) deve parecer mais científico no segundo (“mecanismos básicos”) do que quem parece não científico no primeiro (“clínico”). Pelos dados, vemos que 87% dos “pesquisadores” parecem científicos no segundo item, em oposição a 51% dos “clínicos”. (*Nota:* O fato dos “clínicos” estarem divididos ao meio, em relação aos principais interesses médicos, é irrelevante. O relevante é que eles são menos científicos do que os “pesquisadores”, em seus interesses médicos). A força dessa relação pode ser resumida numa diferença de 36 pontos percentuais.

FIGURA 8-2
Relações bivariadas entre itens sobre orientação científica

A.

		Melhor Papel Acadêmico	
		Clínico	Pesquisador
Principal Interesse Médico	Tratamento completo do paciente	49%	13%
	Mecanismos básicos	51%	87%
		100%	100%
		(268)	(159)

B.

		Preferências de Leitura	
		Efetividade	Argumentos
Principal Interesse Médico	Tratamento completo do paciente	68%	30%
	Mecanismos básicos	32%	70%
		100%	100%
		(78)	(349)

C.

		Preferências de Leitura	
		Efetividade	Argumentos
Melhor Papel Acadêmico	Clínico	85%	64%
	Pesquisador	15%	36%
		100%	100%
		(78)	(349)

Pode-se chegar à mesma conclusão geral em relação às outras relações bivariadas. A força da relação entre preferências de leitura e principal interesse médico pode ser resumida como diferença de 38 pontos percentuais; a força da relação entre preferências de leitura e os dois papéis acadêmicos pode ser resumida numa diferença de 21 pontos percentuais.

Inicialmente, os três itens foram selecionados com base na validade aparente; cada um pareceu dar alguma indicação das orientações científicas dos professores. Examinando as relações bivariadas entre os pares de itens, encontramos apoio para a crença inicial de que todos os itens medem basicamente a mesma coisa.

Relações Multivariadas entre Itens

Constatar uma relação bivariada, esperada, entre pares de itens sugere que eles podem ser incluídos num índice composto, mas esta justificativa não basta. O próximo passo na construção do índice é examinar as relações multivariadas entre os itens. É preciso examinar as relações simultâneas entre muitas variáveis antes de combiná-las num único índice.

A principal razão para construir um índice é desenvolver um método para classificar respondentes nos termos de alguma

variável, como conservadorismo político, religiosidade ou orientação científica. Por exemplo, um índice de conservadorismo político deve identificar respondentes que são muito conservadores, moderadamente conservadores, pouco conservadores e nada conservadores (ou, respectivamente, moderadamente liberais e muito liberais, ao invés das duas últimas categorias). As diferentes gradações em termos da variável resultam da combinação das respostas dadas aos itens do índice. Portanto, quem parece ser conservador em todos os itens é considerado muito conservador no total.

Um índice pode só fornecer gradações significativas se cada item acrescentar algo à avaliação de cada respondente. Falamos na seção anterior que não se deve incluir no mesmo índice dois itens perfeitamente relacionados entre si. Se um deles o for, incluir o outro não acrescentará nada à nossa avaliação dos respondentes. Examinar as relações multivariadas entre itens é um outro modo de eliminar material inútil, além de determinar a força combinada de um certo conjunto de itens quanto à medida da variável considerada.

Voltemos ao exemplo dos professores de medicina, para compreender melhor para que serve o exame multivariado. A Figura 8-3 mostra a relação trivariada entre os três itens. A tabela, agora, é um pouco diferente da Figura 8-2. Agora, a amostra foi dividida em quatro grupos, de acordo com (1) seu melhor papel acadêmico e (2) sua preferência de leitura. Os números entre parênteses são de respondentes em cada grupo. (Portanto, 66 professores disseram que poderiam ensinar melhor como clínicos e também disseram que preferiam artigos sobre a eficácia dos tratamentos). Para cada um dos quatro grupos, apresentamos a porcentagem de respondentes que disseram estar mais interessados nos mecanismos básicos do que no tratamento completo do paciente. (Dos 66 professores mencionados acima, 27% estão primariamente interessados nos mecanismos básicos).

FIGURA 8-3

Relações trivariadas entre os itens sobre orientação científica

	Melhor Papel Acadêmico	
	Clínico	Pesquisador
Porcentagem interessada nos mecanismos básicos	27%	58%
	(66)	(12)
Preferências de Leitura	58%	89%
	(219)	(130)

O arranjo dos quatro grupos se baseia na conclusão já tirada com relação à orientação científica. Respondentes do canto superior esquerdo da tabela são presumivelmente os menos cientificamente orientados dos quatro grupos, quanto a melhor papel acadêmico e preferências de leitura. Os no canto inferior direito da tabela são presumivelmente os mais cientificamente orientados, quanto aos dois itens.

Lembre-se de que expressar interesse primordial em “mecanismos básicos” também foi tomado como indicador de orientação científica. Como esperamos, os no canto inferior direito têm a maior probabilidade de dar esta resposta (89%), e os no canto superior esquerdo a menor probabilidade (27%). Os que deram respostas mistas em relação a papel acadêmico e preferências de leitura estão em posição intermediária no seu interesse por mecanismos básicos (58% em ambos os casos).

Esta tabela nos diz muitas coisas. Primeiro, vemos que as relações originais entre pares de itens não são afetadas significativamente pela presença de um terceiro item. Por exemplo, a relação entre papel acadêmico e principal interesse médico foi resumida como diferença de 36 pontos percentuais. Observando a Figura 8-3, vemos que, somente entre os respondentes mais interessados em artigos sobre a efetividade do tratamento, a relação entre papel acadêmico e principal interesse médico foi de 31 pontos percentuais (58% menos 27% na primeira linha), e o mesmo é verdadeiro para os mais interessados em artigos sobre argumentos básicos (89% menos 58% da segunda linha). A relação original entre papel acadêmico e principal interesse médico é essencialmente a mesma da Figura 8-2, mesmo entre os considerados científicos ou não-científicos quanto às preferências de leituras.

Podemos tirar a mesma conclusão do exame das colunas na Figura 8-3. Lembre que a relação original entre preferências de leitura e principal interesse médico foi resumida como diferença de 38 pontos percentuais. Observando apenas os “clínicos” na Figura 8-3, vemos que a relação entre os outros dois itens é agora 31 pontos percentuais. A mesma relação é encontrada entre os “pesquisadores” na segunda coluna.

A importância destas observações fica mais evidente quando consideramos o que pode ter acontecido. A tabela da Figura 8-4 apresenta dados hipotéticos para ilustrar este ponto.

FIGURA 8-4
Relação trivariada hipotética entre itens de orientação científica

	Melhor Papel Acadêmico	
	Clinico	Pesquisador
Porcentagem interessada nos mecanismos básicos	51% (66)	87% (12)
Preferências de Leitura	51% (219)	86% (130)

Estes dados hipotéticos contam uma história bem diferente da que contaram os dados verdadeiros da Figura 8-3. Neste caso, evidentemente persiste a relação original entre papel acadêmico e principal interesse médico, mesmo quando preferências de leitura são introduzidas no quadro. Em cada linha da tabela, os “pesquisadores” têm probabilidade maior do que os “clínicos” de expressar interesse pelos mecanismos básicos. Mas, observando as colunas, vemos que não há relação entre preferência de leitura e principal interesse médico. Se soubermos se os respondentes acham que podem ensinar melhor como clínicos ou como pesquisadores, conhecermos suas preferências de leitura nada acrescenta à nossa avaliação de suas orientações científicas. Se algo como a Figura 8-4 resultar de nossos dados, concluiremos que preferências de leitura não devem ser incluídas no mesmo índice de papel acadêmico, porque não contribuirá em nada para o índice composto.

Neste exemplo, estavam envolvidos apenas três itens de questionário. Se considerarmos mais itens, é necessário usar tabelas multivariadas mais complexas. Até aqui, limitamos a atenção à análise trivariada dos três itens. Lembre que o propósito desta etapa, na construção de índices, é examinar a interação simultânea dos itens, para determinar se eles podem ser incluídos no mesmo índice.

Escore de Índices

Após determinar os melhores itens para inclusão no índice, o próximo passo é atribuir escores a respostas específicas, assim criando um único índice composto a partir de vários itens. É preciso tomar duas decisões básicas sobre isso.

A primeira é decidir a amplitude desejada dos escores de índice. Uma das principais vantagens de um índice sobre

um item único é a amplitude de gradações oferecida na medição de uma variável. Como observamos antes, conservadorismo político pode ser medido desde “muito conservador” a “nada conservador” (ou “muito liberal”). Então, quão longe, nos extremos, deve se estender o índice? Como antes, esta decisão envolve a questão da variância. Quase sempre, quando se estende os possíveis extremos de um índice, menos casos são encontrados em cada extremidade. Se você tentar medir conservadorismo político no seu maior extremo, pode descobrir que quase ninguém cai nessa categoria.

Então, a primeira decisão está relacionada aos desejos conflitantes de (1) a maior amplitude de medição possível no índice e (2) um número adequado de casos em cada ponto do índice. Você tem que chegar a algum tipo de acordo entre estes desejos.

A segunda decisão concerne a atribuição de escores para cada resposta específica. Basicamente, você deve decidir se atribui peso igual ou diferentes pesos aos itens no índice. Como veremos mais tarde, a construção de escalas é bastante diferente neste aspecto, mas a atribuição de escores na construção de índices é uma questão aberta. Enquanto não há regras firmes a seguir, sugiro — e a prática tende a confirmar — que os itens devam ter pesos iguais, a menos que haja fortes razões para receberem pesos distintos. Ou seja, ponderar diferentemente deve ser justificado, enquanto pesos iguais deve ser a norma.

É claro que a decisão sobre a ponderação deve estar relacionada ao equilíbrio dos itens escolhidos. Se o índice representar a mistura de aspectos pouco diferentes de uma variável, então você deve ponderar igualmente cada um deles. Mas, em alguns casos, você pode achar que dois itens refletem essencialmente o mesmo aspecto, enquanto um terceiro reflete um aspecto diferente. Se quiser respeitar igualmente os dois aspectos no índice, você pode decidir dar ao item diferente um peso igual à combinação dos pesos dos dois itens similares. Nesse caso, você pode atribuir um escore máximo de 2 para o item diferente e escores máximos de 1 para cada um dos dois itens similares.

Embora as razões para ponderar respostas devam considerar questões como essas, normalmente você experimenta diferentes métodos de ponderação, examinando os pesos relativos atribuídos a diferentes aspectos e, ao mesmo tempo, considerando a amplitude e a distribuição dos casos. No fim, o método escolhido representa uma combinação entre essas

diferentes demandas. (*Nota:* Nesta atividade, como na maioria das atividades de *survey*, a decisão pode ser revista, com base num exame posterior. A validação de um índice, a ser discutida daqui a pouco, pode levá-lo a reciclar os esforços e construir um índice completamente diferente).

No exemplo do *survey* na escola de medicina, decidiu-se ponderar cada item igualmente, porque eles foram escolhidos, em parte, por representarem aspectos ligeiramente diferentes da variável orientação científica. Em cada item, os respondentes recebiam um escore de 1 quando davam resposta “científica” e 0 ao dar resposta “não-científica”. Assim, cada respondente podia receber um escore de 0, 1, 2, ou 3, dependendo do número de respostas “científicas” dadas. Este método de ponderação permitiu o que foi considerado uma amplitude útil de variação — quatro categorias de índices — e também um número suficiente de casos em cada categoria para as análises.

Trabalhando com Dados Faltantes

Em quase todo *survey*, alguns respondentes não respondem algumas perguntas (ou escolhem a opção “não sei”). Esses dados faltantes são um problema em todos os estágios de análise, mas especialmente na construção de índices. (As escalas são diferentes também neste ponto). Se alguns respondentes não responderem itens de um índice composto, você terá problemas para atribuir escores a estes entrevistados na construção do índice. Há vários métodos para lidar com este problema.

Primeiro, se relativamente poucos respondentes tiverem dados faltantes, você pode decidir excluí-los da construção e análise do índice. Os principais cuidados neste caso são se a quantidade restante para análise ainda é suficiente e se excluir estes respondentes resulta numa amostra tendenciosa ao usar o índice na análise. A possibilidade de viés pode ser examinada pela comparação, em outras variáveis relevantes, dos respondentes que seriam incluídos no índice e dos que seriam excluídos. (No exemplo da escola de medicina, esta foi a decisão tomada em relação aos dados faltantes).

Segundo, pode haver razões para tratar os dados faltantes do mesmo modo que as outras respostas dadas. Por exemplo, se o questionário pedir aos respondentes para indicar sua participação em várias atividades marcando “sim” ou “não” para cada uma, muitos entrevistados podem marcar “sim” para algumas atividades e deixar as outras em branco. Neste caso,

você pode decidir que uma resposta em branco significa um “não” e registrar os dados faltantes como se tivesse sido marcada a opção “não”.

Terceiro, uma análise cuidadosa dos dados faltantes pode permitir uma interpretação do seu significado. Ao construir uma medida de conservadorismo político, por exemplo, você pode descobrir que quem não respondeu uma certa pergunta era geralmente tão conservador, nos outros itens, quanto aqueles que deram uma resposta “conservadora” a esta pergunta. Outro exemplo: um *survey* recente, que mediu crença religiosa, descobriu que os respondentes que escolheram “não sei” para uma certa crença eram quase idênticos aos “descrentes” nas respostas relativas a outras crenças. (*Nota:* Você não deve tomar estes exemplos como guias empíricos para seus estudos, tome-os apenas como sugestões de formas para analisar os próprios dados). Sempre que a análise de dados faltantes permitir tais interpretações, você pode ponderar de acordo com estas soluções.

Quarto, você pode decidir atribuir um escore intermediário aos dados faltantes. Por exemplo, se um item for ponderado em 0, 1 e 2 para suas três possíveis respostas, você pode atribuir um escore intermediário (1) ao item para os respondentes que não responderam. (É a mesma lógica pela qual a resposta “indeciso” é muitas vezes registrada como estando entre “concordo” e “discordo”).

Quinto, você pode atribuir escores de índice proporcionalmente, baseado nas respostas dadas por um respondente. Por exemplo, suponha que seis itens são combinados num índice, com escores de 0 ou 1 para cada item. O escore máximo que um entrevistado pode receber é 6. Se um respondente responder apenas cinco itens, mas receber 5 nestes itens, poderá receber um escore proporcional de $5/5 \times 6 = 6$ no índice. Quem receber 2 nos quatro itens respondidos poderá receber um escore final de $2/4 \times 6 = 3$. Quando estes cálculos resultarem em frações, deve-se usar algum método de arredondamento, para simplificar os escores finais.

Finalmente, você pode não querer usar qualquer um destes métodos para lidar com dados faltantes, mas a análise posterior pode exigir que todos os respondentes na amostra sejam ponderados. Nesse caso, você pode atribuir escores aos dados faltantes de forma aleatória. Para um item que atribuir os escores de 0, 1 e 2, o primeiro que não respondê-lo poderá receber um escore de 1, o segundo 0 e o terceiro 2 etc. Este é o método mais conservador, do ponto de vista de análise

de pesquisas, porque você está “embaralhando as cartas” contra si mesmo. Se o índice resultante provar ser uma ferramenta poderosa na análise, você poderá concluir que seria ainda mais poderoso se todos houvessem respondido todas as questões. (Claro que, se sua meta for mostrar que o índice não está relacionado a outras variáveis, você terá embaralhado as cartas a seu favor).

A escolha de um determinado método para tratar dados faltantes depende tanto da situação da pesquisa que é impossível sugerir um único melhor método ou hierarquizá-los. Geralmente, sugiro uma análise dos que não responderam, tanto quanto a possíveis vieses na sua exclusão quanto às respostas deles a outros itens do índice. O objetivo final da análise é compreender seus dados.

Validação de Índices

Até agora, discutimos os passos da seleção e ponderação dos itens de um índice composto. Se cada passo for dado cuidadosamente, aumentamos a probabilidade do índice realmente medir a variável a que se refere. O sucesso na criação de um índice útil, contudo, ainda não foi provado; a *validação* do índice ajuda a alcançar este objetivo. A base lógica da validação é: supomos que um índice composto meça uma variável, ou seja, que os escores sucessivos de um índice ordenem os respondentes quanto àquela variável. Um índice de conservadorismo político ordena os grupos quanto ao seu conservadorismo relativo. Se um índice for bem sucedido nisso, então pessoas classificadas como relativamente conservadoras pelo índice devem aparecer como relativamente conservadoras em todos os itens do questionário (ou outros indicadores) que também reflitam orientação política. Há muitos métodos para validar um índice composto.

Análise de Item. O primeiro passo na validação de um índice é uma validação interna chamada *análise de item*. Examina-se o quanto o índice composto está relacionado (ou prediz as respostas) aos itens do questionário incluídos no próprio índice. Se o índice tiver sido construído cuidadosamente, examinando as relações bivariadas e multivariadas entre vários itens, este passo confirmará a validade do índice. Para um índice complexo com muitos itens, este passo provê um teste mais parcimonioso da contribuição independente de cada item para o índice. Se descobirmos que um item não está muito relacionado ao índice, podemos supor que outros itens do índice estão eliminando a contribuição daquele

item. Assim, o item em questão não contribuirá para a potência do índice e deverá ser excluído dele.

A análise de item é um importante primeiro teste para a validade do índice, mas não é suficiente. Se o índice medir adequadamente uma variável, deverá prever com sucesso outros indicadores dela. Para testar isso, precisamos lançar mão de itens não incluídos no índice.

Validação Externa. Pessoas classificadas como politicamente conservadoras no índice devem aparecer como conservadoras nas respostas a outros itens do questionário. Devemos compreender, é claro, que estamos falando de conservadorismo relativo; é impossível dar uma definição final absoluta do que vem a ser “conservadorismo” de maneira definitiva. Mas os respondentes classificados como os mais conservadores pelo índice devem ser os mais conservadores ao responder outras perguntas. Os classificados como menos conservadores no índice devem ser os menos conservadores nas respostas a outros itens. O ordenamento de grupos de respondentes no índice deve predizer o ordenamento deles nas respostas a outras perguntas que lidam com orientação política.

Em nosso exemplo do índice de orientação científica, várias perguntas do questionário ofereceram a possibilidade de validação adicional. A Tabela 8-1 apresenta alguns desses itens. Os itens listados permitem várias lições sobre validação de índices. Primeiro, notamos que o índice prediz fortemente as respostas aos itens de validação, no sentido em que o ordenamento das respostas científicas entre os quatro grupos é o mesmo do ordenamento fornecida pelo próprio índice.

TABELA 8-1
Validação do índice de orientação científica

	Índice da Orientação Científica			
	Baixo			Alto
	0	1	2	3
Porcentagem interessada em freqüentar palestras científicas na escola de medicina	34	42	46	65
Porcentagem que diz que os professores da faculdade devem ter experiência como médicos pesquisadores	43	60	65	89
Porcentagem que preferiria que as atividades dos professores fossem restritas apenas à pesquisa	0	8	32	66
Porcentagem envolvida em pesquisa no ano acadêmico anterior	61	76	94	99

Ao mesmo tempo, cada item dá uma descrição *diferente* da orientação científica como um todo. Por exemplo, o último item de validação indica que a grande maioria dos professores da faculdade pesquisou durante o ano anterior. Se esse fosse o único indicador de orientação científica, concluiríamos que quase todos os professores são cientificamente orientados. Mesmo assim, os professores considerados mais científicos pelo índice têm maior probabilidade de terem pesquisado do que os classificados como relativamente menos científicos. O terceiro item de validação desenha um quadro descritivo diferente. Somente uma minoria dos professores disse que preferia encargos exclusivamente de pesquisa. Ainda assim, as porcentagens desta resposta também correspondem aos escores atribuídos no índice.

Índices Ruins Versus Validadores Ruins. Um dilema que enfrentam quase todos construtores de índices é o fracasso dos itens externos em validarem o índice. Se a análise interna de item mostrar relações inconsistentes entre os itens incluídos no índice e o próprio índice, há alguma coisa errada com o índice. Entretanto, se o índice fracassar em predizer consistentemente os itens de validação externa, a conclusão a tirar é mais ambígua. Você tem que optar entre duas possibilidades: (1) o índice não mede adequadamente a variável em questão, ou (2) os itens de validação não medem adequadamente a variável e, portanto, não são um teste suficiente do índice.

Se você trabalhou longa e conscienciosamente na construção do índice, achará a segunda conclusão bastante atraente. Geralmente, você sente que incluiu os melhores indicadores da variável no índice e que os itens de validação, portanto, são indicadores de segunda classe. Mas, você deve reconhecer que o índice é, propositadamente, uma medida muito poderosa da variável; como tal, ele deve estar de alguma maneira relacionado a qualquer item que, mesmo de longe, se relacione com a variável.

Quando a validação externa fracassa, você deve reexaminar o índice antes de decidir que os itens de validação são insuficientes. Uma maneira de fazer isso é examinar as relações entre os itens de validação e os itens individuais do índice. Descobrir que alguns itens do índice se relacionam aos validadores mas outros não melhora nossa compreensão do índice construído inicialmente.

Não há “receita de bolo” para este dilema; ele é uma agonia com a qual o pesquisador sério deve aprender a conviver. Ao cabo, a sabedoria da decisão sobre o índice será determinada

pela utilidade dele nas suas análises posteriores. Talvez você decida que o índice é bom e os validadores defeituosos e, mais tarde, descubra que a variável em questão (como medida pelo índice) não está relacionada a outras variáveis das maneiras esperadas. Neste ponto, você poderá retornar à composição do índice.

Neste texto, demos considerável importância à construção de índices simples, por duas razões. Primeiro, uma revisão dos temas da literatura sobre pesquisa empírica mostra a popularidade destas medidas entre os pesquisadores de *survey*. Segundo, porque temos muito pouca, se é que alguma, discussão na literatura ou em manuais metodológicos sobre as técnicas para a construção de índices. A construção de índices "simples" talvez tenha sido vista como simples demais para merecer tal discussão, e as técnicas permaneceram parte da tradição oral da pesquisa de *survey*.

"Escalonamento" Likert

Definimos uma escala como uma medida composta construída com base numa estrutura de intensidade entre os itens da medida. Na construção de escalas, os padrões de resposta entre vários itens são ponderados, enquanto na construção de índices as respostas individuais são ponderadas e os escores independentes somados. Por esta definição, o método de medição desenvolvido por Rensis Likert, chamado *escalamento Likert*, representa uma maneira mais sistemática e refinada de construir índices; portanto, este método será discutido aqui e não nas seções sobre escalonamento.

O termo *escala Likert* é associado a um formato de pergunta freqüentemente usado nos questionários de *survey*. Basicamente, mostra-se aos respondentes uma declaração e se pergunta se eles "concordam fortemente", "concordam", "discordam" ou "discordam fortemente". Modificações na redação das categorias de resposta (por exemplo, "aprovam") podem ser feitas.

O valor particular deste formato é a ordinalidade não-ambígua das categorias de resposta. Se fosse permitido aos respondentes inventar ou escolher respostas como "concordo até certo ponto", "concordo mesmo", "concordo em quase tudo", e assim por diante, seria impossível julgar a força relativa da concordância dos vários respondentes. O formato Likert resolve facilmente este dilema.

O escalonamento Likert também se presta a um método bastante direto de construção de índices. Já que se usa categorias idênticas de resposta para os vários itens que medem

uma variável, cada um desses itens pode ser ponderado de maneira uniforme. Com cinco categorias de respostas, escores de 0 a 4 ou 1 a 5 podem ser atribuídos, levando em conta a "direção" do item (por exemplo, atribuir o escore 5 a "concordo fortemente" nos itens positivos e a "discordo fortemente" nos itens negativos). Cada respondente recebe então um escore geral, representando a soma dos escores recebidos pelas respostas aos itens individuais.

O método Likert baseia-se na suposição de que o escore geral, que resulta das respostas aos vários itens que aparentemente refletem a variável estudada, fornece uma medida razoavelmente boa da variável. Estes escores gerais não são o produto final da construção de índices; na verdade, são usados para se fazer uma *análise de itens* levando à escolha dos melhores itens. Essencialmente, cada item individual é correlacionado à grande medida composta. Presume-se que os itens que melhor se correlacionam com a medida composta fornecem os melhores indicadores da variável, e somente estes itens seriam incluídos no índice que será finalmente usado para as análises da variável.

Note que a atribuição uniforme de escores nas categorias de resposta de item-Likert presume que cada item tenha aproximadamente a mesma intensidade dos outros. Esta é o aspecto chave no qual o método Likert difere do escalonamento, como usamos o termo neste livro. Os itens tipo-Likert podem ser usados de várias maneiras e não estamos presos ao método descrito. Eles podem ser combinados com outros tipos de itens na construção de índices simples; igualmente, podem ser usados na construção de escalas. Mas, se todos os itens cuja inclusão for considerada para uma medida composta forem do formato Likert, deve-se considerar o método descrito antes.

Voltemos agora a atenção para as técnicas de escalonamento. Há muitos métodos à disposição do pesquisador de *survey*; consideraremos apenas as escalas Bogardus, Thurstone e Guttman.

Construção de Escalas

Bons índices fazem uma classificação ordinal dos respondentes numa variável. Todos os índices se baseiam na suposição de que uma pessoa com duas indicações de ter inclinação científica, por exemplo, deve ser mais científica do que uma pessoa com apenas uma indicação. No entanto, um índice

pode falhar em considerar que nem todas as indicações de uma variável são igualmente importantes. (É claro que se pode tentar resolver este problema pesando os indicadores diferentemente).

As escalas dão mais segurança com relação à ordinalidade, detectando *estruturas* entre os indicadores. Os itens de uma medida composta podem ter diferentes *intensidades* em termos da variável. As três descrições seguintes de procedimentos de escalonamento ilustram a variedade de técnicas disponíveis.

Escala de Distância Social de Bogardus

Um bom exemplo de escala é a *Escala de Distância Social de Bogardus*. Suponha que você queira saber o quanto os respondentes estão dispostos a se relacionar com albaneses. Pode fazer as seguintes perguntas a eles.

1. Você deixaria albaneses morarem no seu país?
2. Você deixaria albaneses morarem na sua comunidade?
3. Você deixaria albaneses morarem na sua vizinhança?
4. Você deixaria um albanês morar ao lado de sua casa?
5. Você deixaria seu filho (sua filha) se casar com um albanês?

As perguntas estão em ordem crescente de proximidade do contato que os respondentes podem querer ou não com albaneses. Começando com a intenção original de medir a disposição de se relacionar com albaneses, desenvolvemos várias questões indicando diferentes graus de intensidade nesta variável.

As claras diferenças de intensidade sugerem uma estrutura entre os itens. Presumivelmente, respondentes dispostos a aceitar uma espécie de associação estariam dispostos a aceitar todas as associações precedentes na lista, isto é, as com menores intensidades. Por exemplo, quem deixar albaneses morar na vizinhança certamente os aceitará na sua comunidade e na sua nação, mas pode ou não aceitá-los como vizinhos próximos ou parentes. Esta é a estrutura lógica da intensidade inerente aos itens.

Empiricamente, esperamos encontrar que a maioria dos respondentes aceita a co-cidadania e a minoria aceita o casamento. Falamos, nesse sentido, de “itens fáceis” (co-cidadania) e “itens difíceis” (casamento). Mais respondentes concordam

com os itens fáceis do que com os difíceis. Com algumas exceções inevitáveis, a lógica exige que, quando os respondentes recusam uma associação apresentada na escala, também recusam todas as mais difíceis que se seguem.

A Escala de Distância Social de Bogardus ilustra a economia de escalonamento como um dispositivo de redução de dados. Se soubermos *quantas* associações com albaneses um respondente aceitaria, saberemos *quais* associações aceitaria. Um único número pode resumir precisamente cinco ou seis respostas de *survey* sem perda de informação.

Escalas Thurstone

Muitas vezes, a estrutura inerente da Escala de Distância Social de Bogardus não é apropriada para a variável sendo medida. Na verdade, esta estrutura lógica entre vários indicadores raramente é aparente. O *escalonamento de Thurstone* procura desenvolver um formato para gerar grupos de indicadores de uma variável com pelo menos uma estrutura *empírica* entre eles. Um formato básico é o de “intervalos aparentemente iguais”.

Dá-se a um grupo de “juízes” talvez cem itens, que se acredita serem indicadores de alguma variável. Pede-se a cada um para estimar com que força cada item é indicador da variável, atribuindo escores de, por exemplo, 1 a 13. Se a variável for preconceito, pede-se aos juízes para marcar 1 para os indicadores mais fracos de preconceito, 13 para os mais fortes e números intermediários para os que estiverem entre eles.

Depois de os juízes terminarem a tarefa, examinam-se os escores que atribuíram a cada item, para determinar quais itens produziram mais acordo entre eles. Os itens sobre os quais eles mais discordaram são rejeitados como ambíguos. Seleciona-se, entre os itens que produziram acordo geral, um ou mais para representar cada escore da escala de 1 a 13.

Os itens assim selecionados são incluídos no questionário. Espera-se que quem parecer preconceituoso em itens com força 5 parece preconceituoso nos itens com menor força; também não se espera que quem não parecer preconceituoso nos itens com força 6 não o parece em itens com maiores intensidades.

Se os itens da Escala Thurstone forem adequadamente desenvolvidos e ponderados, aparecerá a economia e a eficácia da redução dos dados inerentes à Escala de Distância Social de Bogardus. Um único escore pode ser atribuído a cada respondente (aceita a força do item mais difícil), e este escore representaria adequadamente as respostas a vários itens do questionário.

Como acontece com a escala de Bogardus, um respondente com escore 6 pode ser considerado mais preconceituoso do que um que recebeu escore 5 ou menos.

O escalonamento de Thurstone é raramente usado em pesquisas de *survey* hoje, principalmente devido ao enorme gasto de energia requerido no “juízo” dos itens. Vários (talvez dez ou quinze) juízes têm que gastar um tempo considerável atribuindo escores aos muitos itens iniciais. Já que a qualidade do julgamento depende da experiência com e do conhecimento da variável considerada, a tarefa pode requerer pesquisadores profissionais. Além do mais, os significados dos vários itens que indicam uma certa variável tendem a mudar com o tempo. Portanto, um item que recebe um certo peso num momento pode ter um peso bem diferente mais tarde. Uma Escala Thurstone teria que ser periodicamente atualizada para ser efetiva.

Escalonamento de Guttman

Uma técnica de escalonamento muito popular hoje foi desenvolvida por Louis Guttman. Como as escalas de Bogardus e Thurstone, o *escalonamento de Guttman* baseia-se no fato de que alguns itens podem ser indicadores mais “difíceis” da variável do que outros. Quem aceita um item difícil também aceita os mais fáceis. Se tal estrutura aparecer nos dados examinados, podemos dizer que os itens formam um Escala de Guttman. Um exemplo deve ser suficiente.

No exemplo anterior descrevendo a medição da orientação científica entre professores de uma escola de medicina, foi construído um índice simples. Como veremos, os três itens do índice formam essencialmente uma Escala de Guttman. Esta possibilidade aparece quando procuramos indicadores relativamente “difíceis” e “fáceis” da orientação científica.

O item que perguntava aos respondentes se eles serviriam melhor como clínicos ou pesquisadores é o mais difícil dos três; se este fosse o único indicador da variável, só cerca de um terço deles seria considerado científico. Se o item sobre principal interesse médico (tratamento completo ao paciente versus mecanismos básicos) fosse usado como único indicador, quase dois terços seriam considerados científicos. A preferência de leitura (efetividade do tratamento versus argumentos) é o mais fácil; cerca de 80% dos entrevistados seriam considerados científicos neste item.

TABELA 8-2
Escalonamento da orientação científica

	Preferência de Leitura	Principal Interesse	Papel Acadêmico	Número de Casos
Tipos de escalas:	+	+	+	116
	+	+	-	127
Total = 383	+	-	-	92
	-	-	-	48
Tipos mistos:	-	+	-	18
	+	-	+	14
Total = 44	-	-	+	5
	-	+	+	7

Para determinar se há uma estrutura escalar entre as respostas aos três itens, devemos examinar os vários padrões de respostas possíveis aos três itens simultaneamente. Na Tabela 8-2, todos os padrões possíveis são apresentados de forma esquemática. Para cada um dos três itens, sinais positivos e negativos foram usados para indicar respostas científicas e não-científicas, respectivamente. (“Mais” indica resposta científica, “menos” resposta não-científica.)

Os primeiros quatro padrões de resposta da tabela formam o que chamamos *tipos de escala* — os padrões que formam uma estrutura escalar. Seguindo os respondentes que selecionaram todas as três respostas científicas (linha 1), vemos que os com apenas duas respostas científicas (linha 2) optaram pelas duas mais fáceis; aqueles com apenas uma resposta científica (linha 3) optaram pela mais fácil das três. Finalmente, os entrevistados que não optaram por respostas científicas estão na linha 4.

A segunda parte da tabela apresenta os padrões de respostas que violam a estrutura escalar dos itens. Os dois últimos padrões de resposta são os que saem mais da estrutura escalar: os que só aceitaram o item mais difícil e os que rejeitaram apenas o mais fácil.

A última coluna da tabela indica o número de respondentes no *survey*, em cada padrão de resposta. É imediatamente evidente que a grande maioria (90%) deles se encaixa num dos tipos de escala. Mas a existência de tipos mistos indica que os itens não formam uma Escala de Guttman perfeita.

Neste ponto, precisamos lembrar que uma das principais funções do escalonamento é a redução eficiente dos dados. As escalas provêm uma técnica de apresentação sinóptica dos

dados e, ao mesmo tempo, mantêm o máximo possível da informação original.

Quando os itens de orientação científica formaram um índice, em nossa discussão anterior, os respondentes receberam um ponto no índice para cada resposta científica que forneciam. Se os mesmos três itens forem ponderados como Escala de Guttman, alguns respondentes receberiam escores diferentes dos recebidos no índice. Eles receberiam escores de escala que permitiriam uma reprodução mais exata das respostas originais a todos os três itens.

Quem se encaixar nos tipos de escala recebe os mesmos escores que recebe na construção do índice. Por exemplo, as pessoas que dão todas as três respostas científicas recebem escore 3. Note que, se soubermos que alguém neste grupo recebeu um escore 3, podemos precisamente prever que ele escolheu todas as três respostas científicas. Para as pessoas na segunda fileira da tabela, a atribuição do escore de escala 2 nos levaria a prever exatamente respostas científicas aos dois itens mais fáceis e resposta não-científica ao mais difícil. Em cada um dos quatro tipos de escala, podemos prever exatamente as respostas reais dadas por todos os respondentes.

Entretanto, os tipos mistos na tabela são um problema. O primeiro tipo misto (- + -) recebeu escore 1 no índice para indicar apenas uma resposta científica. Se lhe for atribuído o escore de escala 1, porém, prediríamos que todo mundo neste grupo escolheu apenas o item mais fácil (+ - -), fazendo em consequência dois erros por respondente. Os escores de escala são atribuídos, portanto, para reduzir os erros na reconstrução das respostas originais. A Tabela 8-3 ilustra os escores de índice e de escala que seriam atribuídos a cada padrão de resposta no nosso exemplo.

Como mencionei antes, a ponderação original do índice para os quatro tipos de escala seria mantida na construção da Escala de Guttman e não seriam cometidos erros na reprodução das respostas dadas aos três itens. Os tipos mistos seriam ponderados diferentemente, na tentativa de reduzir os erros. Mas note que um erro é cometido para cada respondente nos tipos mistos. No primeiro tipo misto, prediríamos erroneamente uma resposta científica ao item mais fácil para cada um dos dezoito entrevistados deste grupo, com um total de dezoito erros.

TABELA 8-3
Escore de índice e escala

	Padrões de Respostas			Número de Cartões	Escore de Índice	Escore de Escala	Total de Erros de Escala
Tipos de Escalas:	+	+	+	116	3	3	0
	+	+	-	127	2	2	0
	+	-	-	92	1	1	0
	-	-	-	48	0	0	0
Tipos Mistos:	-	+	-	18	1	2	18
	+	-	+	14	2	3	14
	-	-	+	5	1	0	5
	-	+	+	7	2	3	7

* Nota- Esta tabela apresenta um método comum para se ponderar "tipos mistos", mas o leitor deve saber que outros métodos também são usados.

O grau com que um conjunto de respostas empíricas forma uma Escala de Guttman é determinado pela exatidão com que as respostas originais podem ser reconstruídas a partir dos escores de escala. Para cada um dos 427 respondentes do exemplo, prediremos três respostas, com um total de 1.281 previsões. A Tabela 8-3 indica que faremos quarenta e quatro erros usando os escores de escala atribuídos. A porcentagem de previsões corretas é o *coeficiente de reproducibilidade*, ou seja, a porcentagem de respostas "reproduzíveis". No exemplo, o coeficiente de reproducibilidade é de 1.237/ 1.281, ou seja, 96,6%.

Exceto para o caso de reproducibilidade perfeita (100%), não temos como afirmar, de forma absoluta, se um grupo de itens forma ou não uma Escala de Guttman. Virtualmente, todos os conjuntos de itens desse tipo *aproximam-se* de uma escala. Como regra geral, coeficientes de 90% ou 95% são os padrões comumente usados para determinar se há ou não uma Escala de Guttman. Se a reproducibilidade observada exceder o nível que você estabeleceu, provavelmente você decidirá ponderar e usar os itens como uma escala.³

Um comentário final sobre a Escala de Guttman. Ela é baseada na estrutura observada entre os *dados reais que estão sob exame*. Este ponto importante é freqüentemente mal compreendido pelos pesquisadores. Não faz sentido dizer que um conjunto de itens de questionário (talvez desenvolvido e usado anteriormente por outro pesquisador) constitua uma Escala de Guttman. Podemos apenas dizer que os itens formam uma escala dentro do corpo de dados que está sendo analisado.

Portanto, a escalabilidade é uma questão empírica, dependente da amostra. Se um grupo de itens formar uma Escala de Guttman numa amostra, não há garantia de que formarão tal escala entre os respondentes de outra amostra. Então, itens de questionário em si, e por si sós, nunca formam uma escala, mas sim um conjunto de observações empíricas.

Tipologias

Concluimos este capítulo com uma breve discussão da análise e construção de *tipologias*. Lembre que índices e escalas são construídos para fornecer medidas ordinais de variáveis. Tentamos atribuir escores de índice ou escala aos respondentes, de modo a indicar um grau crescente de preconceito, religiosidade, conservadorismo, e assim por diante. Nestes casos, estamos lidando com dimensões únicas. Em contraste, tipologias são *multidimensionais*.

Muitas vezes você quer resumir as interseções entre duas ou mais dimensões. Por exemplo, examinando orientações políticas separadamente, em termos de temas domésticos e de política externa.

TABELA 8-4
Uma tipologia política

Atitudes para com a Política Doméstica	Atitudes para com a Política Externa	
	Conservador	Liberal
Conservador	A	B
Liberal	C	D

A apresentação 2 x 2 da Tabela 8-4 descreve esta tipologia. As pessoas na célula A da tabela são conservadoras em política externa e política doméstica; as na célula D são liberais em ambas. Os respondentes nas células B e C são conservadores numa questão e liberais na outra.

Freqüentemente, ao tentar construir um índice ou uma escala, você chega a uma tipologia. Os itens que você achou que representavam uma única variável parecem representar duas. Nesse exemplo, você podia estar tentando construir um índice único das atitudes políticas, mas descobriu —

empiricamente — que as políticas externa e doméstica tinham de ser mantidas separadas.

Você deve estar consciente da dificuldade inerente à análise tipológica. Sempre que a tipologia for usada como a *variável independente*, provavelmente não haverá problemas. No exemplo acima, você pode computar e apresentar as porcentagens de pessoas em cada célula que dizem que normalmente votam no Partido Democrata. Assim, você pode facilmente examinar os efeitos das atitudes em relação às políticas doméstica e externa no comportamento eleitoral.

Mas é difícil analisar a tipologia como *variável dependente*. Se quiser descobrir por que os entrevistados caem nas diferentes células da tipologia, você terá problemas. Isto fica aparente ao considerarmos as maneiras pelas quais se pode construir e ler as tabelas. Por exemplo, suponha que você queira examinar os efeitos do sexo sobre atitudes políticas. Com uma dimensão única, você poderia facilmente determinar as porcentagens de homens e mulheres classificados como conservadores e liberais no seu índice ou escala. Entretanto, com uma tipologia você teria que apresentar a distribuição dos homens em sua amostra entre os tipos A, B, C e D. A seguir, teria que repetir o procedimento para as mulheres e comparar as duas distribuições. Suponha que 80% das mulheres sejam classificadas como tipo A (conservadoras nas duas dimensões), comparados a 30% dos homens. Além disso, suponha que apenas 5% das mulheres sejam classificadas como tipo B (conservadoras apenas em relação às questões domésticas), comparados a 40% dos homens. Seria incorreto concluir, de um exame do tipo B, que os homens são mais conservadores em assuntos domésticos do que as mulheres, já que 85% das mulheres, comparados a 70% dos homens, têm essa característica. A escassez relativa de mulheres classificadas como tipo B se deve à sua concentração no tipo A. Vê-se que é muito difícil uma interpretação de tais dados em áreas não descritivas. Você provavelmente acabará examinando as duas dimensões políticas separadamente, especialmente se as variáveis tiverem mais categorias de respostas do que no nosso exemplo.

TABELA 8-5
Nove estilos de vida americanos

Nome	Descrição Típica
Estilo de vida sobrevivente	Muito pobre; baixa educação; muitos idosos; muitos doentes.
Estilo de vida sustentador	À beira da pobreza; freqüentemente encontrado nas favelas urbanas.
Estilo de vida pertencente	Classe média tradicional
Estilo de vida emulador	Lutando para crescer financeiramente; ambicioso; geralmente jovem.
Estilo de vida conquistador	Bem-sucedido na profissão e nos negócios; a imagem do sucesso americano.
Estilo de vida eu-sou-eu	Filhos de conquistadores, renunciando ao sucesso tradicional em busca de um novo estilo de vida.
Estilo de vida experimental	Jovem, educado, bem-sucedido, pessoas mais ou menos místicas que dão grande importância a questões internas.
Estilo de vida socialmente Consciente	Pessoas bem-sucedidas, maduras, influentes, preocupadas e ativas nas áreas de conscientização e responsabilidade social.
Estilo de vida integrado	Pessoas psicologicamente maduras que possuem preocupações internas e externas integradas.

FONTE - Adaptado de MITCHELL, Arnold. *The Nine American Lifestyles*. New York: Warner Books, 1983. p.4-24.

Uma tipologia mais complexa, desenvolvida por Arnold Mitchell e colegas do Instituto de Pesquisas de Stanford, a *tipologia Valores e Estilos de Vida* (VAEV), tem atraído bastante atenção, especialmente nos círculos de *marketing*.⁴ Se os americanos podem ser agrupados ou separados em termos de muitas variáveis (sexo, idade, raça, educação etc.), a pesquisa de Mitchell sobre várias variáveis demográficas e atitudes sugeriu a possibilidade de identificar a maioria deles com um dos nove estilos de vida apresentados na Tabela 8-5.

Como Mitchell e outros demonstraram, as pessoas identificadas com os nove principais estilos de vida têm padrões de consumo bem diferentes, sugerindo que um determinado produto pode ser anunciado de maneira efetiva entre alguns estilos de vida e não entre os outros. Vendas de Cadillacs provavelmente não serão altas entre os sobreviventes, nem as vendas de rifles entre os socialmente conscientes. Além disso, os pesquisadores que usaram o modelo VAEV puderam identificar áreas geográficas onde se concentram certos estilos

de vida, tornando possível, por exemplo, usar mala direta, anunciando apenas nos CEPs com maior probabilidade de vendas para um determinado produto.

Apesar das tipologias apresentarem problemas especiais para a análise de *survey*, o modelo VAEV demonstra que elas não devem ser totalmente abandonadas. É claro que, em última instância, você deve projetar e construir medidas compostas — índices, escalas ou tipologias —, apropriadas às necessidades da sua pesquisa.

Resumo

Este capítulo abordou a lógica e a construção de índices e escalas. As técnicas descritas são comuns na análise de *survey* e extremamente valiosas. Boas medidas compostas como índices e escalas oferecem as seguintes vantagens:

1. Onde indicadores únicos (por exemplo, respostas a um item de questionário) podem produzir uma medida tendenciosa da variável em questão, uma medida composta a partir de vários indicadores diferentes pode eliminar o efeito de viés.
2. Medidas compostas permitem maior amplitude de variação da variável. Enquanto um único item dicotômico fornece apenas dois níveis de intensidade (por exemplo, alto e baixo), a combinação de cinco itens resulta na criação de um índice ou escala variando de um baixo 0 até um alto 5. Uma medida composta adequadamente construída dá maior poder explicativo à análise.
3. Escalas podem prover uma eficiente técnica para reduzir dados. Um conjunto extenso de respostas de questionário pode ser resumido na forma de um único escore de escala, sem perder muito da informação original.
4. Tipologias fornecem classificações multidimensionais, que freqüentemente são eficazes como variáveis independentes, mas problemáticas enquanto variáveis dependentes.

Notas

¹ O exemplo, inclusive as tabelas, foi tirado de BABBIE, Earl. *Science and Morality in Medicine*. Berkeley: University of California Press, 1970.

² Se você tiver dificuldade para entender as tabelas apresentadas neste capítulo, consulte o Capítulo 14, que trata da construção e interpretação de tabelas.

³ A decisão sobre os critérios a usar a este respeito é, claro, arbitrária. Um alto grau de reproducibilidade não assegura que a escala construída de fato meça o conceito em questão, embora aumente a confiança de que todos os itens medem a mesma coisa. Um alto coeficiente de reproducibilidade é mais provável quando menos itens estiverem envolvidos.

⁴ MITCHELL, Arnold. *The Nine American Lifestyles*. New York: Warner Books, 1983.

Leituras Adicionais

LAZARFELD Paul F., ROSENBERG, Morris (Ed.). *The Language of Social Research*. New York: Free Press, 1955. seção 1.

OPPENHEIM, A. N. *Questionnaire Design and Attitude Measurement*. New York: Basic Books, 1966.

SELLTIZ, Claire et al. *Research Methods in Social Relations*. New York: Holt, Rinehart & Winston, 1959. cap.10.

Coleta de Dados

3

CAPÍTULO 9

QUESTIONÁRIOS AUTO-ADMINISTRADOS

CAPÍTULO 10

SURVEYS POR ENTREVISTAS

CAPÍTULO 11

PROCESSAMENTO DOS DADOS

CAPÍTULO 12

PRÉ-TESTES E ESTUDOS-PILOTO

Os quatro capítulos da Parte 3 abordam os vários aspectos da coleta de dados de *survey*. Examinaremos as atividades que resultam num corpo de dados para análise e consideraremos as decisões que você precisa tomar em relação às opções disponíveis.

Os Capítulos 9 e 10 focalizam as técnicas de coleta de dados usadas em questionários auto-administrados e *surveys* por entrevistas, respectivamente. O Capítulo 10 considera as entrevistas cara a cara e as feitas por telefone.

O Capítulo 11 trata dos diferentes métodos usados na conversão dos questionários completos em dados quantitativos manipuláveis. Também estudaremos alguns métodos novos, computadorizados, que criam arquivos de dados à medida que vão sendo coletados.

O Capítulo 12 considera um componente da pesquisa de *survey* profissional freqüentemente esquecido: estudos-piloto e pré-testes. Discutiremos as diferentes técnicas que você pode usar para testar os vários aspectos do desenho de pesquisa, antes de comprometer muitos recursos, e investigaremos como você pode avaliar os resultados desses testes.

Capítulo 9

Questionários Auto-Administrados

Os Capítulos 9 e 10 discutem como pesquisadores de *survey* realmente coletam dados para análise. Depois de construir um questionário apropriado aos objetivos de pesquisa, você precisará distribuir cópias do questionário à sua amostra de respondentes. Este capítulo considera as técnicas para distribuir questionários auto-administrados; o Capítulo 10 aborda os *surveys* por entrevistas.

Apesar do *survey* pelo correio ser a forma típica de pesquisa auto-administrada, há vários métodos adicionais. Em alguns casos, é conveniente administrar o questionário a um grupo de respondentes reunidos num único local ao mesmo tempo. Por exemplo, um *survey* de estudantes cursando Introdução à Psicologia pode ser feito durante a aula. Estudantes do ensino médio podem ser questionados durante o intervalo das aulas.

Recentemente têm sido feitas experiências com a entrega domiciliar de questionários. Os pesquisadores entregam os questionários nas residências dos respondentes e explicam a pesquisa; o questionário é deixado para ser respondido e depois recolhido pela equipe da pesquisa.

A entrega domiciliar e o correio podem ser usados juntos. Em muitas partes do país, o censo dos Estados Unidos é feito assim. Os questionários são enviados pelo correio às famílias; depois um enumerador do censo visita a residência para recolher o questionário e verificar se foi respondido completamente. Usando o método oposto, os pesquisadores entregam os questionários em mãos pedindo que o entrevistado o envie pelo correio, quando respondido, para o escritório da pesquisa.

No geral, a visita do pesquisador, na entrega do questionário, ou na coleta, ou ambos, parece produzir uma taxa maior de questionários respondidos do que quando se usa apenas o correio. Mais experimentação com este método provavelmente levará a técnicas adicionais para aumentar o número de questionários respondidos a custo reduzido.

O resto deste capítulo dedica-se aos *surveys* pelo correio, a forma mais comum de *survey* auto-administrado.

Distribuição e Retorno de Correspondência

Método Básico

O método básico para a coleta de dados pelo correio é enviar o questionário acompanhado de uma carta de explicação e do envelope de retorno. O entrevistado completa o questionário e o devolve ao escritório da pesquisa pelo correio, usando o envelope fornecido para isto.

Método alternativo

Em alguns casos, este processo pode ser ainda mais facilitado pelo uso de um questionário *autopostável*. O questionário é feito de tal modo que o endereço do escritório da pesquisa e a postagem estejam impressos no envelope. Após completar o questionário, ele pode ser deixado no correio sem necessidade de envelope.

Entretanto, você deve planejar com cuidado o uso deste método, porque o correio tem exigências especiais sobre a forma dos materiais que podem ser postados. Questionários, em particular, precisam ser selados de uma certa maneira. A selagem pode ser feita de vários modos. Se o questionário for impresso na forma de livreto, pode-se fazer uma capa com três dobras em vez de duas. Assim, a contracapa tem uma dobra extra, presa com fita adesiva. Quando o questionário é enviado pelo correio, a dobra a mais é posta na parte de trás. Após responder o questionário, o respondente pode desdobrar a contracapa, umedecer o adesivo e envolver o livreto com ela. Se o endereço do escritório da pesquisa e a postagem já estiverem impressas na dobra extra, o questionário pode ser colocado diretamente no correio para retorno.

Este método simplifica a devolução, pois não é necessário incluir envelope de retorno. O respondente não poderá perder o envelope de retorno sem também perder o questionário. Esta

forma de questionário tem o atrativo adicional do “valor de brinquedo”. Até certo ponto, os respondentes podem querer completar o questionário para brincarem com a contracapa.

Se não for possível produzir uma capa com dobra tripla, você pode afixar fitas adesivas ao livreto. Em vez de uma dobra envolvendo o questionário, você pode usar uma fita menor para selá-lo.

Enfim, você pode pedir aos respondentes para fecharem e selarem o livreto, talvez usando um grampeador ou fita adesiva. Este método é um pouco arriscado, por várias razões. Primeiro, forçar os entrevistados a terem mais trabalho provavelmente reduz a taxa de resposta. Na hora, eles podem não ter nada com que selar o envelope, deixar para depois e, eventualmente, esquecer de remeter o questionário. Além disso, a variedade de métodos de selagem que os entrevistados podem inventar provavelmente atrasa o processamento dos questionários retornados. (Já recebi questionários selados com cola, selos, cliques de papel e barbante.) Também, muitos respondentes não selam os envelopes e você pode ter dificuldades para recebê-los pelo correio.

Definitivamente, questionários autopostáveis têm muitas vantagens em termos de facilidade, economia e taxa de resposta obtida, mas devem ser planejados e pré-testados com cuidado. Além disso, é vital que todos os modelos experimentais sejam esclarecidos com os correios.

Opções Postais e Custos Relativos

Ao fazer uma pesquisa pelo correio, você tem várias opções para remeter e receber os questionários. As tarifas postais mudam constantemente, por isso você deve checá-las durante o planejamento da pesquisa e reservar fundos extras para este fim, se houver atraso para chegar ao campo.

O Correio deve ser procurado, antes de começar a pesquisa, para informações sobre custos e modalidades de postagem. Facilidades e dificuldades de cada modalidade de postagem, dependendo do tamanho da amostra, da sua abrangência territorial e das restrições de tempo devem ser pesadas e balanceadas, em relação aos fundos de que você dispõe.

Selos e Cartas Resposta Comerciais

Há duas opções básicas com relação à postagem de retorno para os questionários. Você pode fixar selos no envelope ou no questionário autopostável, ou imprimir uma autorização

de postagem, como nas cartas resposta comerciais. Os pesquisadores têm opiniões divergentes com relação aos méritos destes dois métodos.

A autorização de cartas resposta é similar à autorização da mala direta porque é impressa na correspondência em vez do selo de postagem. (Contate os correios para saber sobre maiores exigências de formato.)

Monitorando os Retornos

A postagem de questionários prepara o palco para uma atividade que pode ser muito valiosa para a pesquisa. À medida que os questionários retornam, não sente preguiçosamente na cadeira — faça um registro cuidadoso dos dados metodológicos.

Uma ferramenta inestimável é um gráfico de taxa de retorno. Marcamos como Dia 1, no gráfico, o dia em que os questionários são enviados; cada dia depois, marcamos no gráfico o número de questionários retornados. Como esta é uma atividade menor, geralmente é melhor fazer dois gráficos. Um deve mostrar o número de questionários retornados a cada dia, subindo e depois descendo à medida que muda o número de retornos. O outro deve informar o número ou a porcentagem *cumulativa* de retornos. Em parte, é uma atividade gratificante que lhe dá uma imagem da coleta de dados bem sucedida. O mais importante é que eles servem como guia de como está indo a coleta de dados. Se você planejar fazer correspondência de acompanhamento, o gráfico dá dicas de quando tais correspondências devem ser enviadas. (As datas de correspondências subseqüentes devem ser anotadas no gráfico.)

Assim que os questionários retornarem, devem ser abertos, examinados e numerados. Os números devem ser atribuídos em série, à medida que os questionários retornam, mesmo se outros números de identificação (ID) já tiverem sido atribuídos. Este procedimento tem muitas vantagens, como ilustram esses dois exemplos.

Suponha que você esteja estudando as opiniões sobre um político. E também que, no meio da coleta de dados, ele seja preso traficando drogas. Conhecendo a data desta informação e quando os questionários foram recebidos, você poderia determinar os efeitos desta informação.

Outra vantagem dos números seriados de ID é seu valor em ajudar na estimativa de vieses de não-respostas no *survey*.

Além de outros testes mais diretos de viés, você pode supor que os respondentes que não responderam o questionário são mais parecidos com os que demoraram a responder do que com os que responderam prontamente. Uma análise de questionários recebidos em diferentes momentos da coleta de dados pode ser usada para estimar vieses na amostra. Por exemplo, se as médias das notas informadas por alunos reduzirem-se constantemente durante a coleta de dados, com os que responderem rapidamente mostrando médias mais altas do que os que responderem mais tarde, você pode concluir que quem não respondeu teve notas mais baixas ainda. Apesar de não ser aconselhável fazer estimativas estatísticas de vieses desta maneira, você pode aproveitar estas estimativas aproximadas.

Se os respondentes tiverem sido identificados para correspondências de acompanhamento, as preparações para elas devem ser feitas à medida que os questionários retornam. O estudo de caso que apresentamos mais tarde neste capítulo discutirá este processo com mais detalhes.

Correspondências de Acompanhamento

A literatura metodológica sobre correspondências de acompanhamento mostra que elas são um método efetivo para aumentar as taxas de retorno nos *surveys* pelo correio. Em geral, quanto mais tempo os entrevistados em potencial gastarem para responder, menor é a probabilidade de eles responderem. Portanto, as correspondências de acompanhamento enviadas na hora certa são um estímulo adicional para o envio da resposta.

Os efeitos das correspondências de acompanhamento são vistos nas curvas de taxas de respostas registradas durante a coleta de dados. As correspondências iniciais são acompanhadas por uma subida e subseqüente queda dos retornos, e as correspondências de acompanhamento trazem novas subidas de retornos. Na prática, três correspondências (a original e duas de acompanhamento) parecem ser o mais eficiente.

É importante o momento de envio destas correspondências. A literatura metodológica oferece guias menos precisos, mas minha experiência diz que duas a três semanas é um intervalo razoável entre as correspondências. (Este período pode ser acrescido de alguns dias se o tempo de correspondência — para e de — for mais do que dois ou três dias.)

Fazer vários *surveys* em seqüência numa mesma população ajuda a desenvolver diretrizes mais específicas em relação às correspondências de acompanhamento. Uma vez, o Escritório de Pesquisas de *Survey* da Universidade do Havá fez vários *surveys* com estudantes, e pudemos aprimorar consideravelmente nosso procedimento de envio e reenvio de correspondência. Na verdade, encontramos um padrão consistente de retornos que nos pareceu transcender a diferença de conteúdo dos *surveys*, da qualidade do instrumento etc. Duas semanas depois da primeira postagem, aproximadamente 40% dos questionários retornaram; duas semanas depois do primeiro acompanhamento, mais 20% chegaram; e duas semanas após o último acompanhamento, chegaram mais 10%. Não podemos supor um padrão similar em *surveys* de populações diferentes, mas este exemplo mostra o valor de se tabelar cuidadosamente as taxas de retorno para cada *survey* feito.

As correspondências de acompanhamento podem ser administradas de várias maneiras. A mais simples é enviar aos que não responderam uma carta de encorajamento a participar. Um método ainda melhor é mandar nova cópia do questionário junto com a carta de acompanhamento. Se os respondentes em potencial não tiverem retornado os questionários após duas ou três semanas, há uma grande probabilidade de os questionários não terem chegado ou terem sido perdidos. Receber uma carta de acompanhamento pode encorajar os respondentes a procurarem o questionário original, mas se ele não for encontrado facilmente, a carta pode ir para o lixo. (As taxas de resposta relatadas antes foram obtidas quando se enviou questionários adicionais.)

Se os indivíduos na amostra do *survey* não forem identificados no questionário, será difícil reenviá-lo apenas para quem não respondeu. Neste caso, você deve enviar a correspondência de acompanhamento a todos os membros iniciais da amostra, agradecendo os que já participaram e encorajando os outros a fazê-lo. (O caso de pesquisa narrado mais tarde neste capítulo descreve outro método que pode ser usado num *survey* anônimo pelo correio.)

Demos muita atenção às maneiras de aumentar as taxas de respostas nos *surveys* pelo correio. Numa pesquisa recente, por exemplo, Richard J. Fox e colegas¹ fizeram uma revisão de oitenta e dois *surveys* cujos desenhos incluíam testes de diferentes técnicas para aumentar a taxa de resposta, e fizeram algumas descobertas consistentes. Por exemplo, patrocínio de universidade parece aumentar taxas de resposta. Também é importante o número de contatos com os respondentes. Ajuda

também enviar uma carta prévia avisando o envio de um questionário e um cartão de acompanhamento. Usar postagem de primeira classe nos envelopes de envio pode produzir taxas um pouco mais altas de respostas do que as malas diretas que passam na máquina de postagem do correio, mas as evidências a respeito não são consistentes. Os autores informaram também que a cor do questionário pode afetar as taxas de resposta (verde é melhor que branco).

Taxas Aceitáveis de Resposta

Uma pergunta que novos pesquisadores de *survey* fazem freqüentemente se refere à porcentagem da taxa de retorno que deve ser alcançada num *survey* pelo correio. Vale a pena repetir que a estatística inferencial usada na análise do *survey* presume que todos os membros da amostra inicial irão completar e retornar seus questionários. Como isto quase nunca acontece, o viés de resposta se torna uma preocupação; você deve testar (e torcer para) a possibilidade de os respondentes serem essencialmente uma amostra aleatória da amostra inicial e, portanto, uma amostra aleatória um pouco menor da população total.²

Mesmo assim, a taxa geral de resposta é um guia da representatividade dos respondentes na amostra. Alcançar uma alta taxa de respostas resulta em chance menor de viés de resposta significativo do que uma baixa taxa de respostas. Mas o que é uma alta taxa de resposta?

Uma rápida revisão da literatura de *survey* mostra uma ampla variação de taxas de resposta. Cada taxa pode vir acompanhada de uma afirmação como "Esta é considerada uma taxa de resposta relativamente alta para um *survey* deste tipo". (Um senador americano fez esta declaração sobre uma pesquisa de constituintes cuja taxa de retorno foi de 4%.) Apesar da grande variedade nas taxas de retorno e das reações a elas, podemos ter algumas regras básicas.

Uma taxa de resposta de pelo menos 50% é geralmente considerada *adequada* para análise e relatório. Uma taxa de resposta de pelo menos 60% é considerada *boa*, e uma taxa de 70% ou mais é *muito boa*. Mas estas são regras rudimentares, sem base estatística, e uma falta de viés de resposta demonstrada é muito mais importante do que uma alta taxa de resposta. Os artigos nos volumes de 1988 da *American Sociological Review* e da *Public Opinion Quarterly* comunicaram vinte e seis taxas de respostas de *surveys* com uma média de 69%.³

Ao computar taxas de resposta, a prática aceita é omitir os questionários que não puderam ser entregues. No relatório metodológico, você deve informar o tamanho da amostra inicial e subtrair o número dos que não puderam ser entregues devido a endereços errados, mortes e coisas assim. O número de questionários completos é então dividido pelo tamanho líquido da amostra para calcular a taxa de resposta. O resultado é que a taxa de resposta será realmente uma medida do seu sucesso em persuadir os membros da amostra a participarem; você não deve contar contra si mesmo os membros da amostra que não conseguiu contatar.

Embora este procedimento seja a prática aceita, você deve conhecer a suposição lógica sobre a qual ele se baseia — os questionários não entregues representam uma amostra aleatória da amostra inicial. Claro, este pode não ser o caso. As pessoas cujos questionários não puderem ser entregues são, pelo menos, provavelmente mais móveis do que as outras, e mobilidade pode estar relacionada a muitas outras variáveis. Os melhores guias ainda são os testes de viés para a falta de respostas.

Um Caso de Pesquisa

Os muitos passos a dar na administração de um *survey* pelo correio podem ser mais bem-compreendidos através de um estudo de caso. Portanto, concluímos o capítulo com a descrição detalhada de um *survey* feito com estudantes da Universidade do Havaí. Como você verá, a pesquisa não representou o ideal teórico de tais pesquisas, mas ela serve muito bem para nosso propósito agora.

Seleção da Amostra

O desenho e a seleção da amostra para esta pesquisa foram apresentados como um estudo de caso no Capítulo 6. Lembre que cerca de 1.100 estudantes foram selecionados a partir do cadastro da universidade, através de um procedimento de amostragem estratificado e sistemático. Para cada estudante selecionado, o computador imprimiu seis etiquetas adesivas para correspondência.

Quando a equipe da pesquisa estava pronta para distribuir o questionário, percebeu que o orçamento da pesquisa era insuficiente para as correspondências para toda a amostra de 1.100 estudantes, porque os custos de impressão dos

questionários foram mais altos do que o esperado. Por isso, foi feita uma amostra sistemática das etiquetas de dois terços da amostra, produzindo uma sub-amostra com 770 estudantes.

Cartões postais

Decidimos que o *survey* seria anônimo, esperando encorajar respostas mais francas a algumas perguntas mais sensíveis. (*Surveys* subsequentes sobre os mesmos temas com a mesma população mostraram que o anonimato era desnecessário.) Portanto, os questionários não teriam a identificação dos estudantes. Ao mesmo tempo, esperávamos que os custos das correspondências de acompanhamento podiam ser reduzidos, insistindo apenas com os que não respondessem.

Desenvolvemos um método especial de cartão postal para alcançar ambos os objetivos. Cada estudante recebeu um questionário sem identificação e um cartão postal endereçado ao escritório da pesquisa, com uma das etiquetas de postagem do estudante fixadas no verso. A carta de apresentação pedia aos estudantes que completassem e remetessem o questionário — assegurando o anonimato — e, simultaneamente, remetessem o cartão. O recebimento dos cartões indicaria que eles haviam retornado os questionários, sem identificar qual questionário era de quem. Este procedimento facilitaria então o envio das correspondências de acompanhamento.

Questionário

O questionário de trinta e duas páginas foi impresso no formato de livreto (*foto-offset* e grampeado). Uma capa de três dobras, descrita antes, permitia o retorno do questionário sem necessidade de envelope adicional.

Carta de Apresentação

Uma carta apresentando a pesquisa e seu propósito foi impressa na capa do livreto. Ela explicava por que a pesquisa estava sendo feita (saber como os estudantes se sentiam com relação a vários assuntos), como eles tinham sido selecionados, a importância de cada um responder e a mecânica de retorno do questionário.

Assegurou-se aos estudantes de que o estudo seria anônimo, e foi explicada a lógica do método do cartão postal. Seguiu uma declaração com relação ao patrocínio da pesquisa e

um número de telefone para quem quisesse mais informações. (Cerca de cinco estudantes ligaram.)

Imprimindo a carta de apresentação no questionário, a equipe da pesquisa eliminou a necessidade de anexar uma carta separada no envelope, simplificando a tarefa de montar as correspondências.

Montagem das Correspondências

A montagem do material para a postagem inicial seguiu os seguintes passos: (1) uma etiqueta de endereçamento para cada estudante foi colada no cartão postal; (2) outra etiqueta foi colada num envelope comercial para a remessa; (3) em cada envelope foram colocados um cartão postal e um questionário, checando se o nome no cartão postal era o mesmo do envelope.

A montagem foi feita através de um procedimento de linha de montagem, envolvendo os vários membros da equipe da pesquisa. Apesar de o procedimento ter sido organizado com antecedência, foi necessária uma certa prática antes de encontrar a melhor alocação de pessoas nas tarefas. Vale notar que houve um atraso para trocar os envelopes por outros maiores, o que teria sido evitado se o processo de montagem tivesse sido feito antes.

Postagem

A distribuição dos questionários foi preparada para ser feita por mala direta. Depois de pôr os questionários nos envelopes, eles foram agrupados por CEP, separados em pacotes e entregues nos correios.

Recepção dos Questionários Retornados

Logo após a postagem inicial, questionários e cartões começaram a chegar ao escritório da pesquisa. Os questionários foram abertos, lidos cuidadosamente e receberam números de identificação como descrito anteriormente.

Mas o processamento dos cartões postais mostrou um deslize no desenho da amostra. Você deve lembrar, da discussão anterior sobre o desenho da amostra (Capítulo 6), que a distribuição final dos estudantes na moldura da amostragem foi feita através do número do Seguro Social, assim fornecendo uma quase estratificação, por região de origem.

Por isso, as etiquetas de postagem foram impressas nesta ordem (por estrato de classe). Mas os números do Seguro Social não foram impressos nas etiquetas, porque isso não era relevante para a pesquisa.

Tendo um cartão postal com um nome e um endereço, foi muito difícil localizar as etiquetas correspondentes entre as restantes nas várias folhas impressas pelo computador. Assim, foi preciso retirar todas as etiquetas das folhas impressas e pô-las em ordem alfabética. Deste modo, localizar a etiqueta de um certo estudante podia ser feito com um mínimo de esforço. (*Nota:* As etiquetas podiam ter sido impressas em ordem alfabética desde o início, se este problema tivesse sido antecipado.) Para cada cartão recebido, procuravam-se as etiquetas daquele estudante e elas eram destruídas.

Correspondências de Acompanhamento

Após duas ou três semanas, as etiquetas de postagem restantes foram usadas para organizar as correspondências de acompanhamento. Foram repetidos os procedimentos de montagem descritos, com uma exceção. Foi incluída na correspondência uma carta de apelo preparada especialmente. A nova carta dizia que muitos estudantes já tinham retornado seus questionários e que era muito importante que todos o fizessem. A carta também dizia que os cadastros do escritório de pesquisa podiam estar errados e, se os estudantes já tivessem retornado os questionários, deviam ignorar aquela correspondência e aceitar os agradecimentos pela ajuda. Se não tivessem ainda participado, eram encorajados a fazê-lo.

Essa correspondência de acompanhamento estimulou os retornos, como esperado, e os mesmos procedimentos de registro foram seguidos. Os cartões foram usados para destruir as etiquetas adicionais de postagem. Infelizmente, as pressões financeiras e de tempo tornaram impossível a remessa de uma terceira correspondência, como tinha sido planejado originalmente, mas as duas correspondências enviadas resultaram numa taxa de resposta de 62%.

Resumo

Este capítulo abordou as bases da coleta de dados através de questionários auto-administrados. Precisamos fazer duas observações.

Primeira, acredito que tais questões corriqueiras são vitais para a pesquisa científica de alta qualidade. O pesquisador que acredita que a qualidade de um projeto de pesquisa é função unicamente de capacidades analíticas e teóricas está muito enganado. Um desenho de pesquisa brilhante executado imprópriamente resultará em fracasso. Nenhum detalhe é pequeno ou corriqueiro demais para ser ignorado.

Segunda, a crônica detalhada de um *survey* pelo correio, com todos os seus problemas, deve trazer à tona a importância dos pré-testes e estudos-piloto relacionados a todos os aspectos do desenho da pesquisa. Somente testes avançados podem revelar problemas que podem eventualmente afundar toda uma pesquisa.

Notas

¹ FOX, Richard J., CRASK, Melvin R., KIM, Jonghoon. Mail *Survey* Response Rates. *Public Opinion Quarterly*, v.52, p.467-491, inverno 1988.

² Para um exame mais detalhado de vieses de não-resposta, ver DONALD, Marjorie N. Implications of Nonresponse for the Interpretation of Mail Questionnaire Data. *Public Opinion Quarterly*, v.24, n.1, p.99-114, 1960, e BROWNLEE, K. A. A Note on the Effects of Nonresponse on *Surveys*. *Journal of the American Statistical Association*, v.52, n.277, p.29-32, 1957.

³ Alguns artigos informam mais de uma amostra (e taxa de resposta), e um par deles se refere a técnicas de teste visando melhorar taxas de resposta.

Leituras Adicionais

DILLMAN, Don A. *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons, 1978.

FOX, Richard J., CRASK, Melvin R., KIM, Jonghoon. Mail *Survey* Response Rates. *Public Opinion Quarterly*, v.52, p.467-491, inverno 1988.

Capítulo 10

Surveys Por Entrevistas

O Capítulo 9 focalizou os questionários auto-administrados; este capítulo apresenta métodos alternativos para a coleta dos dados. Nos *surveys* por entrevistas, em vez de pedir aos respondentes para lerem o questionário e escreverem as respostas, os entrevistadores perguntam oralmente e anotam as respostas. A entrevista é feita em encontros cara a cara ou por telefone. Este capítulo discute os dois métodos de entrevistas.

Importância do Entrevistador

Há várias vantagens em ter um questionário aplicado por entrevistadores, em lugar de respondido pelos próprios respondentes. *Surveys* por entrevistas têm em geral taxas mais altas de respostas do que os pelo correio.¹ Um *survey* por entrevista adequadamente desenhado e executado deve conseguir uma taxa de respostas de pelo menos 80 a 85%. Parece que os respondentes relutam mais em despachar um entrevistador na porta de suas casas, ou mesmo ao telefone, do que em jogar fora um questionário que veio pelo correio.

A presença de um entrevistador geralmente reduz a quantidade de “não sei” e “sem resposta”. Se a minimização de tais respostas é importante para a pesquisa, o entrevistador pode ser instruído a buscar respostas. (“Se tivesse que escolher *uma* das respostas, qual você acha que se *aproximaria* mais dos seus sentimentos?”)

Os entrevistadores também podem evitar confusões com os itens do questionário. Se o entrevistado não entender a intenção da pergunta ou indicar que não a entende, o entrevistador pode clarear o assunto e assim obter respostas relevantes. (Entretanto, tais explicações devem ser estritamente controladas através das especificações formais que discutiremos depois neste capítulo.)

Finalmente, o entrevistador pode *observar* enquanto faz as perguntas. Por exemplo, um entrevistador cara a cara pode anotar a raça do entrevistado se esta pergunta for considerada muito delicada. Observações similares podem ser feitas em relação à qualidade da residência do respondente, à posse de certos bens e outras informações. Mesmo por telefone, os entrevistadores podem perceber o domínio da língua pelo respondente, reações gerais à pesquisa etc.

Num *survey* por entrevista de estudantes, eles receberam um pequeno questionário auto-administrado, relativo a atitudes e comportamentos sexuais, para completar durante a entrevista. Enquanto o estudante completava o questionário, o entrevistador fazia anotações detalhadas com relação aos cuidados pessoais e roupas do entrevistado.

Papel Neutro do Entrevistador

A pesquisa de *survey* é, por necessidade, baseada numa teoria irrealista de cognição e comportamento de *estímulo-resposta*. Deve-se presumir que cada pergunta e cada resposta significam exatamente a mesma coisa para todos os respondentes. Apesar de o objetivo ser impossível, as perguntas são preparadas para se aproximarem o máximo desse ideal.

O entrevistador também precisa se ajustar a essa situação ideal. Sua presença não deve afetar a percepção que o respondente tem da questão, ou da resposta. Portanto, o entrevistador deverá ser um meio *neutro*, através do qual perguntas e respostas são transmitidas.

Se isso for feito, diferentes entrevistadores obterão exatamente as mesmas respostas do mesmo entrevistado. A neutralidade do entrevistador tem importância especial nos *surveys* cara a cara, no domicílio do entrevistado. Para economizar tempo e dinheiro, um entrevistador é designado para todas as entrevistas numa determinada área geográfica — um quarteirão ou um bairro. Se ele fizer qualquer coisa para afetar

as respostas obtidas, o viés decorrente poderá ser interpretado como uma característica da área estudada.

Suponha que um *survey* cara a cara esteja sendo feito para medir atitudes com relação a moradias de baixo custo, para ajudar a escolher o local de um novo programa habitacional do governo. Um entrevistador designado para um certo bairro pode, por gestos ou palavras, comunicar seu desgosto com financiamentos a moradias de baixo custo. Os entrevistados podem, então, tender a dar respostas concordando com a posição do entrevistador. Os resultados do *survey* indicariam que os residentes desta vizinhança resistem fortemente à construção das moradias em sua área.

Em alguns *surveys*, a compatibilidade racial — tentar compatibilizar a raça/etnia de entrevistados e entrevistadores — pode ser recomendável. Numa análise secundária dos Estudos sobre as Eleições Nacionais de 1964 a 1984, Anderson, Silver e Abramson² descobriram que os entrevistados negros tinham maior probabilidade de expressar sentimentos de calor e proximidade em relação a brancos, quando eram entrevistados por brancos, do que quando eram entrevistados por negros. Os autores até sugeriram que uma aparente mudança na hostilidade de negros contra brancos, entre 1976 e 1984, podia ser atribuída a um aumento no número de entrevistadores negros. Reese e colegas³ relataram que a etnia tem um impacto similar no caso de entrevistadores e entrevistados hispânicos.

Estes exemplos realçam a importância da neutralidade do entrevistador no processo de coleta de dados e a dificuldade de alcançá-la completamente.

Regras Gerais para Entrevistar

A maneira de conduzir as entrevistas depende, de certo modo, da população e do conteúdo do *survey*. Mesmo assim, é possível sugerir diretrizes gerais que se aplicam à maioria das, se não a todas, situações de entrevistas.

Aparência e Comportamento

Como regra geral, os entrevistadores devem se vestir como as pessoas que vão entrevistar. Um entrevistador vestido ricamente provavelmente terá dificuldade em obter cooperação

e boas respostas de entrevistados mais pobres, e um mal vestido terá dificuldades semelhantes com entrevistados mais ricos.

Havendo diferença entre o trato e as roupas do entrevistador e do entrevistado, ela deve ser na direção da limpeza, simplicidade e modéstia. Mesmo não sendo aceitas por todas as camadas sociais, a limpeza e a simplicidade da classe média continuam sendo as normas básicas com maior probabilidade de serem aceitas pela maioria dos entrevistados.

Roupas e trato são sinais típicos das atitudes e orientações de uma pessoa. Uma jovem usando jaqueta de couro preta, botas de engenheiro e tatuagem de caveira na testa certamente comunicará — correta ou incorretamente — que é antiautoritária politicamente, sexualmente permissiva, fisicamente agressiva, a favor das drogas, e assim por diante. Possivelmente, não obterá respostas francas em relação aos *Hell's Angels* ou a roqueiros *punk*.

Apesar de tema da aparência ser o mais óbvio no caso das entrevistas cara a cara, lembre também que comunicamos identidade através da fala. Os entrevistadores por telefone não estão isentos dessas preocupações. Considere estas introduções (não recomendadas) a entrevistas telefônicas:

Ôi, escuta só essas perguntas e responde. Deu prá entender?...
Ei, tô falando com você.

Olha aí, tenho que te fazer algumas perguntas, tá legal, e eu preciso de, olha aí, mais ou menos, pegar suas respostas, tá legal? Isto não é demais, cara?

Bom, estou fazendo uma pesquisa de *survey*, é isto aí. E preciso de suas respostas prá ganhar uma grana.

Quanto a comportamento, entrevistadores devem ser pelo menos agradáveis. Já que vão xeretar a vida pessoal e as atitudes dos entrevistados, devem comunicar um interesse genuíno em conhecê-las sem parecerem espíões. Entrevistadores devem ser relaxados e amigáveis sem serem casuais demais ou importunos. Uma das capacidades naturais mais importantes para entrevistadores é a de encontrar logo o tipo de pessoa com quem o entrevistado se sentirá mais à vontade; em outras palavras, o tipo de pessoa com quem o entrevistado gostará mais de conversar. Há dois aspectos com relação à importância de um comportamento agradável. A entrevista terá mais sucesso se o entrevistador conseguir ser o tipo de pessoa com quem o

entrevistado se sente mais à vontade. Ao mesmo tempo, como se pede aos entrevistados para gastar seu tempo dando informações pessoais sobre si mesmos, eles merecem a experiência mais agradável que o autor da pesquisa e o entrevistador puderem lhes proporcionar.

Familiaridade com o Questionário

Se os entrevistadores não estiverem familiarizados com o questionário, a pesquisa sofrerá e será posto um peso injusto sobre o entrevistado. A entrevista provavelmente tomará mais tempo do que o necessário e, em geral, será incômoda. O entrevistador não conseguirá familiaridade passando os olhos no questionário duas ou três vezes. Ele deve ser estudado cuidadosamente, pergunta por pergunta, e o entrevistador precisa praticar sua leitura em voz alta. (O treinamento do entrevistador será discutido com mais detalhes mais tarde neste capítulo.)

O entrevistador deve poder ler os itens do questionário sem errar nem gaguejar nas frases e palavras. Um bom guia para entrevistadores é o estilo de um ator lendo as falas num filme. Elas devem ser lidas naturalmente, como se fizessem parte de uma conversa natural, mas a conversa tem que acompanhar exatamente o fraseado do questionário. O entrevistador não deve tentar memorizar o questionário.

Do mesmo modo, o entrevistador deve estar familiarizado com as instruções que acompanham o questionário. Inevitavelmente, alguns itens não se ajustarão exatamente à situação de algum entrevistado, surgindo a dúvida de como tais perguntas devem ser interpretadas numa determinada situação. As instruções fornecidas ao entrevistador devem orientá-lo adequadamente em casos como este, mas o entrevistador deve estar suficientemente familiarizado com a organização e o conteúdo das instruções para usá-las eficientemente. É melhor deixar uma pergunta sem resposta do que gastar cinco minutos procurando esclarecimentos e/ou tentando interpretar as instruções relevantes.

Seguindo Exatamente a Redação da Questão

O Capítulo 7, que discute a conceituação e a construção de instrumentos, mostrou o efeito da redação da pergunta sobre a resposta obtida. Portanto, uma pequena mudança de

redação pode levar o entrevistado a responder sim ao invés de não. Você pode redigir as perguntas cuidadosamente, visando obter a informação que precisa, assegurando-se de que os entrevistados interpretarão os itens de forma apropriada, e depois ver seu trabalho jogado fora se os entrevistadores fizerem as perguntas com suas próprias palavras.⁴

Registrando as Respostas de Maneira Exata

Sempre que os questionários tiverem perguntas abertas, ou seja, perguntas pedindo a própria resposta do entrevistado, é muito importante que o entrevistador anote as respostas exatamente como foram faladas. Não se deve tentar resumir, parafrasear ou corrigir erros de gramática. A resposta deve ser escrita exatamente como foi dada.

Registrar as respostas de maneira exata é especialmente importante porque o entrevistador não sabe como as respostas serão codificadas antes de serem processadas; você pode não perceber isto até ter a oportunidade de ler cem ou mais respostas. Por exemplo, o questionário pode perguntar como os entrevistados se sentem com relação à situação do trânsito em suas comunidades. Um entrevistado pode responder que há muitos carros nas ruas e que deveria se fazer algo para limitar essa quantidade. Outro pode dizer que há necessidade de mais ruas. Se o entrevistador resumi-las como “tráfego congestionado”, você não poderia aproveitar as importantes diferenças entre as respostas originais.

Algumas vezes um entrevistado é tão desarticulado que a resposta fica ambígua demais para permitir interpretação, mas o entrevistador é capaz de entender a intenção da resposta através dos gestos ou tom de voz do respondente. Neste caso, a resposta exata deve ser anotada, e o entrevistador deve acrescentar comentários à margem, dando sua interpretação e as razões para ela.

De modo geral, é útil ter comentários marginais explicando os aspectos da resposta não transmitidos verbalmente, como a aparente incerteza, raiva, embaraço etc. do entrevistado ao dar a resposta. Em cada caso, a resposta verbal exata deve ser registrada.

Sondando Respostas

Algumas vezes, os entrevistados respondem uma pergunta inapropriadamente. Por exemplo, a pergunta pode apresentar uma declaração e pedir que o entrevistado “concorde fortemente, concorde mais ou menos, discorde mais ou menos, ou discorde fortemente”, e o entrevistado pode responder “acho que é verdade”. Neste caso, você deve completar esta resposta perguntando “Você diria que concorda *fortemente* ou concorda *mais ou menos*?” Se necessário, pode explicar que foi instruído a marcar uma dessas categorias. Se o entrevistado se recusar peremptoriamente a escolher, você deve aceitar sua decisão polidamente e escrever a resposta exatamente como dada por ele.

Sondar respostas é, mais comumente, necessário para obter respostas a perguntas abertas. Por exemplo, face à pergunta prévia sobre condições de tráfego, o entrevistado pode dizer “muito ruim”. O entrevistador pode conseguir uma resposta mais elaborada de vários modos. Algumas vezes, o melhor é o *silêncio*; se o entrevistador se senta calmamente com o lápis em posição, o entrevistado provavelmente acrescentará alguns comentários. (Esta técnica é muito usada por repórteres de jornais.) Boas sondagens verbais podem ser “Como assim?” ou “De que maneira?” Talvez a mais usada seja “Mais alguma coisa?”

Freqüentemente, o entrevistador precisa tornar as respostas mais informativas, para propósitos analíticos. Entretanto, em todos os casos, é imperativo que as sondagens sejam completamente *neutras*. Não se deve, de maneira alguma, afetar a resposta dada. Se você antecipar que uma pergunta pode necessitar sondagem para obter a resposta apropriada, deve fazer uma ou mais sugestões, próximas à pergunta no questionário. Esta prática tem duas vantagens. Primeiro, você terá mais tempo para refletir e escolher a sondagem melhor e mais neutra. Segundo, todos os entrevistadores farão a mesma coisa, quando necessário. Portanto, mesmo se a sondagem não for inteiramente neutra, todos os entrevistados estarão, pelo menos, sujeitos ao mesmo estímulo. Esta diretriz lógica é a mesma discutida com relação à redação das perguntas. Apesar de uma pergunta não dever ser carregada ou tendenciosa, todos os entrevistados devem pelo menos ser submetidos à mesma pergunta, mesmo uma tendenciosa.

Treinamento do Entrevistador

A maioria dos *surveys* por entrevistas requer vários entrevistadores, e algum dia você terá que treinar e supervisionar uma equipe de entrevistadores. Portanto, vamos discutir o que está envolvido neste trabalho.

Mesmo tendo uma equipe de entrevistadores inteligentes e experientes, é essencial um curso de treinamento dos entrevistadores.⁵ Todos os *surveys* e questionários diferem uns dos outros, e os entrevistadores devem ser retreinados para cada novo *survey*. O tempo gasto nisso depende do escopo e da natureza do *survey* e da experiência relativa dos entrevistadores. Uma pesquisa domiciliar grande usando uma combinação de entrevistadores experientes e inexperientes pode normalmente exigir duas semanas de treinamento: uma semana de instrução em aulas e outra de prática no campo. Para uma pesquisa curta por telefone, os entrevistadores podem algumas vezes ser treinados em uma ou duas sessões longas à noite.

Instruções Gerais

O treinamento deve começar com comentários gerais sobre a natureza do *survey* e seu propósito. Você nunca deve considerar os entrevistadores como autômatos que não pensam ou como técnicos que podem ser instruídos para executar tarefas mecanicamente. Em geral, se os entrevistadores compreenderem por que o *survey* está sendo feito e que ele é um empreendimento importante, serão mais cuidadosos e diligentes no trabalho.

A descrição geral deve incluir o patrocinador do *survey*, seu propósito básico e como a amostra de entrevistados foi selecionada. Os entrevistadores também devem receber uma descrição geral dos outros passos do *survey*, inclusive como o questionário foi desenhado, como os dados serão processados e analisados. Se eles entenderem onde se encaixam nesse processo, provavelmente se sentirão mais integrados à equipe da pesquisa e responderão de acordo com esse sentimento.

Os detalhes administrativos devem ser explicados logo no início. Quanto tempo durará a fase de entrevistas? Quantas horas por semana deverão trabalhar os entrevistadores? Quando serão pagos? O entrevistador preocupado com estes detalhes não conseguirá se concentrar nas minúcias do questionário.

Mesmo se todos os entrevistadores forem experientes, as sessões de treinamento devem cobrir as diretrizes e regras gerais de se entrevistar. É útil preparar um conjunto de normas gerais para distribuir e discutir com os entrevistadores. Os mais experientes freqüentemente podem especificar algumas das regras através de exemplos de experiências passadas, e os novatos terão assim uma idéia melhor do que os espera numa verdadeira entrevista.

Estudando o Questionário e as Especificações

Os entrevistadores, o supervisor e/ou o diretor do projeto devem percorrer todo o questionário passo a passo. Deve ser explicado o propósito de cada item e discutidas todas as possíveis ambigüidades. É crucial que esta parte do treinamento seja conduzida de maneira informal, mais como discussão do que como aula. Os entrevistadores que não compreenderem um determinado ponto devem ser encorajados a pedir imediatamente explicações mais detalhadas. Sua resposta nunca deve demonstrar insatisfação, senão os entrevistadores poderão sentar-se em silêncio, sem entender bem o que se espera deles.

O exame do questionário deve incluir o estudo das especificações. Você deve ler a pergunta e as categorias de repostas, explicar a intenção da pergunta, descrever algumas situações simples e a ação correta do entrevistador e, a seguir, descrever algumas questões mais complicadas, passar para a seção correspondente das especificações e mostrar como elas resolveriam cada situação.

A maior parte da discussão durante esta fase tomará a forma de "O que eu faço se o entrevistado disser...?" Onde for apropriado, você deve usar as especificações para mostrar como a situação pode ser resolvida. Enfim, deve-se descrever a forma mais desejável de lidar com a situação. Os entrevistadores devem saber exatamente como você lidaria com a situação.

Mas esse assunto tem outro lado. Freqüentemente, surge uma situação que você não pode resolver facilmente, cuja ocorrência pode não ter passado por sua cabeça, e as especificações não falarão sobre ela. Nestes casos, não tome decisões rápidas; em vez disso, prometa que a questão será resolvida na próxima reunião. Trabalhe, então, para achar uma solução apropriada, determinar como ela se encaixa na lógica da intenção da pergunta, e prepare-se para explicar sua lógica. A solução deve ser explicada aos entrevistadores na reunião seguinte, e

as especificações atualizadas para levá-la em conta. O pior a fazer é dar respostas arbitrárias, nítidas, sem o cuidado de integrá-las à lógica geral da pergunta e das especificações. Tais respostas vão se somando até os entrevistadores terem muitas instruções específicas para aprender, não mais tendo uma compreensão lógica geral para guiá-los nas próprias interpretações.

Ensaie as Entrevistas em Sala de Aula

Depois de estudar o questionário e as especificações, deve-se treinar a prática de entrevistas. Para começar, deve-se fazer duas ou três entrevistas na frente da turma. O melhor começo é você entrevistar um supervisor ou vice-versa. Como estas entrevistas servirão como modelos para os entrevistadores, devem ser feitas precisamente de acordo com as instruções gerais e específicas dadas antes. Se estiver sendo entrevistado, informe suas próprias condições e atitudes, ou assuma uma identidade fictícia, mas consistente. Quando estiver entrevistando, deve agir exatamente como em uma entrevista verdadeira.

Durante o ensaio da entrevista, às vezes algum fato específico precisará de explicação para a turma, mas isto não deve quebrar o clima de estar fazendo uma entrevista verdadeira. Ao fim de cada entrevista, a turma deve ser encorajada a discuti-la, fazer perguntas, e assim por diante.

Depois das entrevistas de exercício em frente à turma, os entrevistadores devem ser emparelhados para ensaiar, entrevistando um ao outro. Estas entrevistas não precisam ser feitas na frente da turma, devendo ocorrer simultaneamente. Ande pela sala, parando e escutando as entrevistas, mas na maioria das vezes é melhor tomar nota e depois discutir as questões com a turma ao fim da sessão de exercício. Os entrevistadores devem ser reemparelhados e fazer mais exercícios. Esta parte do treinamento deve alternar a prática e a discussão.

Exercitando as Entrevistas de Campo

Depois das instruções em sala de aula, os entrevistadores devem ensaiar entrevistas no campo, indo às casas das pessoas ou telefonando para elas, dependendo do desenho da pesquisa. Esta parte do treinamento deve ser exatamente como a situação real. A designação dos entrevistados e/ou endereços deve ser feita como no *survey* real. Normalmente, o responsável pela seleção da amostra fornece uma lista dos respondentes não selecionados para a amostra principal.

Os entrevistadores devem receber exatamente o mesmo material do *survey* real. Farão contato com os entrevistados e os entrevistarão. (Os entrevistados não devem saber que estão fazendo uma entrevista de treino, isso impediria todo o realismo do exercício.) Neste momento, os entrevistadores devem exercitar todos os procedimentos administrativo: preencher as tabelas de horários, os registros etc. O ensaio da entrevista de campo fornece um estudo-piloto de toda a operação de entrevistas.

Normalmente, basta os entrevistadores ensaiarem cerca de cinco entrevistas de campo, apesar de este número poder variar de acordo com a natureza do *survey*. Um aspecto da operação de entrevistar pode ser modificado no ensaio. No *survey* real, os entrevistadores podem ter que fazer um número mínimo especificado de visitas ao entrevistado, antes de decidir que ele não está "disponível". Para facilitar os ensaios, você pode omitir este aspecto. Para compensar, cada entrevistador deve receber número suficiente de endereços para assegurar que serão feitas entrevistas suficientes.

À medida que os entrevistadores terminam as entrevistas de exercício, devem rever seu trabalho com o supervisor de entrevistas. O supervisor deve percorrer todo o questionário com cada entrevistador para localizar e discutir quaisquer problemas e responder novas perguntas que possam surgir. No caso dos *surveys* por entrevista cara a cara, pode ser uma boa idéia fazer os entrevistadores voltarem para relatar as duas ou três primeiras entrevistas e depois retornarem ao campo para completar o resto.

Apesar de normalmente ser uma boa idéia estabelecer um número específico de entrevistas de exercício para cada entrevistador, a regra geral é que todos os entrevistadores devem fazer tantas entrevistas de exercício quantas necessárias para demonstrar que compreenderam completamente a tarefa ou que provavelmente não serão capazes de compreendê-la. Assim, alguns entrevistadores poderão ter que fazer mais entrevistas de exercício do que outros. Esta situação é bem aceitável, mesmo se as entrevistas de exercício subseqüentes tiverem que se superpor ao início das entrevistas verdadeiras. É melhor alguns entrevistadores comecem mais tarde do que cometerem erros durante as entrevistas da amostra final. Erros podem implicar ter que retirar os respondentes entrevistados erroneamente, reentrevistá-los, ou ajustar o desenho da amostra.

A Operação de Entrevistar

A operação de entrevistar deve ser organizada e controlada tão cuidadosamente como a sessão de treinamento. Apesar de os detalhes da operação variarem de *survey* para *survey*, podemos comentar alguns pontos gerais.

Equipe

O número de entrevistadores requeridos é determinado com base no (1) número de entrevistas a serem feitas, (2) tempo médio necessário para cada entrevista, (3) tempo programado para toda a operação de entrevistas e (4) número de entrevistadores qualificados disponíveis.

Uma boa regra básica é determinar o número requerido com base nos pontos 1, 2 e 3 e então recrutar e treinar duas vezes este número de pessoas. Durante o treinamento, muitos candidatos desistirão voluntariamente; outros serão convidados a desistirem. Geralmente, é melhor começar a entrevistar com poucos bons entrevistadores do que com muitos ruins; estes custam mais dinheiro, usam mais tempo do supervisor e produzem dados de pior qualidade. Após retirar os entrevistadores ruins, você deve iniciar as entrevistas verdadeiras com mais entrevistadores do que tinha antecipado que precisaria. Entrevistadores continuarão a desistir durante a pesquisa, e outros serão convidados a deixar a equipe. Quase sempre, a equipe de entrevistadores chega abalada ao final da operação; se pede aos melhores para ficarem um pouco mais, para o encerramento.

Alguns *surveys* precisam de mais de um supervisor. Apesar de uma pessoa dever ser responsável pela supervisão de toda a operação, ela pode receber assistência de uma equipe de supervisores. Como regra básica geral, basta um supervisor para cada dez entrevistadores. Num grande *survey* por entrevistas com vários supervisores, o responsável por tudo não deve supervisionar os entrevistadores. Ele tem trabalho bastante administrando a logística da operação, coordenando com o diretor do projeto, possivelmente recrutando e treinando mais entrevistadores e, é claro, supervisionando os supervisores.

A Entrevista Verdadeira

Num *survey* domiciliar cara a cara, os entrevistadores recebem os questionários e uma lista dos endereços onde devem

fazer as entrevistas. Eles são responsáveis por encontrar o endereço, bater na porta e se apresentarem. Se a entrevista for bem curta, pode ser melhor conduzi-la na porta da residência, mas qualquer entrevista que exija mais do que alguns minutos é mais bem-feita no domicílio do entrevistado. Além de tornar as coisas mais fáceis para o entrevistador, entrevistar na residência do respondente também proporciona mais privacidade.

Entrevistas por telefone são bem diferentes e variam de acordo com os sistemas disponíveis para cada *survey*. Uma entrevista por telefone pode ser bem semelhante a uma entrevista domiciliar, com a diferença que o entrevistador telefona para o entrevistado em vez de ir à sua casa. Esta era mais ou menos a situação em 1973, quando este livro foi escrito pela primeira vez.

Na atualidade, as entrevistas por telefone evoluíram muito, principalmente devido às possibilidades oferecidas pelos computadores. Conseqüentemente, o que era visto como conveniente, mas ruim em termos de qualidade, se tornou uma técnica muito usada e bem aceita. Portanto, falemos um pouco sobre o sistema ETAC – Entrevista por Telefone Assistida por Computador.

Imagine-se usando um fone de ouvido de telefonista, sentado em frente a um terminal de computador. Um computador central seleciona aleatoriamente um número de telefone e o disca automaticamente. (Seleção aleatória evita o problema de ligar para números que não constam no catálogo.) À medida que o número está sendo discado, sua apresentação aparece na tela do vídeo à sua frente (“Alô, meu nome é...”). Quando o respondente atende, você diz alô, apresenta a pesquisa, e faz a primeira pergunta que aparece na tela: “Você pode me dizer quantas pessoas moram neste endereço?”

Sempre que o entrevistado responde, você digita a resposta no teclado do terminal, seja a resposta literal a uma pergunta aberta ou a categoria de código correspondente a uma resposta a uma pergunta fechada. A resposta é imediatamente armazenada no computador central. A pergunta seguinte aparece na tela, você a faz e a resposta entra no computador e a entrevista continua.

Esta tecnologia não é ficção científica e está sendo cada vez mais usada por pesquisadores de *survey*, acadêmicos, governamentais ou comerciais. Boa parte do desenvolvimento desta técnica para *surveys* ocorreu no Departamento de Pesquisas

de *Survey* da Universidade da Califórnia, em Berkeley, algumas vezes em colaboração com o Departamento de Agricultura dos EUA e outros órgãos governamentais.

J. Merrill Shanks e Robert Tortora (1985, p.4) descreveram algumas das maneiras como eles integraram os computadores ao processo de *survey*. Além do seu uso no cenário descrito acima, o computador é uma ferramenta valiosa nos diferentes estágios do desenvolvimento de questionários: esboço, teste, revisão e formatação. Além disso, a logística do processo de entrevistas (por exemplo, treinamento, cronograma e supervisão dos entrevistadores) também pode ser gerenciada por computador.

Shanks e Tortora vão além do ETAC para falar mais geralmente sobre métodos de *survey* assistidos por computador. Muitas das técnicas discutidas com relação às entrevistas telefônicas podem ser aplicadas às entrevistas domiciliares cara a cara. O desenvolvimento dos computadores portáteis permite aos entrevistadores montarem um escritório na residência do entrevistado, lendo perguntas da tela e digitando as respostas como descrito acima. Além disso, algumas vezes é adequado usar computador para questionários auto-administrados. Em algumas situações, pode-se pedir aos entrevistados que se assentem em frente a um terminal e digitem suas respostas às perguntas que aparecem na tela.

Supervisão e Revisão

Deve-se estabelecer procedimentos para o relatório regular dos entrevistadores ao diretor da pesquisa ou aos supervisores, no caso de um estudo muito grande. Num *survey* por entrevista cara a cara, talvez o melhor seja fazer uma reunião regular, semanal, com cada entrevistador. No horário marcado, o entrevistador apresenta todas as entrevistas completas, repassa-as a você, e recebe outros endereços. O relatório é muito mais simples no caso de entrevistas por telefone, com todas as entrevistas feitas de um local central, mas se os entrevistadores usarem os telefones de suas casas, deve-se usar os procedimentos mais elaborados que acabamos de descrever.

Você pode iniciar perguntando e discutindo problemas ocorridos nas entrevistas. Repasse cada questionário completo, procurando respostas omitidas, erros aparentes, respostas ilegíveis e qualquer coisa mais que torne o questionário difícil de usar e interpretar. Procure também qualquer coisa que sugira

que o entrevistador não compreendeu a tarefa ou parte dela. Cada erro deve ser mostrado e discutido. O entrevistador deve sair da reunião sabendo precisamente como a situação devia ter sido resolvida para o mesmo erro não ocorrer novamente.

É bem razoável que o supervisor altere as designações e revisões para certos entrevistadores, baseado em seus desempenhos iniciais. Um entrevistador que estiver fazendo um bom trabalho pode receber mais entrevistas para fazer e você não tem que ser tão cuidadoso no controle do questionário completo, na presença dele. Ao mesmo tempo, não permita o desenvolvimento de um "efeito de halo" em relação aos entrevistadores. Mesmo se um entrevistador fizer um trabalho perfeito no início, continue a monitorar o trabalho dele.

Todos os questionários completos devem ser inteiramente revisados. Em alguns casos, esta tarefa pode ser atribuída a editores especiais, mas normalmente os supervisores devem fazer a maior parte da revisão, se não toda.

Redistribuindo Entrevistas

Os entrevistadores devem ser treinados quanto à importância de completar as entrevistas designadas, devendo estar preparados, em alguns casos, para persuadir entrevistados indecisos a participarem. Contudo, devem ser avisados para não exagerarem nesta tentativa.

Freqüentemente, um entrevistador telefona ou chega em momento inoportuno e é recusado pelo entrevistado, ou algo com relação à aparência ou postura do entrevistador pode levar a uma recusa inicial. O entrevistador não deve forçar o entrevistado a uma rejeição peremptória; pode ser que um telefonema do supervisor para o entrevistado, para discutir a pesquisa, ganhe sua cooperação.

Alguns entrevistadores são mais bem-sucedidos em obter cooperação do que outros. São capazes de estabelecer melhor e mais rapidamente a relação com os entrevistados. Tais pessoas são identificáveis no transcorrer da pesquisa e é aconselhável designá-las para as entrevistas difíceis. Se isto for feito, sua eficiência não deve ser julgada da mesma maneira que a dos outros entrevistadores, já que entrevistas difíceis levam mais tempo. Também pode ser correto pagá-lhes mais, baseado na dificuldade da tarefa e sua importância especial para o *survey*.

Verificação de Entrevistas

Todas ou parte das entrevistas devem ser verificadas pelo supervisor. Esta verificação pode assumir várias formas. Como checagem mínima, você deve telefonar para o respondente, identificar-se e verificar se a entrevista foi realmente feita. Um procedimento mais rigoroso de verificação é refazer algumas perguntas-chave e comparar as respostas com as relatadas no questionário. Mas essa reentrevista não deve ser muito extensa, porque gasta seu tempo, o do entrevistado (ele já deu tempo e informação voluntariamente), e pode deixar este preocupado com o sigilo do *survey*.

Há duas razões para verificar entrevistas. Alguns entrevistadores podem cometer erros honestos ao contatar os respondentes designados, especialmente com um desenho complexo de amostra domiciliar; também é possível discar o número errado. Outros entrevistadores simplesmente trapaçiam. Você não deve presumir que todos os entrevistadores irão religiosamente tentar contatar e entrevistar as pessoas designadas. Na verdade, é raro o *survey* sem trapaça. Alguns entrevistadores fazem apenas entrevistas rápidas e preenchem o questionário baseado nelas. Outros só entrevistam os respondentes mais convenientes. Outros completam questionários tomando uma cervejinha em casa.⁶

Sempre que descobrir trapaça, você deve acabar com ela, mas, ao mesmo tempo, deve tentar descobrir por que ela ocorreu. Você poderá concluir, se você for honesto consigo mesmo, que a trapaça se deveu em parte ao seu próprio manejo da operação de entrevistar. Você pode ter exigido demais dos entrevistadores, dado instruções ambíguas ou, simplesmente, dado a impressão de considerar o trabalho de campo abaixo do seu status. Tais descobertas podem não salvar o *survey* já em andamento, mas podem ajudá-lo em outros *surveys*.

Demitir Entrevistadores

Na prática, é muito difícil demitir entrevistadores. Ninguém quer fazê-lo, especialmente após trabalhar tão perto deles durante o treinamento e ver que eles parecem estar tentando fazer o melhor possível. Mesmo assim, alguns entrevistadores acabam se mostrando incapazes de desempenhar a tarefa. Mantê-los no trabalho compromete as finanças, o cronograma e a qualidade da pesquisa.

Sempre que alguém não estiver dando certo como entrevistador, lembre-se de que entrevistar exige habilidades especiais. A falta destas habilidades interpessoais não significa que a pessoa seja desqualificada para outras tarefas. Pode ser que uma pessoa seja um desastre total como entrevistadora e excelente em outra função no *survey*. O entrevistador e o projeto podem lucrar mais com uma transferência de setor do que com a simples demissão.

Resumo

Este capítulo tocou de leve em alguns dos importantes elementos da operação de entrevistas eficiente. A intenção dos comentários não foi treiná-lo como supervisor de entrevistas, foi informá-lo sobre o que está envolvido. Uma boa operação de entrevistas requer supervisores e entrevistadores cujos treinamento e experiência elaborem muito mais os pontos cobertos tão resumidamente aqui.

Ao mesmo tempo, o diretor do projeto deve ter uma função ativa no processo de entrevistas. Você precisa saber o que está envolvido para ajustar suas demandas às realidades das entrevistas no campo e estar numa posição melhor para avaliar o significado dos dados que receber. Como foi dito, sua presença durante o processo de entrevistas pode ter um importante efeito sobre a moral e a qualidade do trabalho a ser executado.

Notas

¹ Numa análise de taxas de resposta em 517 *surveys*, John Goyder concluiu que a diferença nas taxas de resposta entre os *surveys* por entrevistas e os feitos pelo correio decresceram em tempos recentes e que as diferenças anteriores eram primariamente função de um acompanhamento mais extenso no caso dos *surveys* por entrevistas. Veja GOYDER, John. Face-to-Face Interviews and Mailed Questionnaires: The Net Difference in Response Rate. *Public Opinion Quarterly*, v.49, p.234-252, verão 1985.

² ANDERSON, Barbara A., SILVER, Brian D., ABRAMSON, Paul R. The Effects of the Race of the Interviewer on Race-Related Attitudes of Black Respondents in SRC/CPS National Election Studies. *Public Opinion Quarterly*, v.52, p.289-324, outono 1988.

³ REESE, Stephen D., DANIELSON, Wayne A., SHOEMAKER, Pamela J., CHANG, Tsan-kuo, HSU, Huei-ling. Ethnicity-of-Interviewer Effects Among Mexican-Americans and Anglos. *Public Opinion Quarterly*, v.50, p.563-572, inverno 1986.

⁴ Anos atrás, me pediram para falar de um *survey* sobre qualidade de vida que eu desenhei para os idosos do Havaí. Perguntado como lidaríamos com o problema dos entrevistados que não falavam inglês, dei a "resposta de manual", explicando que teríamos uma pessoa para traduzir o questionário para a segunda língua e aí pedir a outra pessoa que o traduzisse de volta para o inglês. Assim, poderíamos comparar as duas versões em Inglês, fazer as correções necessárias na tradução e tê-lo administrado por um entrevistador que falasse a segunda língua.

Estava no auditório uma mulher que tinha feito muitas das entrevistas iniciais para nós, que ofereceu outra solução. "Se eu achar que o entrevistado não entende o inglês muito bem, eu apenas refaço a pergunta de maneira que ele possa entender." Enquanto eu aprofundava cada vez mais na cadeira, ela ilustrou: "Como esta questão que diz 'Em geral como você caracterizaria seu estado atual de saúde: Excelente, bom, médio ou ruim?' eu apenas digo, 'Ei, você doente?'".

⁵ O impacto do treinamento do entrevistador na qualidade dos dados coletados foi demonstrado por BILLIET, Jacques, LOOSVELDT, Geert. Improvement of the Quality of Responses to Factual *Survey* Questions by Interviewers. *Public Opinion Quarterly*, v.52, p.190-211, verão 1988.

⁶ Meu colega Gary Marx provavelmente estabeleceu o recorde nesta categoria. Num *survey* nacional sobre atitudes dos negros durante a década de 1960, ele reparou numa série de aproximadamente vinte questionários — todos do mesmo entrevistador — que relatou as atitudes antibrancos mais violentamente hostis de toda a equipe de entrevistadores. Antes de Gary investigar a questão, ficamos sabendo que o entrevistador e alguns amigos tinham sido presos numa conspiração para explodir o Monumento a George Washington e a Estátua da Liberdade.

Leituras Adicionais

- BRADBURN, Norman M., SUDMAN, Seymour e Associados. *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass, 1979.
- GORDEN, Raymond L. *Interviewing: Strategy, Techniques, and Tactics*. Homewood, IL: Dorsey Press, 1969.
- GROVES, Robert M., MATHIOWETZ, Nancy A. Computer Assisted Telephone Interviewing: Effects on Interviewers and Respondents. *Public Opinion Quarterly*, v.48, p.356-369, primavera 1984.
- KAHN, Robert L., CANNELL, Charles F. *The Dynamics of Interviewing*. New York: John Wiley & Sons, 1967.
- LAVRAKAS, Paul J. *Telephone Survey Methods: Sampling, Selection, and Supervision*. Beverly Hills, CA: Sage, 1987.

OKSENBERG, Lois, COLEMAN, Lerita, CANNELL, Charles F. Interviewers Voices and Refusal Rates in Telephone *Surveys*. *Public Opinion Quarterly*, v.50, p.97-111, primavera 1986.

RICHARDSON, Stephen A. et al. *Interviewing: Its Forms and Functions*. New York: Basic Books, 1965.

SHANKS, J. Merrill, TORTORA, Robert D. Beyond CATI: Generalized and Distributed Systems for Computer-Assisted *Surveys*, preparado para o Bureau of the Census, First Annual Research Conference, Reston, VA, mar. 20-23, 1985.

WEEKS, Michael F., KULKA, Richard A., PIERSON, Stephanie A. Optimal Call Scheduling for a Telephone *Survey*. *Public Opinion Quarterly*, v.51, p.540-549, inverno 1987.

Capítulo 11

Processamento dos Dados

O que o microscópio foi para a biologia e o telescópio para a astronomia, o computador tem sido para a moderna ciência social. Ainda hoje está se descobrindo a contribuição dos computadores para a pesquisa de *survey*. Este capítulo o introduz a esta contribuição.

O propósito é descrever métodos de conversão de dados de pesquisa de *survey* numa forma *legível-pela-máquina*, uma forma que possa ser lida e manipulada por computadores e máquinas semelhantes usadas na análise de dados quantitativos. Se você estiver conduzindo um projeto de pesquisa paralelamente à leitura deste livro, seus dados, até agora, estão na forma de questionários completos, folhas de transferência codificadas, ou coisas assim. Ao completar o estágio descrito neste capítulo, os dados estarão gravados em disquetes, cassetes, cartões perfurados ou outros dispositivos legíveis e analisáveis por uma máquina.

Dada a velocidade do desenvolvimento e disseminação dos computadores, é impossível prever o tipo de equipamento disponível para seu uso. Este capítulo dará uma visão geral dos vários estágios na evolução dos computadores na pesquisa de *survey*, para dizer algo sobre o equipamento com que você trabalha, qualquer que ele seja. Além disso, os primeiros equipamentos e técnicas algumas vezes revelam mais claramente a lógica da análise de dados do que os equipamentos e técnicas avançados de hoje, da mesma forma que você pode aprender mais facilmente os fundamentos do funcionamento de um

automóvel a partir de um Fusca do que do modelo mais recente e sofisticado de um Maserati.

Computadores na Pesquisa de Survey

A história da computação na pesquisa social começou em 1801, quase na mesma época em que germinavam as sementes da moderna pesquisa social. Nesse ano, o francês Joseph-Marie Jacquard fez uma revolução na indústria têxtil que teve um efeito nos mais improváveis recantos da vida.

Para facilitar o trançado de padrões intrincados, Jacquard inventou um tear automático que recebia instruções de cartões perfurados. À medida que uma série de cartões passava pelo “leitor” do tear, toquinhos de madeira passavam pelos buracos dos cartões e o tear traduzia aquela informação em padrões de trançado. Para criar novos padrões, Jacquard só precisava furar orifícios apropriados em novos cartões e o tear respondia de acordo.

O ponto a observar é que *informação* (neste caso, um padrão desejado de trançado) podia ser *codificada e armazenada* sob a forma de buracos num cartão e subsequente *recuperada* por uma máquina que lia os buracos e agia baseada no significado atribuído a eles. Esta é a lógica fundamental que veremos ser repetida de diferentes formas na discussão abaixo.

O evento seguinte da nossa história seletiva do computador ocorreu nos Estados Unidos, durante o censo de 1890. A Constituição americana exige um censo completo da população da nação a cada 10 anos, começando com a contagem de 1790 de quase 4 milhões de americanos. À medida que a população cresceu, também cresceu a tarefa de contá-la. O censo de 1880 contou mais de 62 milhões de pessoas, mas o Bureau do Censo levou *nove* anos para terminar as tabulações. Era preciso uma ruptura tecnológica antes do censo de 1890. O Bureau procurou sugestões.

Um ex-funcionário do Bureau, Herman Hollerith, que tinha trabalhado no censo de 1880, teve uma idéia. Como jovem professor de engenharia no M.I.T., propôs adaptar os cartões de Jacquard à tarefa de contar a população. À medida que as contagens locais eram feitas, eram transformadas em orifícios em cartões. Então, uma máquina de tabulação, criada por Hollerith, lia os cartões e fazia a contagem da população do país.

O sistema de Hollerith foi testado num concurso com outras propostas e verificou-se que era o mais rápido. Assim, o Bureau do Censo alugou por US\$750.000,00 o equipamento da nova Companhia de Máquinas de Tabulação de Hollerith e o total da população de 1890 foi conhecido em seis semanas, em contraste com os nove anos gastos para informar o censo de 1880. A Companhia de Hollerith continuou a desenvolver novos equipamentos, fundiu com outras empresas pioneiras e, eventualmente, mudou o nome para International Business Machines Corporation – IBM.

Armazenando Dados em Cartões

Durante a década de 1950, os cartões perfurados — chamados comumente de *cartões IBM* — foram adaptados à armazenagem e recuperação de dados de pesquisas de *survey*; ainda são usados algumas vezes para isso. Mesmo se você nunca tiver trabalhado com dados armazenados em cartões perfurados, a lógica desse método pode ajudá-lo a compreender como os tipos de dados discutidos neste livro são geralmente armazenados e recuperados nos dias de hoje.

Um cartão perfurado, como o que se vê na Figura 11-1, é dividido em oitenta colunas verticais, geralmente numeradas da esquerda para a direita. Cada coluna vertical é dividida em dez espaços, numerados 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 de cima para baixo. Você armazena dados nos cartões fazendo buracos nas colunas. Mais tarde, os computadores e outros equipamentos de processamento de dados recuperaram os dados, localizando e lendo colunas especificadas no cartão.

Uma máquina de perfuração faz buracos nos espaços das colunas dos cartões. Usando um teclado semelhante ao de uma máquina de escrever, o operador da perfuradora faz os buracos especificados (0, 1, 2) nas colunas especificadas do cartão.

Os dados são postos em formato legível-pela-máquina pela atribuição de uma ou mais colunas específicas de um cartão de dados (um *campo*) a uma variável, e pela atribuição de buracos dentro daquela coluna às várias categorias da variável. Por exemplo, o sexo de uma pessoa pode ser posto na coluna 5 do cartão. Se for mulher, perfura-se o 1 desta coluna; se for homem, perfura-se o 2. Repare que colunas correspondem a *variáveis*, e os diferentes orifícios numa coluna correspondem aos *atributos* que constituem a variável. A idade de uma pessoa pode ser designada nas colunas 6 e 7 (um código de duas colunas);

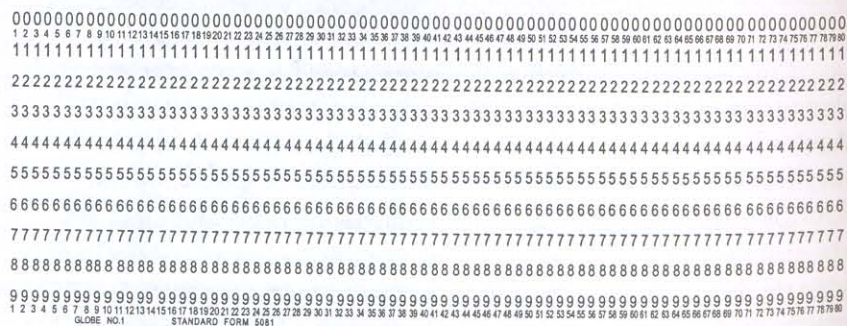
se ela tiver trinta e cinco anos de idade, seriam perfurados 3 e 5, respectivamente, nestas colunas. As idades podem também ser registradas em categorias e armazenadas numa única coluna, por exemplo, o orifício 1 para os menores de 20 anos, o 2 para os entre 20 e 29 etc.

Portanto, um cartão representa os dados relevantes a uma unidade de análise. Se as unidades de análise forem editoriais de jornais, numa análise de conteúdo, cada cartão representaria um editorial. As colunas de cada cartão seriam designadas a variáveis específicas descrevendo aquele editorial. Por exemplo, duas colunas podem ser designadas para armazenar os dois últimos dígitos do ano no qual o editorial foi publicado.

Na pesquisa de *survey*, um cartão de dados pode representar um questionário, com as colunas representando os vários itens do questionário. A coluna 34 pode ser usada para armazenar respostas à pergunta "Você já fumou maconha alguma vez?" O buraco 1 pode representar "sim" e o 2 o "não". As principais *idéias* a reter são: cada cartão representa uma única unidade de análise da pesquisa, e cada coluna (ou conjunto de colunas) é usada para a mesma variável em cada cartão. Portanto, *todas* as respostas dos entrevistados à pergunta da maconha seriam armazenadas na coluna 34.

FIGURA 11-1

Cartão perfurado padrão para registrar dados



Recuperando os dados a partir dos cartões

Vários equipamentos podem ler os cartões perfurados. A máquina mais simples entre estes é o contador-classificador [*counter-sorter*]. Embora seja improvável que você venha a usar

um contador-classificador, deve compreender sua lógica de funcionamento, pois isto esclarece o que fazem as máquinas mais complexas. O contador-classificador é preparado para ler uma determinada coluna. Quando os cartões passam por ele, são jogados em caixas correspondentes aos códigos perfurados na coluna especificada; um contador indica o número de cartões em cada caixa. Se o sexo for registrado na coluna 5 e o classificador tiver sido preparado para aquela coluna, mulheres e homens serão classificados nas caixas 1 e 2, respectivamente, e o contador registrará o total geral de cartões em cada caixa.

O contador-classificador pode ser usado também para examinar as relações entre variáveis. Suponhamos que você queira determinar se a probabilidade dos homens fumarem maconha é maior do que a das mulheres. Tendo separado os respondentes por sexo, você prepararia o contador-classificador para ler a coluna que tem as respostas sobre fumar maconha — por exemplo, coluna 34. Todos os cartões "homens" passariam novamente pelo contador-classificador para determinar se os respondentes já fumaram maconha, como indicado na coluna 34. O mesmo procedimento seria repetido para as mulheres, e as distribuições das respostas seriam comparadas.

Durante anos, o contador-classificador foi a principal máquina usada para a análise dos dados de *survey*, mas ele tinha três limitações básicas. Primeiro, era limitado a apenas contar e classificar os cartões. Embora se pudesse usar os resultados para análises extremamente sofisticadas, a própria máquina não fazia manipulações sofisticadas dos dados. Segundo, o contador-classificador era muito lento, quando comparado às máquinas inventadas depois. Terceiro, estava limitado a examinar um cartão por unidade de análise, ao observar as relações entre variáveis. Na verdade, estava limitado a oitenta colunas de dados por unidade de análise.

A Entrada dos Computadores

A maioria das análises de dados, hoje, é feita em computadores, indo desde grandes computadores centrais "*mainframe*" a pequenos computadores pessoais. É desnecessário dizer que os computadores estão revolucionando a maioria das áreas da vida social moderna. Ao mesmo tempo, algumas pessoas se preocupam com o fato de os computadores estarem "tomando conta". Espero que a discussão a seguir deixe claro que os computadores são apenas *ferramentas*,

assim como máquinas de escrever e calculadoras de bolso — mas ferramentas muito poderosas.

Através de programas de manipulação, o computador evita as limitações do contador-classificador. Primeiro, ele vai além da simples contagem e classificação, para executar complicadas computações e fornecer apresentações sofisticadas de resultados. O computador pode ser programado para examinar várias variáveis simultaneamente e computar uma grande variedade de estatísticas. Segundo, dados armazenados em discos ou fitas magnéticas podem ser lidos muito mais rapidamente do que era possível com os cartões e o contador-classificador. Computadores também podem calcular estatísticas complexas muito mais rápida e exatamente do que seres humanos. E finalmente, o computador não está limitado às oitenta colunas de um cartão perfurado.

Hoje, a maioria das análises dos dados sociais científicos quantitativos usam programas que instruem o computador a simular o processo de contagem-classificação descrito acima — e muito mais. Há vários programas para a análise de dados da ciência social. Alguns destes programas: ABtab, AIDA, A.STAT, BMDP, CRISP, DAISY, DATA-X, Dynacomp, INTERSTAT, MASS, MicroCase, Microquest, Microstat, MicroSURVEY, Minitab, POINT FIVE, P-STAT, SAM, SAS, SNAP, SPSSX, Statgraf, Statpac, StatPro, STATS PLUS, Statview, Survey Mate, SYSTAT, STAT80, STATA, SURVTAB, TECPACS. Em qualquer programa, a lógica básica do armazenamento e análise de dados será a mesma.

Até o fim dos anos 70, todas as análises computacionais eram feitas em computadores grandes e caros — às vezes chamados computadores *mainframe* — mantidos por departamentos centralizados de computadores; a maioria das análises ainda o é. Para usar tais instalações, você precisa fornecer os dados e as instruções para a análise desejada. Você leva os dados ao centro de computadores na forma de cartões ou em fita magnética, ou usa o banco de dados da biblioteca do centro.

No centro de computadores, você submete seus dados e instruções à análise do operador da máquina, que entra com a tarefa no computador (alguns centros são montados para o próprio pesquisador fazer o trabalho). Sua tarefa seguinte é esperar. Algum tempo depois — minutos (raramente), horas ou dias, dependendo da carga de trabalho do centro — o operador lhe entrega o trabalho finalizado, em geral na forma de um *impresso* dos resultados da análise.

Recentemente, dois desenvolvimentos melhoraram este processo. Sistemas põem leitores de cartão e impressoras

próximos aos pesquisadores, que podem, então, submeter seus trabalhos e receber os resultados em locais distantes do centro de computação. Seu departamento pode ter este equipamento, que permite que pedidos e resultados possam ser enviados de e para o centro de computação através de cabos coaxiais.

O *compartilhamento de tempo* é outro progresso no uso de computadores para análises. Computadores sempre operaram compartilhando tempo, com vários usuários usando o mesmo computador. Mas, no início, as instalações de computação eram partilhadas de maneira seriada; o computador executava o seu trabalho, depois o meu, depois o de outra pessoa. Hoje, os computadores podem executar várias tarefas simultaneamente. Podem ler meu pedido, analisar o seu e imprimir o de outra pessoa ao mesmo tempo. Sua enorme velocidade permite ensanduichar pequenas operações (que podem demorar apenas milésimos de segundo) no meio de trabalhos maiores. Esta capacidade permitiu aos computadores lidar com pedidos de centenas de usuários simultaneamente, às vezes dando a impressão que cada usuário tem toda a atenção deles. Tais interações geralmente ocorrem nos *terminais de computador* — dispositivos tipo máquinas de escrever que imprimem em papel ou mostram as operações em monitores de vídeo TRC (tubos de raios catódicos).

O uso de terminais de computador em regime de compartilhamento de tempo permite várias operações. Programas de edição de texto permitem aos usuários escrever, editar e imprimir relatórios. Se o seu centro de computação tem um sistema de compartilhamento de tempo, você encontrará, provavelmente, vários *video games* gravados em algum lugar do sistema.

Discutiremos aqui somente o compartilhamento de tempo na análise de dados da ciência social. Vejamos um padrão freqüente. Seus dados são mantidos pelo centro de computação, em disco ou em fita. Em vez de perfurar cartões para a análise, você digita instruções num terminal. Depois de preparar o conjunto de instruções, você digita uma instrução final (que varia de computador para computador), pedindo a execução do trabalho. Aí, o computador começa a executar, produzindo depois resultados para serem apanhados no centro ou impressos em seu terminal.

O compartilhamento de tempo de computador em pesquisas de *survey* ficou ainda mais prático com o desenvolvimento de terminais portáteis de computador, não maiores que pequenas máquinas de escrever. Estes terminais se comunicam com o

computador por linhas telefônicas convencionais. Portanto, você pode levar um terminal portátil para casa ou numa viagem. Para usá-lo, discar um número especial no seu telefone comum. O computador atende, geralmente com um ruído eletrônico. Você encaixa o fone em dois bocais de borracha, chamados *acopladores acústicos*, localizados no terminal. (Cada vez mais terminais portáteis têm equipamentos de comunicação — chamados *modems* — embutidos, de maneira que o terminal pode ser diretamente conectado a uma tomada de telefone padrão.) Uma vez conectado, o que você digitar no terminal é transmitido por linha telefônica ao computador, que faz o que você pede e envia a resposta ao seu terminal.

Os sistemas de compartilhamento de tempo também podem conectá-lo a *redes* nacionais de computadores. Os possíveis usos de tais redes são inimagináveis, e os serviços disponíveis atualmente somente arranham a superfície do seu potencial. Por exemplo, você pode discar um número especial e conectar-se a um computador que lhe permite procurar rapidamente, em edições passadas de um jornal, artigos sobre algum tópico de seu interesse. Se quiser, você pode ter os artigos impressos em seu terminal, ou pode deixar mensagens na rede para ser interceptadas por outros usuários.

Mais relevante à nossa discussão é que as redes de computadores permitem a um pesquisador sentado em sua casa conseguir uma cópia de um conjunto de dados num computador em outra cidade e analisar estes dados usando um programa de um computador em uma terceira cidade. Os resultados desta análise podem ser enviados a você e também ser armazenados para exame por um colega numa outra cidade. Esse procedimento ainda não é comum, mas é provável que venha a sê-lo em sua carreira de pesquisador; antes do fim dela, você se lembrará disso como sendo primitivo.

Microcomputadores

O progresso mais útil e empolgante até hoje é o *microcomputador*. Provavelmente você está familiarizado com microcomputadores e pode estar usando um. Essas máquinas são computadores completos, pequenos, não muito maiores do que uma máquina de escrever (ou até menores). As operações são exibidas num tubo de raios catódicos (TRC), que parece um aparelho de televisão, ou algum outro tipo de tela de vídeo. (Em alguns sistemas, você pode usar uma televisão convencional para servir como tela de vídeo.) Os primeiros sistemas de

microcomputador usavam fitas cassete convencionais para armazenar dados; elas foram substituídas por “discos flexíveis” de 5,25 polegadas ou “microcassetes” de 3,5 polegadas — lâminas redondas de fita magnética envoltas em papelão ou plástico. Os sistemas mais avançados usam “discos rígidos” que armazenam dados em milhões de bytes (*megabytes*). O microcomputador no qual digitei este texto tem um disco rígido embutido que armazena vinte megabytes.

Os microcomputadores já se provaram eficazes no desempenho de várias tarefas. O mesmo microcomputador usado para escrever este livro é usado para toda a correspondência; além disso, organiza minha agenda e imprime meus compromissos e listas de coisas a fazer, controla minha conta bancária, armazena meu diretório telefônico e desempenha tarefas numerosas demais para listar. Quando estou viajando, posso me registrar num hotel, tirar o telefone da tomada, ligar meu computador e me comunicar com pessoas e computadores em todo o país. Quando quero descansar, posso jogar xadrez com ele ou salvar o mundo de alienígenas invasores.

Em 1986, escrevi: “Os microcomputadores estão apenas começando a florescer na arena da pesquisa em ciências sociais. A maior limitação é sua *pequena* memória, mas os micros modernos estão superando este problema de forma impressionante.” Este problema já está solucionado, como mostra o seguinte exemplo.

Quando eu era aluno de graduação, há vinte e cinco anos, o Centro de Pesquisas de *Survey* de Berkeley tinha um computador IBM 1620, que ocupava aproximadamente o mesmo espaço de seis ou sete refrigeradores. Sua capacidade de memória era de 24 kilobytes (24K bytes), ou aproximadamente 24.000 caracteres de informação. Recentemente, comprei um computador/calculadora a bateria mais ou menos do tamanho de uma carta de baralho. Sua memória é de 32K bytes! O computador de mesa que usei para escrever este livro tem mais de 5 milhões de bytes (cinco *megabytes*) de memória (mais de 200 vezes a memória do velho IBM 1620), e seu disco rígido interno armazena 80 megabytes.

À medida que a capacidade de memória dos microcomputadores explodiu, programas de análise de dados foram desenvolvidos para analisar os dados da ciência social de acordo com a capacidade dos micros. A evolução tecnológica está ainda longe de um fim. Apresentando uma edição especial da *Sociological Methods & Research* sobre “Microcomputadores e a Pesquisa Social”, David R. Heise diz:

Microcomputadores são baratos, confiáveis, portáteis, computacionalmente poderosos e fáceis de usar. Este perfil os torna significativamente diferentes dos computadores *mainframe* e garante sua ampla difusão. Até o fim da década, os microcomputadores mudarão a maneira dos cientistas sociais fazerem pesquisa, a maneira pela qual lecionam e a maneira como fazem trabalho aplicado. Os microcomputadores também criarão novos tópicos para análise social à medida que a revolução dos microcomputadores alcançar diversos setores da sociedade.¹

O resto deste capítulo discute as etapas (e opções) envolvidas na conversão dos dados em formas aptas para análise por computador. Discutiremos o processo de codificação e enumeraremos os vários métodos para transformar os dados para forma legível-pela-máquina.

Codificação

Para os computadores fazerem seus milagres, precisam poder ler os dados coletados na pesquisa. Além disso, eles são melhores com números. Se os respondentes de um *survey* lhe disserem que pensam que o maior problema da sua região, hoje, é “a ameaça da guerra termonuclear”, o computador não compreenderá esta resposta. Você deve traduzir a resposta através de um processo chamado *codificação*. A discussão sobre análise de conteúdo do Capítulo 2 lidou com o processo de codificação de modo relevante a esta questão. Lembre que o analista de conteúdo precisa desenvolver métodos para classificar parágrafos, editoriais, livros, canções, portanto, classificação ou atributos específicos. Na análise de conteúdo, a codificação é inerente à coleta de dados ou à observação. Quando se usa outros métodos de pesquisa, freqüentemente é preciso um processo de codificação após os dados terem sido coletados. Por exemplo, os itens abertos do questionário resultam em respostas não-numéricas que precisam ser codificadas antes da análise.

Como na análise de conteúdo, a tarefa aqui é reduzir uma grande variedade de itens idiossincráticos de informação a um conjunto mais limitado de atributos compondo uma variável. Por exemplo, suponha que você perguntou aos entrevistados “Qual é sua ocupação?” As respostas variarão consideravelmente. Apesar de ser possível atribuir um código

numérico separado para cada ocupação respondida, tal procedimento não facilitaria a análise, que tipicamente depende de vários sujeitos terem o mesmo atributo.

A variável ocupação tem vários esquemas de código preestabelecidos (nenhum muito bom, porém). Um destes esquemas distingue entre ocupações profissionais e gerenciais, ocupações de escritório, ocupações semi-especializadas etc. Outro distingue diferentes setores da economia: manufatura, saúde, educação, comércio, e assim por diante. Outros combinam os dois esquemas.

O esquema de código ocupacional escolhido deve ser adequado aos conceitos teóricos que estão sendo estudados na pesquisa. Para algumas pesquisas, pode ser suficiente codificar todas as ocupações como de colarinho branco ou colarinho azul. Para outros, basta autônomo ou empregado. Um pesquisador da paz pode querer saber apenas se a ocupação depende do *establishment* militar.

Apesar do esquema de codificação dever ser ajustado para satisfazer requisitos particulares da análise, deve-se prestar atenção a uma regra geral. Se os dados forem codificados com alto grau de detalhamento, as categorias do código podem ser combinadas em qualquer análise que não necessitar deste grau de detalhe. Mas, se eles forem codificados em poucas categorias mais abrangentes, não será possível recriar os detalhes originais durante a análise. Assim, é aconselhável codificar os dados de modo mais detalhado do que você planeja usar na análise.

Desenvolvendo Categorias de Códigos

Há duas abordagens básicas do processo de codificação. Primeiro, você pode começar com um esquema de codificação relativamente bem-desenvolvido, baseado no seu objetivo de pesquisa. Portanto, o pesquisador da paz pode codificar as ocupações em termos de suas relações com o *establishment* militar. Ou, suponha que você está fazendo observação participante de uma nova religião emergente e tenha anotações cuidadosas sobre as razões da filiação dos novos membros. Talvez lhe pareça que os novos membros parecem considerar a religião como um substituto da família. Então, você pode rever as anotações cuidadosamente, codificando os comentários de cada novo membro quanto a isto ter sido, ou não, mencionado. Ou codificar esses comentários de acordo com o fato de eles terem, ou não, uma família.

Se você tiver ajuda durante a codificação, sua tarefa seria refinar suas definições de categorias de código e treinar os codificadores para estarem aptos a atribuir as categorias apropriadas às respostas dadas. Você deve explicar o significado das categorias de código que desenvolveu e dar vários exemplos de cada uma. Para assegurar que os codificadores compreendem completamente o que você pensa, codifique vários casos. Depois, peça aos codificadores para codificar os mesmos casos, sem dizer-lhes como você os codificou, comparando o trabalho deles ao seu. As discrepâncias indicarão comunicação imperfeita do seu esquema de codificação aos codificadores. Mesmo se esse exercício mostrar perfeito acordo entre você e eles, *cheque os códigos* de pelo menos uma parte dos casos durante a codificação.

Se não tiver ajuda na codificação, consiga alguma verificação de sua própria confiabilidade como codificador. Ninguém é perfeito, especialmente um pesquisador quente na pista de uma descoberta. Na pesquisa sobre a nova religião emergente, suponha que você tenha a impressão que as pessoas sem família convencional têm maior probabilidade de considerar a nova religião como um substituto da família. O perigo é que, sempre que descobrir uma pessoa que diz não ter família, você pode tentar, inconscientemente, encontrar evidência, nos comentários dela, de a religião funcionar como substituto da família. Então, se possível, tente achar alguém para codificar alguns dos seus casos, para ver se esta pessoa faria as mesmas atribuições que você. (Veja como esta situação se relaciona com a característica da *intersubjetividade* na ciência.)

A segunda abordagem à codificação deve ser usada sempre que não tiver certeza de como seus dados devem ser codificados, porque você não sabe quais variáveis os dados representam entre os sujeitos da pesquisa. Por exemplo, suponha a pergunta "O que você acha da idéia de uma moratória nuclear?" Apesar de haver antecipado a codificação das respostas como positiva, negativa ou neutra, é improvável que você possa antecipar toda a gama de variações de respostas. Nesses casos, prepare uma lista de 50 ou 100 respostas reais para a pergunta aberta e, depois, faça uma revisão na lista, notando as diferentes dimensões refletidas pelas respostas. Talvez você descubra que várias das respostas positivas contêm referências ao custo da corrida armamentista ou que várias respostas negativas se referam à ameaça do comunismo em escala mundial.

Depois de desenvolver um esquema de codificação baseado na lista das 50 ou 100 respostas, você deve certificar-se

de que cada resposta listada se encaixa numa das categorias de código. Assim, você estará pronto para começar a codificar as demais respostas. Se tiver ajuda para a codificação, aqui se aplicam os comentários anteriores com relação a treinamento e checagem; se não tiver, aplicam-se os comentários sobre pedir a alguém para checar seu trabalho.

Tal como o conjunto de atributos que compõem uma variável e as categorias de respostas a uma pergunta fechada num questionário, as categorias de código devem ser exaustivas e mutuamente excludentes. Cada pedaço de informação codificado deve se encaixar *numa e somente numa* categoria. Os problemas surgem quando uma resposta parece se encaixar igualmente em mais de uma categoria de código e quando não se encaixa em nenhuma.

Construção do Livro de Códigos

O produto final do processo de codificação é a conversão dos itens de dados em códigos numéricos. Estes códigos representam os atributos que compõem uma variável, os quais, por sua vez, são localizações de cartões e colunas designadas num arquivo de dados. Um *livro de códigos* é um documento que descreve as localizações das variáveis e lista os vínculos entre os códigos e os atributos que compõem as variáveis. Um livro de códigos serve para duas funções essenciais. Primeira, ser o guia principal no processo de codificação. Segunda, ser o seu guia para localizar variáveis e interpretar os códigos no arquivo de dados, durante a análise. Se você decidir correlacionar duas variáveis na análise dos dados, o livro de códigos lhe mostrará onde encontrar as variáveis e o que os códigos representam.

A Figura 11-2 ilustra partes do livro de códigos criado em 1987 para um *survey* dos professores e do pessoal administrativo de uma universidade. O formato do livro de códigos reflete um programa, o MicroCase, usado para processar e analisar os dados do *survey*. Outros programas usam formatos similares.

Estes formatos se aproximam das colunas de cartões perfurados descritos antes e são usados para localizar os itens de dados. Podemos, por exemplo, determinar o sexo do respondente referindo à VAR. 5. O MicroCase, como outros programas similares, permite também dar um nome abreviado a cada

item de dado; assim, podemos usar POS ACAD, em vez de VAR. 2, ao acessar os dados das posições acadêmicas dos respondentes. Esta possibilidade é particularmente útil por permitir dar nomes que você lembrará mais tarde, ao analisar os dados.

FIGURA 11-2
Exemplo parcial de um livro de códigos

VAR. 1: Fun Atú

Qual é sua função atual na Universidade Chapman?

1. Administração
2. Professor permanente
3. Professor contratado, em vias de tornar-se permanente
4. Professor contratado temporariamente
5. Diretoria
6. Função mista
7. Outra

VAR. 2: Pos Acad

Qual a sua posição acadêmica?

1. Professor Assistente
2. Professor Associado
3. Professor
4. Professor Adjunto
5. Outro
6. Não aplicável

VAR. 3: Camp Acad

Qual é o seu campo acadêmico?

1. Negócios
2. Educação
3. Humanidades
4. Ciências do Movimento e Exercícios
5. Ciências Naturais
6. Ciências Sociais
7. Misto ou outros
8. Não aplicável

VAR. 4: Grau

Qual o mais alto grau acadêmico que você possui?

1. Bacharel
2. Mestre
3. Doutor
4. Outro

VAR. 5: Sexo

Qual o seu sexo?

1. Feminino
2. Masculino

VAR. 6: Não/Rumo

A universidade não tem rumo.

1. Concordo fortemente
2. Concordo
3. Discordo
4. Discordo fortemente
5. Não sei

VAR. 7: Para/Pesq

A universidade está se afastando do ensino e se movendo na direção da pesquisa.

1. Concordo fortemente
2. Concordo
3. Discordo
4. Discordo fortemente
5. Não sei

Todo livro de códigos deve conter a redação completa das perguntas feitas; como notamos, a redação das perguntas tem um forte impacto nas respostas dadas. Finalmente, o livro de códigos deve indicar os atributos de cada variável. Assim, na VAR. 7, vemos que foi pedido aos professores que “Concordassem fortemente”, “Concordassem”, “Discordassem”, “Discordassem fortemente” ou dissessem que não sabiam. Muitos programas, inclusive o MicroCase, permitem que você dê nomes abreviados aos atributos (por exemplo, CF para “Concordo fortemente”, C para “Concordo” etc.), que serão usados para mostrar os resultados da análise, tornando mais fácil a leitura das tabelas, entre outras vantagens.

Opções de Codificação e Inserção de Dados

Há não muito tempo, quase toda entrada de dados era feita por perfuração manual, com cartões perfurados analisados por equipamentos de unidade de gravação ou transmitidos para computadores para análises mais complexas. Nos anos recentes, houve grandes avanços na entrada de dados. Usando terminais de computadores ou microcomputadores, os dados são, em geral, digitados diretamente nos arquivos de dados armazenados nos discos dos computadores. Mas, como antes, esta entrada de dados está intimamente relacionada à codificação e vários métodos podem ser usados para efetivar esta ligação. Vejamos algumas delas.

Folhas de Transferência

O método tradicional de processamento de dados envolve a codificação dos dados e a transferência dos códigos atribuídos para *folhas de transferência* ou *folhas de código*. Tais folhas são tradicionalmente divididas em oitenta colunas, correspondentes às colunas dos cartões de dados, mas podem ser adaptadas a outras configurações. A Figura 11-3 ilustra a folha de transferência. Os codificadores escrevem os números correspondentes às categorias de código desejadas nas colunas apropriadas das folhas. Estas são, então, usadas para entrar os dados nos arquivos do computador. Esta técnica ainda é útil quando se está processando questionários particularmente complexos ou outros documentos-fonte de dados.

FIGURA 11-3

Uma folha parcial de transferência de códigos

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		4	2	5	1	1	2	8	2	2	2
0	0	3	6	6	1	2	3	1	1		1	3	6	6	1	3	9	2	2	3
0	0	4	5	3	4	4	2	2	1		1	4	1	0	3	6	0	4	2	1

Codificação nas Margens

A codificação nas margens elimina a necessidade de folhas de código. A margem externa de cada página do questionário ou outro documento-fonte de dados é deixada em branco ou é marcada com espaços correspondentes a nomes ou números de variáveis. Em vez de os códigos serem transferidos para uma folha separada, eles são escritos nos espaços apropriados nas margens. Então, os documentos-fonte codificados nas margens são usados para a entrada dos dados.

Inserção Direta de Dados

Se os questionários tiverem sido desenhados adequadamente, você pode entrar com os dados diretamente no computador sem usar folhas de código separadas ou codificação nas margens. O questionário pré-codificado indica as colunas e os valores das perguntas e respostas, e a entrada de dados pode ser direta. Este procedimento foi muito aprimorado por alguns dos novos programas de computador, como ilustrado na Figura 11-4. Este exemplo foi tirado do programa MicroCase, mas outros programas seguem procedimento similar. (Os números em negrito são aqueles que você teria que digitar; o resto é feito pelo computador.)

Repare que, na Figura 11-4, o computador pede a entrada de dados repetindo a pergunta completa, lembrando-o de todos os possíveis números de código digitáveis como respostas àquela pergunta. Na prática, você pode entrar um conjunto de dados mais rapidamente se referindo à informação na tela, de vez em quando, para assegurar-se de que não se perdeu no processo de entrada de dados. Quando se usa o método de entrada direta de dados, os documentos precisam ser *revistos* antes da entrada.

Um revisor deve ler inteiramente cada documento para confirmar que todas as perguntas foram respondidas (digite 99 ou outro código padrão para os dados faltantes) e esclarecer respostas obscuras.

Se a maior parte do documento permitir entrada direta (por exemplo, perguntas fechadas apresentadas de forma clara), você pode também codificar alguns itens abertos e ainda assim entrar com os dados diretamente. Neste caso, marque o código de uma determinada pergunta num local específico, próximo à pergunta, para facilitar a tarefa de entrada dos dados.

O formato do documento é muito importante para uma efetiva entrada direta de dados. Se a maioria das perguntas e categorias de respostas forem apresentadas no lado direito da página, mas um conjunto apresentado no lado esquerdo, quem estiver entrando os dados irá, freqüentemente, pular o conjunto desviante. (Nota: muitos respondentes cometerão o mesmo erro, portanto, um questionário cuidadosamente desenhado para a entrada de dados será também mais eficaz na coleta de dados.)

Para a análise de conteúdo e situações semelhantes em que a codificação ocorre durante a coleta de dados, faz sentido registrar os dados de forma compatível com a entrada direta. Uma forma pré-codificada pode ser melhor, ou, em alguns casos, os dados podem ser registrados diretamente em folhas de transferência.

FIGURA 11-4

Ilustração da inserção de dados computadorizada

Caso 1

Qual é sua função atual na Chapman?

1. Admin
2. Permanente
3. Contr via Perm
4. Tempor
5. Diretoria
6. Mista
7. Outros

Var. 1. Fun Atu

2

Qual é sua posição acadêmica?

1. Prof. Assist.
2. Prof. Assoc.
3. Professor
4. Prof. Adj.
5. Outro
6. Não Apl.

Var. 2. Pos Acad

3

Entrada de Dados pelos Entrevistadores

A maneira mais fácil de entrar dados de *survey* — entrevista por telefone assistida por computador, ou ETAC — foi discutida no Capítulo 10: entrevistadores com fones de ouvido sentam em frente a terminais de computador, que mostram as perguntas a fazer, e entram as respostas assim que são dadas. Assim, os dados entram diretamente nos arquivos assim que são gerados. Os dados de perguntas fechadas estão prontos para análise imediata. Pode-se entrar também os dados de questões abertas, depois de uma etapa extra.

Suponha que o questionário pergunte “Qual é o maior problema do país hoje?” O computador pede ao entrevistador para fazer a pergunta. Em vez de esperar a inserção de um código numérico, o computador permite ao entrevistador digitar o que o entrevistado disser, por exemplo “crime nas ruas, especialmente os crimes cometidos por...” Depois, os codificadores recuperam as respostas abertas uma a uma, atribuindo-lhes códigos numéricos como discutimos antes.

Codificação para Folhas de *Scanner* Óptico

Algumas vezes, a entrada de dados pode ser feita eficazmente usando um *scanner* óptico, uma máquina que lê marcas de lápis preto numa folha de códigos especial e cria arquivos de dados correspondentes a estas marcas. Os codificadores podem transferir dados codificados para essas folhas especiais preenchendo os espaços apropriados. As folhas são depois postas num *scanner* óptico e os arquivos de dados são criados automaticamente.

Apesar de o *scanner* óptico proporcionar grande exatidão e velocidade à entrada manual de dados, tem também desvantagens. Primeiro, alguns codificadores acham difícil transferir dados para as folhas especiais. Pode ser difícil localizar a coluna certa e, depois de encontrada, o codificador precisa procurar o espaço certo para preencher.

Segundo, o *scanner* óptico tem tolerâncias relativamente rígidas. A menos que as marcas pretas sejam suficientemente pretas, ele pode cometer erros. (Você não terá como saber se isso aconteceu antes de iniciar a análise.) Além disso, se as folhas óptico-sensíveis forem dobradas ou estragadas, o *scanner* pode recusar-se a lê-las.

Uso Direto das Folhas de *Scanner* Óptico

Às vezes, é possível usar folhas de *scanner* óptico de forma um pouco diferente, evitando talvez as dificuldades que apresentam para os codificadores. Pode-se instruir os respondentes a completarem seus questionários diretamente nessas folhas. Folhas padrão podem ser fornecidas junto com instruções sobre seu uso, ou folhas especiais ser preparadas para uma pesquisa. As perguntas podem ser apresentadas com as categorias de respostas e os entrevistados orientados para colorir de preto os espaços ao lado das respostas escolhidas. Se as folhas forem apresentadas adequadamente, o *scanner* óptico poderá ler e entrar as respostas diretamente. Este método pode ser ainda mais viável no registro das observações experimentais ou na compilação de dados numa análise de conteúdo.

Pré-Codificação para a Inserção de Dados

Ao preparar o formato de um questionário, é essencial prestar muita atenção ao método usado para processar os dados, principalmente se você planejar entrar os dados diretamente a partir do questionário. As seguintes sugestões tornarão esta tarefa mais fácil.

O questionário deve ser pré-codificado para facilitar a entrada dos dados (ou codificado para esta função). Os itens devem ser atribuídos a campos de dados, nomeados ou numerados previamente, e anotações no questionário devem indicar essas atribuições.

Além das anotações sobre os campos de dados, você pode querer indicar as atribuições numéricas para as várias respostas num campo determinado, por exemplo, 1 = Concordo fortemente, 2 = Concordo, e assim por diante. Sempre que um campo tiver três ou menos categorias de respostas, pode-se em geral omitir a pré-codificação. Quando houver mais de três categorias, é normalmente mais seguro exibir as marcas de código. A Figura 11-5 apresenta exemplos de pré-codificação apropriados à marcação direta.

Há várias coisas a comentar na Figura 11-5. Primeiro, as atribuições de campo estão entre parênteses; Var14 a Var17 foram atribuídas à aprovação/desaprovação de figuras políticas, e a Var18, à identificação do partido político. As atribuições de respostas para a questão 9 estão apresentadas acima das colunas de categorias de respostas. Na questão 10, que tem um só

conjunto de respostas, as atribuições estão apresentadas nos espaços ao lado para marcar as respostas.

Finalmente, veja que todas as respostas às duas perguntas estão do mesmo lado da página. Isto é importante para facilitar a entrada dos dados. Você pode entrar os dados das duas perguntas sem precisar mover os olhos para frente e para trás na página. Quando se muda a posição das respostas na página, você trabalha mais lentamente e há mais probabilidade de perder itens.

Sempre que usar pré-codificação em questionários auto-administrados, é bom mencioná-la em nota introdutória ao questionário. ("Os números entre parênteses ao lado das categorias de respostas devem ser ignorados; eles foram postos apenas para ajudar no processamento das suas respostas.") Em qualquer caso, a pré-codificação deve ser discreta, para não confundir o respondente. Se o questionário for tipografado, a pré-codificação deve ser feita em letras pequenas.

FIGURA 11-5

Ilustração da pré-codificação de um questionário

9. Ao lado de cada figura política listada abaixo, indique se você Aprova Fortemente (AF), Aprova (A), Desaprova (D), Desaprova Fortemente (DF) a filosofia política da pessoa em questão, ou marque (NS) se você Não Souber.

(Var14-Var17)	AF	A	D	DF	NS
	1	2	3	4	5
a. George Bush	[]	[]	[]	[]	[]
b. Michael Dukakis	[]	[]	[]	[]	[]
c. Jesse Jackson	[]	[]	[]	[]	[]
d. George Wallace	[]	[]	[]	[]	[]

10. Qual é seu partido político, se você tiver?

	(Var18)
Partido Democrata	1 []
Partido Republicano	2 []
Partido da Paz e Liberdade	3 []
Partido Verde	4 []
Partido Independente Americano	5 []
Outro	6 []
Nenhum	7 []

Mesmo se o questionário não for marcado diretamente, o formato deve levar em consideração os métodos de processamento que se usará. A maioria das sugestões feitas aqui facilitará a codificação manual. Se o questionário for lido por uma máquina de leitura óptica, verifique se o formato está compatível com a máquina.

Limpeza dos Dados

Independentemente do método de processamento de dados usado, agora você tem um conjunto de dados legíveis-pela-máquina que pretende representar a informação coletada na pesquisa. O passo seguinte é eliminar os erros — "limpar" os dados. Não importa como, ou quão cuidadosamente, os dados foram entrados, erros são inevitáveis. Dependendo do método de processamento dos dados, os erros podem resultar de codificação incorreta, leitura incorreta de códigos escritos, leitura incorreta das marcas pelo *scanner* etc. Devem ser feitos dois tipos de limpeza: *limpeza de possíveis-códigos* e *limpeza de contingências*.

Limpeza de Possíveis-Códigos

Qualquer variável tem um conjunto especificado de atributos legítimos, traduzidos num conjunto de códigos possíveis. Na variável *sexo*, pode haver três códigos: 1 para homem, 2 para mulher e 0 para sem resposta. Se um caso foi codificado como 7 na coluna de sexo, com certeza houve um erro.

A limpeza de possíveis-códigos pode ser feita de duas maneiras. Primeiro, como vimos, os programas de computador para entrada de dados podem procurar erros à medida que estão sendo entrados. Por exemplo, se você digitar um 7 para sexo no MicroCase, o computador soa um bip e recusa o código errado. Outros programas de computador testam códigos ilegítimos nos arquivos de dados que não foram checados durante a entrada.

Segundo, se não tiver esses tipos de programas, você pode limpar possíveis-códigos examinando a distribuição das respostas a cada item no conjunto de dados. Assim, se seu conjunto de dados tiver 350 pessoas codificadas 1 para sexo (mulheres), 400 codificadas 2 (homens) e uma codificada 7, você suspeitará que o 7 é um erro.

Sempre que descobrir erros, localize o questionário pertinente, determine o código certo e faça a correção.

Limpeza de Contingências

A limpeza de contingências é mais complicada. A estrutura lógica dos dados pode determinar restrições especiais às respostas de certos respondentes. Por exemplo, um questionário pode perguntar o número de filhos que uma mulher deu à luz. Todas as respondentes mulheres devem ter uma resposta codificada (ou um código especial se não responderem), e nenhum entrevistado homem deve ter resposta registrada para esta pergunta (ou um código especial indicando que a pergunta não se aplica). Se um respondente homem for codificado como tendo dado à luz três crianças, ou se cometeu um erro que deve ser corrigido, ou sua pesquisa ficará muito mais famosa do que você jamais sonhou.

Geralmente, a limpeza de contingências é feita através de programas de computador que requerem um conjunto bastante complicado de declarações se-então. Em alguns casos, os programas que permitem entrada direta de dados podem ser instruídos a checar automaticamente as contingências necessárias, tal como checam códigos legítimos.

Apesar da limpeza dos dados ser um passo essencial no processamento dos dados, ela pode ser evitada em certos casos. Você pode achar plausível excluir sem perigo os pouquíssimos erros que ocorrem em algum item se a exclusão destes casos não afetar significativamente os resultados, ou algumas respostas contingentes inadequadas podem ser ignoradas. Se alguns homens receberem o status da maternidade, você pode limitar sua análise desta variável às mulheres. Mas não use esses comentários como racionalizações para uma pesquisa malfeita. Dados "sujos" quase sempre produzem resultados enganosos na pesquisa.

Resumo

A discussão sobre as opções e requisitos do processamento dos dados enfatizou a necessidade absoluta de um planejamento adequado da pesquisa. Você não pode primeiro coletar os dados e depois começar a pensar como eles serão

processados. O método de processamento dos dados deve ser escolhido antes de desenhar o questionário.

Os pesquisadores costumam suspirar de alívio quando os questionários são completados e guardados em segurança no escritório da pesquisa. O perigo de não conseguir os dados foi evitado; agora tudo o que é preciso é pô-los em ordem para a análise. Esta sensação de bem-estar pode ser infundada. Mesmo que os dados tenham sido coletados no formato correto, pode ser impossível pô-los numa ordem que faça sentido. Você só pode ficar aliviado se pensou o processamento dos dados com cuidado e antecipação e se desenhou uma técnica de coleta de dados adequada.

Nota

¹ HEISE, David (Ed.). *Microcomputers and Social Research*. Beverly Hills, CA: Sage, 1981. v.9, n.4, p.395.

Leituras Adicionais

COZBY, Paul C. *Using Computers in the Behavioral Sciences*. Palo Alto, CA: Mayfield, 1984.

EDWARDS, Perry, BROADWELL, Bruce. *Data Processing*. 2.ed. Belmont, CA: Wadsworth, 1982.

HEISE, David (Ed.). *Microcomputers in Social Research*. Beverly Hills, CA: Sage, 1981.

Capítulo 12

Pré-Testes e Estudos-Piloto

Todo manual de pesquisa aconselha fazer algum tipo de teste do desenho da pesquisa antes da pesquisa maior. Os argumentos a favor do pré-teste são convincentes. Ninguém quer investir grandes quantias de dinheiro e muito trabalho num grande projeto e não alcançar os objetivos da pesquisa devido a algum erro imprevisto.

Neste capítulo, distinguiremos dois tipos de teste: pré-testes e estudos-piloto. *Pré-testes* se referem ao teste inicial de um ou mais aspectos do desenho do estudo, como o questionário, o desenho da amostra, um programa de computador para a análise etc. *Estudos-piloto* se referem a exames em miniatura de todo o desenho da pesquisa.

Fazendo os Pré-Testes

Os pré-testes são os testes iniciais de um ou mais aspectos do desenho da pesquisa. Geralmente, isso significa a administração do esboço do questionário a um grupo de sujeitos, mas o conceito do pré-teste é mais amplamente aplicável.

Fazendo o Pré-Teste do Desenho da Amostra

Freqüentemente, um desenho de amostra parece razoável no papel, mas revela-se desastroso na prática. Se o desenho pede a criação de uma moldura de amostragem, pode-se criar uma parte desta moldura no pré-teste. Por exemplo, num desenho

de amostra por conglomerado de áreas, você pode tentar atualizar as estimativas dos tamanhos dos quarteirões baseado em um ou mais resultados de censo antes de enviar sua equipe para atualizar todos os tamanhos dos quarteirões. Se o desenho pedir a listagem das residências, você pode enviar enumeradores ao campo para selecionar quarteirões e revelar dificuldades inesperadas.

Se for selecionar sua amostra a partir de uma lista, estude-a cuidadosamente, procurando problemas. Se a intenção for estratificar, você deve testar o procedimento de estratificação. A estratificação pode ser feita manualmente? Tente fazê-la. Precisaremos de um programa de computador? Teste.

Os mesmos passos devem ser seguidos na seleção da amostra. Se for usar um programa de computador, ele deve ser testado com dados parciais ou mesmo dados hipotéticos. Se for fazer seleção manual, um pré-teste com dados parciais ou mesmo hipotéticos apontará os problemas e talvez sugira que a tarefa é impossível ou extremamente difícil; a seleção por computador pode ser mais viável.

Portanto, o pré-teste do desenho de uma amostra pode indicar se este desenho é possível, permitir avaliar as dificuldades e estimar o tempo e custo necessários. Os pré-testes têm outra função. Você nunca saberá antecipadamente todas as decisões que precisará tomar durante o *survey*. Apesar de ter sido explícito na definição da população do *survey*, a arte da amostragem desvendará problemas escondidos de definição. Por exemplo, numa amostra de área, você entrevistará os transeuntes (por exemplo, turistas) assim como os residentes? Filhos adotivos serão considerados membros da família? Ao fazer uma pesquisa numa universidade, você descobre logo que “estudantes” e/ou “professores” são criaturas muito difíceis de definir. Um pré-teste certamente apontará tais problemas (pelo menos alguns deles), e você poderá tomar decisões com relação à definição logo no começo. Com um pré-teste, suas decisões podem ser consideradas mais cuidadosamente, e você pode assegurar-se de que elas serão seguidas consistentemente até o final do estudo.

Pré-Testando o Instrumento da Pesquisa

A finalidade de pré-testar o instrumento da pesquisa é óbvia, mas as opções metodológicas podem não ser tão claras. Os seguintes pontos explicam melhor.

1. Todo o instrumento ou só uma parte dele pode ser pré-testado. Talvez a preocupação básica seja a aplicabilidade de um conjunto de perguntas. Se for assim, você pode concentrar a atenção em vários pré-testes desta parte do questionário e em modificações dela. Embora isto seja legítimo, lembre que o contexto no qual as perguntas aparecem afeta sua recepção. Portanto, se vale a pena fazer testes iniciais de partes do questionário, é melhor acrescentar um ou mais pré-testes de todo o instrumento.

2. Preferencialmente, o instrumento deve ser pré-testado tal como será usado na pesquisa; questionários auto-administrados e cronogramas de entrevistas devem ser pré-testados apropriadamente. Não obstante, isto não proíbe o teste inicial do instrumento por um método diferente. Recomenda-se pré-testar um rascunho inicial de um questionário auto-administrado através de entrevistas. Isto permite identificar problemas, se, durante a entrevista, um entrevistador detectar confusão e sondar a natureza desta. Entretanto, o instrumento deve ser pré-testado, em última instância, tal como será usado.

3. Um formato de perguntas abertas pode ser usado de forma proveitosa durante o pré-teste, para determinar categorias de respostas apropriadas para o que mais tarde se tornará uma pergunta fechada. Pede-se aos respondentes darem suas próprias respostas a uma pergunta, que, em seguida, serão adicionadas pelo pesquisador para criar categorias de respostas. Mas é importante testar as respostas fechadas também. Portanto, deve-se fazer um novo pré-teste para descobrir possíveis problemas nas categorias padronizadas.

4. A seleção de sujeitos para os pré-testes dos instrumentos pode ser feita de modo flexível e variado. Amostragem controlada não é necessária neste ponto. (Estudos rápidos são uma questão completamente diferente.) Digo isso porque os pesquisadores, freqüentemente, omitem os pré-testes devido às dificuldades na amostragem. A única diretriz que recomendo para seleção dos sujeitos é que estes devem ser pessoas razoavelmente adequadas para

as perguntas consideradas. Idealmente, cada item individual deve ser pré-testado nos membros da equipe de pesquisadores, embora tal teste não seja suficiente. Se a pesquisa tiver como alvo uma população em particular, quaisquer membros desta população ou pessoas semelhantes a ela podem servir como sujeitos para o pré-teste. Em pré-testes mais rigorosos do instrumento de pesquisa, deve-se dar pouca atenção à representatividade estrita; ao contrário, deve-se tentar alcançar a mais ampla variação nos tipos de respondentes — incluindo os que podem representar uma pequena minoria da população. Numa pesquisa de atitudes políticas, deve-se pré-testar o instrumento em sujeitos tirados de todo o espectro da população. Isto significa que os extremos — direita e esquerda — serão sobre-representados no teste, para assegurar que o instrumento fará sentido e será útil na compreensão de todos os tipos de respondentes na população. (É claro que o instrumento também deverá ser testado entre os “moderados”.) O objetivo do pré-teste é melhorar o instrumento da pesquisa e não fornecer descrições da população.

5. Pré-testar permite comparar diferentes métodos para obter os dados desejados. Diferentes rascunhos do questionário podem ser testados simultaneamente, e revisões e retestes sucessivos devem ter o mesmo efeito.

6. Pode ser bom usar os mesmos sujeitos mais de uma vez no pré-teste do instrumento. A razão disto é que o perfil geral do sujeito assim derivado fornece uma base geral para avaliar as respostas a itens específicos. Por outro lado, o rascunho final deve ser pré-testado em novos sujeitos para levar em conta o inevitável processo de aprendizagem que ocorre entre os sujeitos anteriores.

Estas são algumas das opções e diretrizes relevantes à realização de pré-testes. Enfatizo que o pré-teste deve ser um processo cumulativo em múltiplos estágios; você não deve pensar em termos *do* pré-teste. Mais à frente, consideraremos a avaliação dos pré-testes do questionário.

Pré-Teste da Coleta de Dados

Muito já foi escrito sobre as técnicas de coleta de dados, e pesquisadores experientes provavelmente já estabeleceram seus próprios métodos básicos. Mesmo assim, cada pesquisa tem suas particularidades com relação à coleta de dados.

Se for usar questionário enviado pelo correio, teste, pessoalmente, os procedimentos de montagem e remessa dos questionários. Somente assim podem ser devidamente organizadas as várias etapas do processo. Carimbar selos, colar etiquetas, dobrar cartas, pôr nos envelopes pode ser mais difícil na prática do que na teoria.

Um pesquisador experiente provavelmente pode organizar o trabalho, preparar a área de trabalho e designar tarefas à equipe num ensaio mental. Mesmo assim, só um teste físico pode demonstrar que efetivamente questionários não cabem em envelopes, ou que envelopes de retorno não cabem nos envelopes de remessa etc. Se o número de identificação do questionário deve emparelhar com o número na etiqueta de endereçamento, só um ensaio físico apontará os riscos (e, espera-se, a prevenção) de misturar estes números.

Continuar esta lista de exemplos não faz sentido, porque a natureza dos problemas a serem descobertos varia muito com a natureza do desenho da pesquisa. Basta dizer que o pesquisador que não pré-testar a remessa dos questionários corre um risco considerável.

A necessidade de pré-testar os métodos de coleta de dados é ainda maior num *survey* por entrevistas. Como isso pertence à área de treinamento, teste e supervisão do entrevistador, a discussão deste pré-teste já foi apresentada no Capítulo 10.

Pré-Testando o Processamento de Dados

Os pesquisadores tendem a ficar muito tensos na fase de coleta de dados do *survey* (os dados chegarão?) e relativamente mais relaxados nos estágios posteriores. Sentem que, assim que têm os questionários em mãos, a pressão acaba e os problemas podem ser resolvidos mais tranquilamente. Isso é verdade em muitos casos, mas essa atitude pode ser seriamente disfuncional se impedir o pré-teste dos procedimentos de processamento de dados.

É claro que a natureza do pré-teste do processamento de dados depende da natureza do próprio processamento dos dados. Você deve pré-testar as operações de codificação e entrada de dados. Os resultados deste pré-teste têm importantes implicações no formato do instrumento da pesquisa. Frequentemente, umas poucas pequenas modificações no *layout* do questionário melhoram muito a eficiência da codificação e da entrada dos dados. Pré-testar mostra as necessidades específicas da pré-codificação, exigindo mais, ou indicando que menos pode bastar. A única maneira de saber é fazer os codificadores praticarem a codificação, com dados hipotéticos ou com questionários completos pré-testados. Mas todo o processo do treino da codificação será em vão se este treinamento for atrasado até a hora que os questionários finais estiverem chegando.

Estes comentários se aplicam se os questionários forem desenhados para entrada direta de dados. Além disso, um pré-teste apontará as necessidades relacionadas à revisão dos questionários antes da entrada dos dados. Se a entrada for feita a partir de folhas de transferência codificadas, o teste mostrará requisitos especiais de codificação.

É preciso enfatizar a necessidade de pré-teste se os dados forem entrados por escaneamento óptico. Um pré-teste mostrará as tolerâncias da máquina com relação a folhas danificadas e espaços mal marcados. Um grande *survey* transferiu os dados do questionário para folhas óptico-sensíveis sem perceber, até o fim da codificação, que os codificadores usaram lápis com o grau errado de sombreamento. A equipe do projeto perdeu duas semanas, em tempo integral, escurecendo as marcas iniciais.

Se houver questões abertas, deve-se dar atenção antecipada à organização da codificação. Um pré-teste sugerirá a ordem mais eficaz para a codificação das questões abertas e fechadas.

Também deve-se pré-testar o processamento pós entrada de dados, tais como limpeza e transferência de dados para fita. Embora possa haver programas de computador de uso geral para estas tarefas, somente um teste verdadeiro assegurará que os dados estão organizados adequadamente para o uso destes programas. Se a pesquisa exigir a construção de arquivos de dados em múltiplos níveis (por exemplo, arquivos do domicílio, arquivos da família e arquivos pessoais), pode ser necessária uma manipulação mais complexa dos dados; esta manipulação precisa ser pré-testada.

Finalmente, pré-testes do(s) estágio(s) do processamento de dados devem mostrar o caminho para uma inter-relação eficiente dos elementos separados. A catalogação de questionários completos, revisão inicial, codificação, entrada de dados, limpeza e quaisquer outros elementos envolvidos no processamento da pesquisa precisam ser coordenados. Isso significa que os produtos de um estágio devem ser insumos apropriados do próximo e que deve haver coordenação de decisões do tipo quem levará os questionários de uma sala para outra, onde os questionários serão guardados num certo momento etc.

Pré-Teste da Análise

Você pode achar esquisito falar de pré-teste da análise, especialmente considerando minha insistência na análise como atividade aberta e frequentemente heurística. Mas sugiro esta etapa seriamente. Embora inevitavelmente você faça coisas com os dados que inicialmente não pretendia fazer, precisa assegurar-se de que será capaz de fazer as coisas que *pretendia* fazer.

Apesar de dar mais espaço à análise relacionada aos estudos-piloto, devo mencionar dois tipos de pré-testes analíticos. Primeiro, você deve dar todos os passos, desde os dados brutos até o produto final, de construção de tabelas, de índices, de escalas, regressões, análise fatorial etc., que tenha em mente; verificar cada etapa, dos registros dos dados à apresentação escrita. A razão dessa verificação exaustiva é assegurar que você *pode* chegar lá a partir daqui. Frequentemente, a forma da coleta da informação não se presta à análise pretendida. Vejam alguns exemplos.

Um *survey* de população coleta dados sobre cada membro da residência. Os dados são padronizados, com um registro separado para cada membro. Entretanto, neste formato, pode ser impossível determinar facilmente o *número* de pessoas em cada residência. Mesmo sendo possível contar o número de registros de pessoas para cada uma, pode ser impossível usar esta informação como variável analítica. (*Solução*: peça aos codificadores que contem as pessoas e codifiquem a contagem de pessoas num arquivo de residência separado.) Igualmente, pode ser difícil relacionar os dados da residência (por exemplo, a idade do chefe da família) aos arquivos das pessoas individuais, a menos que se tomem medidas antecipadas para permiti-lo.

Em outro exemplo, suponha que no pré-teste da análise você descubra que quer apresentar idade como uma *média*, para permitir comparações com outros dados; se você perguntou a idade em intervalos, em vez de a idade exata, pode ter problemas.

O segundo tipo de pré-teste analítico tem a ver com o uso de *hardware* e *software* de manipulação de dados — testar o computador e os programas que você pretende usar na análise. Você pode descobrir que o formato do questionário precisa ser modificado para se adequar a seus propósitos analíticos; descobrir isso durante o pré-teste é melhor do que depois de ter coletado todos os dados.

Fazendo Estudos-Piloto

Como tentei mostrar, é importante pré-testar os vários aspectos do desenho e da análise da pesquisa. Idealmente, devemos fazer extensos pré-testes de cada aspecto. Além disso, estarmos sempre atentos para as implicações do pré-teste de um aspecto para com outros aspectos e procurar inter-relacionar todos eles. O melhor meio para assegurar inter-relações válidas é fazer um estudo-piloto — um exame miniaturizado de toda a pesquisa, da amostragem ao relatório, que deve diferir do *survey* final apenas quanto ao tamanho, estudando menos casos (e usando menos tempo).

Estudo-Piloto da Amostragem

Diferentemente do pré-teste, o estudo-piloto deve fazer uma amostra representativa da população alvo. Portanto, a amostra do estudo-piloto deve ser selecionada exatamente da mesma maneira programada para o *survey* final. Mas com uma exceção.

Como se pode querer evitar estudar os mesmos respondentes no estudo-piloto e no *survey* final, pode-se evitar isto selecionando as duas amostras ao mesmo tempo, o que pode ser feito de duas maneiras. Primeira, selecionando a amostra final — ou, pelo menos, as unidades primárias de amostragem (por exemplo, quarteirões de censo) — e, depois, do restante da população, selecionando a amostra do estudo-piloto. Segunda, selecionando uma amostra inicial com elementos suficientes para ambas as amostras e subamostrando desta lista para o estudo-piloto.

Se a sobreposição não for considerada um problema, a amostra do estudo-piloto deve ser selecionada da mesma maneira planejada para o *survey* final. Em qualquer dos casos, terá sido testado o desenho da amostra (ou extraída a amostra final).

Estudo-Piloto do Instrumento de Pesquisa

O estudo-piloto deve envolver a administração de um instrumento de pesquisa o mais idêntico possível ao planejado para o *survey* final. Mas, se for usado um procedimento de postagem muito elaborado, pode não ser viável produzir o número necessário para o estudo-piloto e depois um lote revisado para o *survey* final.

O questionário do estudo-piloto deve conter todas as questões planejadas, com a redação, o formato e a seqüência que o pré-teste indicou serem os melhores. O estudo-piloto não deve ser um veículo para testar novos itens ainda não pré-testados.

Dito isso, vejamos uma exceção. É razoável que o instrumento do estudo-piloto contenha mais questões do que o planejado para o *survey* final. Frequentemente, pré-testes não controlados são insuficientes para determinar qual, de vários métodos para obter os dados, será mais útil. Portanto, em alguns casos, pode-se incluir mais de um método no mesmo questionário (se eles não se mostrarem uma repetição grosseira), com a intenção de avaliá-los através do estudo-piloto.

Uma maneira melhor, apesar de mais complexa, para lidar com esta situação é usar diferentes versões do questionário em subamostras do estudo-piloto. Assim, os respondentes do estudo-piloto estarão reagindo ao que pode se tornar o instrumento do *survey* final, fornecendo um teste melhor.

Se um único instrumento do *survey*-piloto for usado para avaliar métodos alternativos de obtenção dos dados, é preciso tomar cuidado para não criar um instrumento grande demais. As lições aprendidas com a administração de um instrumento de estudo-piloto longo e repetitivo podem ser inadequadas para o instrumento final mais curto.

Coleta e Processamento de Dados do Estudo-Piloto

Como em outras fases do estudo-piloto, a coleta e o processamento dos dados devem ser um exame miniaturizado do

desenho do *survey* final. Se os instrumentos de pesquisa forem comparáveis, devem ser administrados exatamente como você planeja fazer no *survey* final. Onde eles diferirem, é preciso estar especialmente consciente das implicações destas diferenças, tentando aprender a partir da inferência em vez da experiência. Os questionários completos devem ser codificados, e os dados entrados, transferidos, limpos e analisados exatamente como planejado para a pesquisa final.

Análise do Estudo-Piloto

Pode parecer estranho, mas a pesquisa *organizada* deve começar com um rascunho do relatório escrito. Ele deve conter os argumentos lógicos da pesquisa, tabelas em branco ou hipotéticas e a amarração verbal que une tudo isso. O pesquisador organizado assegura assim familiaridade com qual informação é necessária, de quem ela é necessária, e a forma na qual é necessária. A análise do estudo-piloto, assim como a análise final, deve consistir em preencher os espaços vazios empíricos, anotar os acontecimentos inesperados e elaborar estes desenvolvimentos. A análise do estudo-piloto deve ser feita com todo o vigor e a imaginação que serão usados na análise final. Como a amostra do estudo-piloto é extraída com a intenção de representar a população alvo, os resultados da análise do estudo-piloto devem ser essencialmente os mesmos do *survey* final. Este deve ser uma replicação mais substancial do primeiro.

Entretanto, os desenhos de pesquisas nunca são perfeitos, nem o raciocínio lógico que os embasa. Portanto, a análise do estudo-piloto nunca realmente resulta exatamente no que é esperado. Mais freqüentemente, aponta erros no raciocínio e/ou no desenho. Os métodos para determinar estes erros são discutidos a seguir. Agora, quero enfatizar a importância de levar o *survey* piloto o mais longe possível, para revelar o maior número possível de erros, antes de comprometer os principais recursos no *survey* final.

Se faltarem dados necessários ao estudo-piloto, devem-se fazer as aproximações possíveis para permitir a maior elaboração possível na análise pretendida. Se, por exemplo, a análise sugerir que se deve levar em conta níveis de renda, e você não tiver coletado dados sobre renda, você deve usar escolaridade ou ocupação para fazer estimativas aproximadas. Se não tiver dados aproximados, construa tabelas hipotéticas, apresentando todos os possíveis

resultados verdadeiros para renda. Determine como cada resultado possível seria explicado e quais outros dados deveriam ser considerados na explicação. Se esta abordagem sugerir mais análises, faça-as com os dados do estudo-piloto; se necessário, faça aproximações ou crie novas tabelas hipotéticas e tente explicá-las.

É muito fácil perceber um problema e suas implicações imediatas e deixar para corrigir o problema no *survey* final. Este procedimento é perigoso. Para tirar proveito de um estudo-piloto, você deve considerar as implicações de segunda, terceira e quarta ordem do problema e suas possíveis soluções.

Avaliando Pré-Testes e Estudos-Piloto

Esta seção oferece algumas diretrizes e critérios para avaliar pré-testes e estudos-piloto. Apesar de não oferecer leis para resolver este assunto, tentarei ligar os objetivos da pesquisa aos critérios de avaliação. Limitarei a discussão aos usos da análise na avaliação do instrumento da pesquisa.

O cerne de um *survey* são a coleta, manipulação e compreensão dos dados. Acima de tudo, pré-testes e estudos-piloto visam assegurar a criação de dados úteis. Consideremos maneiras de reconhecer dados úteis.

Clareza da Pergunta

Para serem úteis, as perguntas devem fazer sentido para os respondentes, mesmo se as implicações mais evidentes das questões não forem tão evidentes para eles. Igualmente, as categorias de respostas (se fornecidas) devem fazer sentido, tanto em si mesmas quanto na relação com a pergunta, e umas com as outras. A seguir, mostramos alguns sinais de perigo que podem surgir no pré-teste ou no estudo-piloto.

Omissão de Resposta. Geralmente, todo respondente pula algumas perguntas, e toda pergunta é pulada por alguém; contudo, quando uma pergunta produz muitos "sem resposta", temos uma dica de problemas no desenho do *survey*.

Múltiplas Respostas. Mesmo quando se pede aos respondentes para selecionar só uma resposta de uma lista de alternativas, alguns persistirão em selecionar mais de uma. Se uma pergunta produzir muitas respostas múltiplas, você deve suspeitar que ou suas categorias de respostas não são mutuamente

excludentes, ou a pergunta está sendo mal compreendida. A solução deste problema varia com o tipo de múltiplas respostas. Se as mesmas duas categorias forem escolhidas juntas freqüentemente, talvez elas possam ser melhor distinguidas ou, alternativamente, combinadas. Se várias combinações de respostas estiverem sendo selecionadas, deve haver algum erro mais básico, e toda a pergunta deve ser reexaminada.

Respostas "Outras". Freqüentemente, é correto oferecer ao respondente a alternativa de dar sua própria resposta em perguntas fechadas. Seja oferecida ou não esta opção, ter um grande número de respostas escritas (ou oferecidas pelo entrevistado) indica que as categorias fornecidas não são suficientemente exaustivas. Se as respostas "outras" se ajustam conceitualmente a uma ou mais categorias lógicas, sem sobreporem-se às existentes, talvez devam ser acrescentadas à lista de alternativas.

É claro que não se deve acrescentar respostas à lista, se isto criar conflitos com os objetivos do estudo ou desta pergunta. Por exemplo, os pesquisadores em comunicação nos Estados Unidos algumas vezes pedem aos respondentes para indicar sua fonte primária de informação política, oferecendo opções de uma lista que inclui jornais, amigos, rádio, revistas e periódicos políticos. A televisão é propositadamente omitida da lista, supondo que a maioria a escolheria caso fosse incluída. Mesmo se respondentes voluntariamente responderem televisão, você pode não querer incluí-la na lista. Entretanto, se muitos responderem televisão, você pode decidir excluí-la especificamente na pergunta básica (por exemplo, "Fora a televisão, qual dos seguintes...?").

Respostas Qualificadas. Freqüentemente, os respondentes qualificam suas respostas, tanto ao escolher uma categoria fornecida quanto ao responder uma pergunta aberta. Solicitados a informar sua "renda anual", eles podem assinalar que estão informando a renda do ano passado ou a estimativa deste ano. Solicitados a informar a idade do pai, podem observar que estão relatando a idade do *padrasto* ou a idade do pai quando faleceu. Perguntados sobre suas filiações a partidos políticos, podem assinalar que eles marcaram o partido ao qual são filiados, mas que, em geral, votam independentemente desta filiação; por outro lado, podem assinalar exatamente o contrário: que não estão filiados a um partido mas geralmente votam no partido indicado.

Respostas qualificadas como estas apontam uma falta de clareza nas perguntas e/ou nas respostas fornecidas. Uma resposta como essa pode ser suficiente para exigir uma revisão da pergunta, se a falta de clareza puder resultar em compreensão errônea da pergunta. Várias respostas como essas exigem uma revisão.

Comentários Diretos. Freqüentemente, os respondentes apontam diretamente problemas de redação ou formato da pergunta (por exemplo, "Esta é uma péssima pergunta"). Embora você possa desconsiderar alguns desses comentários, deve desconfiar das perguntas que os produzam em quantidade superior à normal.

Formato do Questionário

Todos os sinais de perigo discutidos anteriormente podem apontar erros no formato assim como erros na redação. Por exemplo, respostas faltantes podem indicar que o "fluxo" do questionário é inadequado. Isto é especialmente verdade para perguntas contingentes, que freqüentemente são (e *devem ser*) postas no lado da página do questionário. Melhores instruções, caixas, setas e dispositivos similares devem resolver estes problemas.

Algumas vezes, os respondentes respondem uma série de perguntas sim/não simplesmente marcando alguns sins. Podemos perguntar a que tipos de organizações pertencem os respondentes e providenciar uma lista de diferentes tipos com "sim/não" ao lado de cada uma. Se os respondentes marcarem somente sim em alguns casos, o problema consiste em não sabermos se as organizações ao lado das quais eles não marcaram nada são aquelas às quais eles não pertencem, ou se simplesmente eles as pularam inadvertidamente ao percorrer a lista. Instrução mais explícita (por exemplo, "por favor, responda a *cada uma*") pode aliviar este problema. Se você estiver particularmente preocupado com a precisão das respostas numa situação dessas, pergunte sobre a filiação a cada organização em perguntas separadas, em vez de numa lista; entretanto, isto aumentará o tamanho do questionário.

Variância nas Respostas

Uma preocupação primária na avaliação dos testes de um instrumento de pesquisa deve ser a distribuição das respostas

provocadas por cada pergunta. As respostas estão distribuídas mais ou menos igualmente entre as várias categorias de respostas fornecidas ou a maioria selecionou as mesmas respostas? A consideração deste aspecto dos resultados dos pré-testes e estudos-piloto, apesar de essencial, depende dos objetivos do estudo e das questões específicas.

Se a pergunta for feita para medir a informação fatural que permita uma avaliação independente (por exemplo, pedir a estudantes que informem suas médias escolares, quando você conhece a distribuição da população), então as respostas aos testes-piloto dão uma dica da clareza da pergunta e/ou da honestidade dos respondentes (é claro, supondo que os dados do teste sejam tirados de uma amostra representativa).

Se uma pergunta for, tal como redigida, intrinsecamente interessante, você pode ficar satisfeito com qualquer distribuição de respostas que obtiver. Pode ser interessante e valioso saber, por exemplo, que todos na população aprovam uma lei proposta, e você pode querer replicar esta descoberta no *survey* final.

Entretanto, em pesquisa social, estamos mais interessados nas *relações entre as variáveis*. Queremos saber por que algumas pessoas são favoráveis a um produto e outras não. Queremos saber por que algumas pessoas são conservadoras e outras não. Se não houver variação nas respostas dadas a uma pergunta, não seremos capazes de explicar as respostas. Não podemos explicar as diferenças que não aparecem na coleta de dados.

Procurando relações entre variáveis, a parte mais interessante da pesquisa social, é preciso obter variância nas respostas. Lembre-se das discussões anteriores deste livro, onde declaramos que os dados sociais são criados, em vez de coletados neutramente, assim como as discussões onde declaramos que conceitos sociais e variáveis não existem num sentido absoluto. É neste sentido que você precisa *criar* variância nas respostas fornecidas pelo instrumento de pesquisa. Se suas questões abordando religiosidade sugerem que todo mundo é religioso, é preciso ser mais rigoroso no critério de religiosidade, reescrevendo as perguntas de maneira a empurrar mais respondentes à apostasia. Dizemos isto sem sutilezas, porque é um ponto algumas vezes dificilmente aceito. Infelizmente, a literatura tradicional sobre o método científico criou a crença infundada que os pesquisadores devem conceituar as variáveis e as posições nessas variáveis que reflitam mais de perto a *realidade*. Portanto,

muitos novos pesquisadores gastam muito tempo tentando definir precisamente conceitos como “religiosidade”, para depois definirem o que constitui a “pessoa religiosa”.

A base epistemológica deste livro é a convicção de que conceitos como “religiosidade” e “religioso” não podem ser definidos de maneira final. O melhor que podemos esperar no presente é criar várias conceituações úteis de religiosidade, permitindo falar aproximadamente de pessoas *mais* ou *menos* religiosas. (O mesmo é verdade sobre conceitos como status social, liberalismo, alienação, conhecimento etc.) Conceituações são “úteis” na medida em que nos ajudam a entender os dados empíricos e facilitam a construção de teorias coerentes.

Para explicar por que algumas pessoas são mais religiosas que outras, é preciso definir e medir religiosidade de modo a permitir classificar alguns respondentes como mais religiosos e outros como menos religiosos, mesmo se este método de classificação não coincidir completamente com sua crença pessoal sobre o que é, definitivamente, uma “pessoa religiosa”.

Se você quiser explicar a variância, deve tentar maximizá-la entre os respondentes. Numa pergunta com respostas dicotômicas, tente obter uma distribuição meio a meio das respostas. Numa pergunta com mais alternativas de respostas, procure uma distribuição igual. Isso lhe dará maior potencial para a análise posterior. Então, ao avaliar os resultados do pré-teste ou do estudo-piloto, você pode querer manipular a variância produzida por certas perguntas. Vejamos algumas sugestões para obter a variância desejada.

1. A variância pode ser manipulada mudando a ênfase da pergunta. Por exemplo, suponha que queremos examinar a dedicação dos alunos ao aspecto de treino vocacional de uma educação superior nas artes liberais. Podemos pedir-lhes que concordem ou discordem da seguinte declaração: “Aprender uma habilidade vocacional é uma parte importante da minha educação superior.” Se quase todos os alunos num estudo-piloto concordarem, não seria possível analisarmos as respostas, devido à falta de variância. A variância pode ser manipulada “endurecendo” a declaração: “Aprender uma habilidade vocacional é a parte mais importante da minha educação superior.” Devemos esperar que menos estudantes concordem com a segunda declaração do que com a primeira.

Igualmente, se a declaração modificada for perguntada num estudo-piloto e pouquíssimos estudantes concordarem com ela, podemos considerar mudá-la para a forma mais fácil.

2. Alterar a ênfase das categorias de respostas pode ter o mesmo efeito. Suponha que se pergunte, num *survey* sobre saúde pública, “Com qual frequência você lava as mãos antes de comer?” e se dessem as seguintes respostas: “Sempre, de vez em quando, nunca.” Se poucos ou nenhum respondente num estudo-piloto disser “sempre”, esta resposta pode ser trocada por “quase sempre”.

3. Expandir a lista de categorias de respostas quase sempre aumenta a variância nas respostas. Aumentar as possibilidades de resposta de “concordo/discordo” para “concordo fortemente, concordo, discordo, discordo fortemente” quase sempre espalha as respostas. (No entanto, podemos acabar tendo que explicar a diferença entre “concordo fortemente” e “concordo”, por um lado, e “discordo fortemente”, por outro.)

De maneira similar, os respondentes algumas vezes relutam em escolher uma resposta extrema, especialmente em itens atitudinais ou de orientação. Solicitados a caracterizar suas orientações políticas como “muito liberal”, “moderadamente liberal”, “moderadamente conservador” ou “muito conservador”, relativamente poucos escolhem um dos extremos. Se os termos “esquerda radical” e “direita radical” forem acrescentados aos extremos da lista, esperamos que mais pessoas selecionarão “muito liberal” e “muito conservador”; pelo menos, mais respondentes selecionarão “esquerda radical” ou “muito liberal” do que escolheriam “muito liberal” na primeira lista.

Os exemplos mostrados envolvem a expansão das categorias de resposta em direção aos extremos da variável. Algumas vezes, é útil empurrar as respostas para apenas uma direção, ou seja, expandir apenas um extremo da variável. Assim, no exemplo, os pesquisadores da universidade poderiam acrescentar apenas “Nova Esquerda” ou “socialista” à lista inicial das categorias de respostas.

A variância nas respostas é um critério muito importante para avaliar o valor das perguntas. Entretanto, o modo de avaliar a variância depende do propósito da pesquisa ou da questão. Se o propósito for descrever, pode ser interessante e importante a pouca variância, quando a forma da pergunta for claramente adequada, ou você pode tomar a falta de variância como uma indicação de que a pergunta está mal trabalhada e que a descrição produzida é enganosa. Mas, se o propósito for explicar, é preciso ter variância. Seus respondentes devem parecer diferentes se você tiver que explicar suas diferenças.

Validação Interna de Itens

Até agora, discutimos apenas a avaliação independente de itens individuais. Os itens também podem ser avaliados examinando suas relações com outros itens. Suponha que você incluiu no estudo-piloto do questionário vários itens que acredita que meçam “alienação”. Uma função da avaliação do estudo-piloto é determinar se cada item serve a esse propósito. Se todos os itens de fato medirem a mesma disposição nos respondentes, as respostas a um item devem se correlacionar com as respostas a outros.

Tipicamente, todos ou a maioria dos itens concebidos para medir uma variável se relacionam empiricamente uns aos outros, mas a força das relações varia entre os pares de itens. Se um item for muito fracamente relacionado aos outros, você pode concluir que ele não mede de fato a variável e tirá-lo do questionário. (A possibilidade de ele ser o único item que mede a variável é um dos problemas com os quais os pesquisadores precisam aprender a viver.) No entanto, você não deve abandonar automaticamente todos os itens desse tipo. É possível decidir substantivamente que o item em questão representa uma dimensão um pouco diferente da variável, não relacionado empiricamente mas conceitualmente importante, ou decidir que a falta de relação é descritivamente interessante e desejar replicar as descobertas no *survey* final.

No extremo oposto, uma correlação muito alta entre dois itens pode sugerir que é desnecessário incluir os dois no questionário final. Eliminar um ou mais itens redundantes economiza espaço que pode ser usado para itens que não foram considerados no estudo-piloto do questionário.

A avaliação do estudo-piloto se baseia, em última instância, nos mesmos critérios de avaliação de um projeto de pesquisa encerrado — sua capacidade de contar uma história útil e interessante. Embora os comentários anteriores apontem alguns critérios avaliativos específicos, o critério destes é o sucesso no estudo-piloto da análise e do relatório. Como dissemos, você deve analisar e escrever os dados do seu estudo-piloto como se eles representassem o *survey* final. Basicamente, as seções do desenho do estudo-piloto que permitem construir uma história interessante e útil devem ser mantidas; as que impedem tal construção devem ser modificadas.

Os problemas de redação de perguntas, formato, variância, validade e relação entre itens são, em última instância, avaliados em termos das disfunções que trazem para o fio da história. Só uma análise e um relatório completos podem determinar isso. Você deve executar todas as etapas programadas da análise, construção de tabelas, índices, escalas, fatores, correlações, regressões etc., e explicar suas descobertas. Repito: este é o único método adequado para avaliar os resultados do estudo-piloto.

Analisando seriamente os dados do estudo-piloto, você tem condições de reconhecer que certos dados foram desconsiderados, num momento em que ainda é possível modificar o desenho, além de antecipar as surpresas analíticas que o esperam na análise do *survey* final. Se sua hipótese básica de pesquisa não for apoiada pelos dados do estudo-piloto, você deve analisar estes dados da maneira mais minuciosa possível, tentando descobrir por que a hipótese não foi apoiada. A medição das variáveis foi adequada? Outros fatores interferiram com a relação esperada entre as variáveis? A hipótese era sem fundamento? Apesar de não poder responder satisfatoriamente todas essas perguntas, a análise do estudo-piloto deve dar-lhe pistas sobre o que será examinado no *survey* final. Pelo menos, as descobertas do estudo-piloto devem dar-lhe tempo para moderar suas declarações sobre a hipótese durante o almoço no restaurante universitário. (Você pode também retirar sua oferta de apresentar um trabalho na reunião profissional anual.)

Terminemos o capítulo de modo realista. Embora os pré-testes e estudos-piloto descritos aqui o levem certamente a fazer uma pesquisa de *survey* mais profissional e mais valiosa, poucos, se é que algum, *surveys* passados podem ser dados como exemplos das regras e diretrizes discutidas. Você deve conhecer as razões de pesquisadores no passado não terem feito, em geral, uso adequado dos pré-testes e estudos-piloto e por que os pesquisadores futuros provavelmente também não o farão.

A boa pesquisa de *survey* é, quase sempre, muito onerosa e gasta muito tempo; pré-testes e estudos-piloto — especialmente bons estudos-piloto — parecem acrescentar mais tempo e custo. Como a pesquisa de *survey* (especialmente *surveys* por entrevistas) é muito trabalho-intensiva, o pesquisador de *survey* deve sempre lidar com o problema do gerenciamento de pessoal. Pode ser impossível fazer um bom estudo-piloto e manter seus entrevistadores, codificadores, digitadores e outros esperando vários meses ou anos enquanto você faz extensas análises de estudos-piloto. É bem provável que o orçamento e o tempo não permitam muitos pré-testes e um bom estudo-piloto. Também, questões, atitudes, molduras de amostragem etc. podem mudar bastante entre o estudo-piloto e o *survey* final atrasado. Estas são apenas algumas das razões por que pré-testes e estudos-piloto não são usados atualmente de forma adequada e provavelmente não o serão em todo seu potencial no futuro.

Ao mesmo tempo, estas restrições ainda permitem um uso melhor destes métodos do que é feito normalmente, e as seguintes sugestões devem ser consideradas. Primeiro, o pré-teste deve ser um processo contínuo, em múltiplos estágios, durante o desenho da pesquisa, como foi dito. Não pense em termos de fazer o pré-teste como um ritual de pesquisa, mas, em vez disso, use cada oportunidade para pré-testar cada aspecto do desenho do estudo sob quaisquer condições de teste disponíveis.

Segundo, se o tempo for insuficiente para uma análise detalhada dos resultados do estudo-piloto antes da revisão e impressão do questionário final, considere os seguintes passos. Assim que os questionários do estudo-piloto começarem a chegar, devem ser codificados e entrados. Não espere até que todos retornem. Começar imediatamente pode apontar dificuldades

no processamento que podem ser modificadas prontamente. Assim que um lote substancial de questionários, digamos uns cem ou mais, tiver sido processado, comece uma análise inicial. As distribuições marginais devem ser feitas imediatamente para permitir uma avaliação inicial dos itens individuais. Estas conclusões tentativas podem ser verificadas mais substancialmente à medida que novos lotes de questionários do estudo-piloto fiquem disponíveis.

Assim que um número suficiente de questionários for processado, deve-se iniciar uma análise explicativa mais detalhada; também se deve fazer isto com arquivos de dados parciais. Se não for possível fazer uma análise exaustiva antes dos prazos finais para revisar e imprimir o questionário definitivo, a análise do estudo-piloto não deve terminar quando o questionário revisto for mandado para a impressão. Se o desenho da amostra e o método de coleta de dados tiver sido organizado com antecedência suficiente, a impressão do questionário final em geral lhe dará um dos poucos "pontos mortos" no cronograma do projeto. Este tempo pode ser usado proveitosamente para uma análise mais detalhada do estudo-piloto, mesmo após você ter se comprometido com a forma do questionário final. Algumas vezes, você pode corrigir dados faltantes ou errôneos modificando os procedimentos de coleta dos dados sem mudar o próprio questionário. Além disso, esta análise detalhada lhe dará uma vantagem inicial na análise final.

Tendo notado que os procedimentos ideais delineados neste capítulo não podem ser seguidos religiosamente, devemos enfatizar o valor de se aderir a eles o máximo possível. Tentei sugerir diretrizes e procedimentos para fazer e avaliar pré-testes e estudos-piloto na pesquisa de *survey*. Como este tópico recebeu pouca ou nenhuma atenção na literatura prévia sobre pesquisa, o presente trabalho deve ser visto como fragmentário e tentativo. Este esforço inicial será útil até futuros escritores melhorarem esse quadro com relatórios de suas próprias experiências.

Análise da Pesquisa de *Survey*

4

CAPÍTULO 13
LÓGICA DA MEDIÇÃO E DA ASSOCIAÇÃO

CAPÍTULO 14
CONSTRUINDO E COMPREENDENDO TABELAS

CAPÍTULO 15
O MODELO DE ELABORAÇÃO

CAPÍTULO 16
ESTATÍSTICA SOCIAL

CAPÍTULO 17
TÉCNICAS MULTIVARIADAS AVANÇADAS

CAPÍTULO 18
O RELATÓRIO DA PESQUISA DE *SURVEY*

A Parte 4 trata de vários tópicos da análise dos dados de *survey*. Nos próximos capítulos, pegaremos a pesquisa no ponto onde você já tem seus arquivos de dados computadorizados, para discutir os passos a dar até a publicação das suas descobertas.

O Capítulo 13, que discute a lógica da medição e da associação, dá continuidade aos Capítulos 7 e 8. Nos concentramos, de novo, no significado das respostas a itens de questionário de *survey*, com respeito ao uso destas respostas como indicadores de variáveis como religiosidade, preconceito, alienação etc. Mas, desta vez, examinamos este tema do ponto de vista empírico, enquanto a discussão anterior se limitou às manipulações conceituais.

O Capítulo 14 aborda a lógica das tabelas de contingência. Alguns pesquisadores consideram rudimentar este formato analítico, mas lhe damos atenção especial, como base para compreender a lógica das análises de *survey*. A discussão começa com a análise univariada, depois passa para a análise bivariada e para análises multivariadas mais elaboradas. Neste processo, passamos da lógica da descrição para a lógica da explicação. A leitura deste capítulo deve lhe permitir construir tabelas de contingência a partir dos seus próprios dados de *survey* e compreender as tabelas construídas e publicadas por outros pesquisadores.

A chave da Parte 4 é o Capítulo 15, que discute o *modelo de elaboração*, onde examinamos a lógica básica da explicação científica no contexto do *survey*. Basicamente, buscamos compreender as relações empíricas entre duas variáveis através da introdução controlada de variáveis adicionais. Por exemplo, vendo que as mulheres de classe baixa são mais religiosas que as de classe alta, examinamos como a manipulação de variáveis adicionais pode esclarecer melhor o porquê disso.

Na maior parte, este livro trata mais da *lógica* da pesquisa do que das estatísticas envolvidas. Acredito que, entendendo bem a lógica da pesquisa, você está em condições de dominar as manipulações estatísticas relevantes. Nessa perspectiva, o Capítulo 16 dá uma visão geral das estatísticas mais comuns, não para oferecer uma explicação completa das computações destas estatísticas (vários excelentes textos de estatística fazem isso), mas para situá-las na lógica da análise, tal como discutida em todo o livro. A este respeito, é dada atenção especial a testes de significância estatística que parecem atalhos convenientes para a compreensão.

Como disse, examinamos primariamente a análise de *survey* através de tabelas de contingência, porque acredito que assim se vê mais claramente a lógica da análise. Ao mesmo tempo, outros modos de análise, mais complexos, têm mais potencial explicativo e são mais apropriados em muitas situações. Vários desses métodos são discutidos de maneira geral no Capítulo 17. Não tentei escrever um livro de receitas, mas espero ter conseguido situar esses métodos dentro da mesma lógica que informa a discussão das tabelas de contingência. Espero que você termine essa leitura com uma compreensão geral das restrições e potenciais destes métodos e possa aproveitar outros textos relevantes para empregá-los de modo apropriado.

O Capítulo 18 desenvolve um tópico que recebe relativamente pouca atenção — o relatório das descobertas do *survey*. Ele discute diretrizes gerais para a comunicação da pesquisa e considera os possíveis diferentes tipos de apresentações.

Capítulo 13

Lógica da Medição e da Associação

O cerne da análise do *survey* se reparte entre a descrição e a explicação. O analista de *survey* mede variáveis e depois examina as associações entre elas. Mas, como vimos no Capítulo 1, há muita confusão sobre a natureza deste processo.



Como já apresentamos os vários aspectos do desenho do *survey*, especialmente a conceituação e o desenho do instrumento, convém retornarmos à imagem tradicional do método científico. Já criticamos essa imagem, usando um exemplo de *survey*; examinemos agora, esquematicamente, a imagem tradicional e, depois, uma imagem alternativa, mais correta, da ciência na prática. A seguir, consideraremos a noção de Paul Lazarsfeld sobre a “interpermutabilidade de índices”.

A Imagem Tradicional

A perspectiva tradicional do método científico baseia-se numa seqüência de passos, resumidos na seguinte lista, que se acredita que os cientistas sigam em seu trabalho.

1. Construção da teoria
2. Derivação de hipóteses teóricas
3. Operacionalização de conceitos
4. Coleta de dados empíricos
5. Teste empírico de hipóteses

Antes de refutar esta perspectiva tradicional, vamos examiná-la melhor.

Elementos do Modelo Tradicional

Construção da Teoria. Face a um aspecto interessante do mundo natural ou social, o cientista supostamente deduz uma teoria abstrata para descrevê-lo. Supõe-se que este exercício é puramente lógico. Suponha que você esteja interessado em comportamento desviante. Você constrói, baseado na teoria sociológica existente, uma teoria do comportamento desviante. Entre outras coisas, essa teoria inclui uma variedade de conceitos relevantes às causas do comportamento desviante.

Derivação das Hipóteses Teóricas. Com base na sua teoria geral do comportamento desviante, você, supostamente, deriva hipóteses relacionando os vários conceitos que formam a teoria. Este é, também, um procedimento puramente lógico. No exemplo acima, suponha que você derive, logicamente, a hipótese de que a delinqüência juvenil é uma função da supervisão. Especificamente: à medida que aumenta a supervisão, diminui a delinqüência juvenil.

Operacionalização de Conceitos. O próximo passo na visão tradicional é a especificação de indicadores empíricos que representarão seus conceitos teóricos. Enquanto conceitos teóricos devem ser um tanto abstratos e talvez um pouco vagos, seus indicadores empíricos precisam ser específicos e precisos. Portanto, em nosso exemplo, você pode operacionalizar o conceito "juvenil" como pessoas com menos de dezoito anos de idade, "delinqüência" como ter sido preso por um ato criminoso, e "supervisão" como a presença de um adulto em casa, não trabalhando.

O efeito da operacionalização é converter a hipótese teórica em empírica. No caso, a hipótese empírica seria a seguinte: entre as pessoas com menos de dezoito anos de idade, as que moram em casa com um adulto que não trabalha têm menor probabilidade de serem presas por atos criminosos do que as sem um adulto não trabalhando em casa.

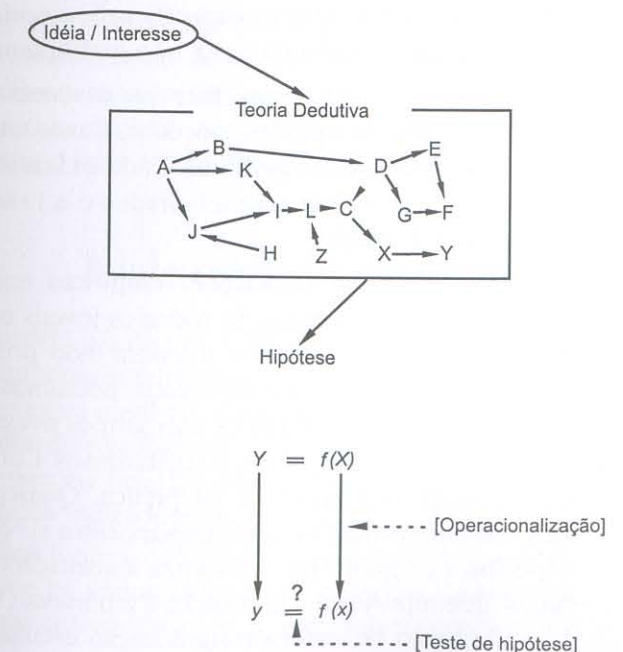
Coleta de Dados Empíricos. Baseado na operacionalização dos conceitos teóricos, você coleta dados relacionados aos indicadores empíricos. No nosso exemplo, você pode fazer um *survey* com pessoas com menos de dezoito anos de idade. Entre outras coisas, o questionário perguntaria a cada uma se ela mora em casa com um adulto que não trabalha e se já foi presa por cometer um ato criminoso.

Teste Empírico de Hipóteses. Após coletar os dados, o passo final é o teste estatístico da hipótese. Você determina empiricamente se os jovens com adultos que não trabalham em casa têm menor probabilidade de serem presos por atos criminosos do que os que não têm esses adultos. A confirmação ou desconfirmação da hipótese empírica é então usada para aceitar ou rejeitar a hipótese teórica.

Uma Apresentação Esquemática

A imagem tradicional da ciência é mostrada de modo esquemático na Figura 13-1. Você começa com um interesse particular sobre o mundo, cria uma teoria geral sobre ele e usa a teoria dedutiva para gerar uma hipótese sobre a associação entre duas variáveis. Esta hipótese é representada na forma $Y = f(X)$. Esta expressão é lida "Y é uma função de X", significando que os valores de Y são determinados ou causados por valores de X. No nosso exemplo, delinqüência (Y) é uma função de supervisão (X). A seguir, você operacionaliza as duas variáveis especificando medidas empíricas para representá-las no mundo real, coleta dados relevantes a tais medidas e, finalmente, testa empiricamente a relação esperada.

FIGURA 13-1
A imagem tradicional da ciência



Dois Problemas Básicos

Esta descrição da visão tradicional do método científico pode convencê-lo de que pesquisa científica é uma atividade rotineira. Tudo o que é preciso fazer é seguir religiosamente os passos 1 a 5 e se terá descoberto a verdade.

Infelizmente, a pesquisa científica e o mundo empírico que estudamos não são tão nítidos. A visão tradicional do método não representa o que realmente acontece na pesquisa científica — social ou outra. Dois problemas básicos impedem que o sonho se torne realidade.

Primeiro, conceitos teóricos raramente permitem operacionalizações não-ambíguas. Conceitos são abstratos e genéricos, enquanto a especificação de indicadores empíricos se faz por aproximação. No exemplo dado, é improvável que a variável “supervisão” seja representada adequadamente pela presença de um adulto que não trabalha em casa. A simples presença deste adulto não garante a supervisão de um jovem e, em algumas casas sem tal adulto, pode haver outros modos de fazê-la.

Igualmente, ser preso por um ato criminoso não é o mesmo que o conceito abstrato de “delinqüência”. Alguns jovens podem ter comportamento delinqüente sem serem presos. Outros podem ser presos erradamente. Além disso, a especificação de “juvenil” como uma pessoa abaixo de dezoito anos de idade é arbitrária. Outras especificações podem ser feitas e provavelmente nenhuma seria não-ambiguamente correta.

Além disso, não basta dizer que devíamos ter especificado “melhores” indicadores dos conceitos. Conceitos teóricos quase nunca têm indicadores perfeitos. Todo indicador empírico tem defeitos; todos podem ser melhorados e a busca de melhores indicadores é infinita.

Segundo, as associações empíricas entre as variáveis quase nunca são perfeitas. Se todos os jovens com adultos não trabalhando em casa nunca tivessem sido presos e todos os sem tal adulto tivessem sido presos, podíamos concluir que a hipótese foi confirmada. Se os dois grupos tivessem exatamente o mesmo número de presos, a rejeitaríamos. Contudo, nenhuma das alternativas é provável, na prática. Quase todas as variáveis são empiricamente relacionadas entre si “em algum grau”. Especificar o “grau” que representa a aceitação da hipótese e o “grau” que representa sua rejeição é arbitrário. (Veja no Capítulo 16 a discussão de testes de significação estatística.)

Na realidade, usamos indicadores aproximados dos conceitos teóricos para descobrir associações parciais. Estes problemas conspiram entre si contra você. Suponha que você especifique um grau de associação para a aceitação da hipótese e a análise empírica não o atinge. Naturalmente, você se perguntará se indicadores diferentes dos conceitos podiam ter resultado no grau especificado de associação.

Fiz estes comentários para indicar que a visão tradicional do método científico não é apropriada à pesquisa na prática. Não se faz pesquisa seguindo obstinadamente o caminho entre os passos 1 e 5. Isto não deve desanimá-lo, contudo, e sim desafiá-lo. Não deve ser considerado como negação da possibilidade da pesquisa científica, mas como base para a pesquisa esclarecida, realmente científica.

Medição e associação são conceitos inter-relacionados que devem ser trabalhados lógica e simultaneamente. Em vez de dar ordenadamente uma série de passos fixos, você deve dar estes passos para frente e para trás, incessantemente. Suas construções teóricas são, muitas vezes, feitas em torno de associações previamente observadas entre indicadores empíricos, construções teóricas parciais podem sugerir o exame de novos dados empíricos, e assim por diante. Espera-se que após cada atividade você entenda um pouco melhor o tema do seu interesse. É raro o “experimento crítico” que determina de uma vez por todas o destino de toda uma teoria.

Portanto, a pesquisa científica é uma empreitada sem fim, visando o entendimento de algum fenômeno. Para isso, você mede e examina associações, sempre consciente da natureza das suas inter-relações. As seções seguintes explicam melhor a natureza dessas inter-relações.

A Intercambialidade de Índices

Paul Lazarsfeld, ao discutir a “intercambiabilidade de índices”, criou uma importante ferramenta conceitual para a compreensão da relação entre medição e associação, assim como uma solução parcial dos dois problemas discutidos na seção anterior.¹ Essa discussão nasceu do reconhecimento de que qualquer conceito tem vários indicadores possíveis.

Retornemos à noção de uma hipótese teórica $Y = f(X)$. Lazarsfeld viu que há vários indicadores possíveis de supervisão, que podemos chamar de x_1, x_2, x_3 etc. Embora possa haver

razões para acreditar que alguns possíveis indicadores são melhores do que outros, eles são essencialmente intercambiáveis. Portanto, enfrentamos o dilema de quais indicadores usar no teste da hipótese $Y = f(X)$.

A solução para o dilema é usar todos os indicadores. Assim, testamos as hipóteses empíricas: $y = f(x_1)$, $y = f(x_2)$, $y = f(x_3)$ etc. Em vez de ter um teste para a hipótese, temos vários, como indicado esquematicamente na Figura 13-2.

Surge um novo dilema. Se o cientista que segue a visão tradicional do método enfrenta o problema da única associação empírica talvez não ser perfeita, o cientista que usa este método está diante de várias associações empíricas, nenhuma das quais perfeita, algumas delas podendo entrar em conflito com outras. Portanto, mesmo havendo especificado um grau de associação como suficiente para confirmar a hipótese, talvez você descubra que os testes usando x_1 , x_3 e x_5 satisfazem o critério especificado, mas os testes usando x_2 e x_4 não. Como resolver este dilema?

Nos termos da noção da "intercambiabilidade de índices", a hipótese teórica é aceita como uma proposição *geral* se for confirmada por todos os testes empíricos específicos. Se, por exemplo, a delinqüência juvenil for uma função da supervisão, no sentido geral deste conceito, ela deve estar relacionada empiricamente a todo indicador empírico de supervisão.

Se, no entanto, o cientista descobrir que somente certos indicadores da supervisão se relacionam à delinqüência juvenil, então os tipos de supervisão aos quais a proposição se aplica foram especificados. Na prática, isto pode ajudar a reconceituar "supervisão" em termos mais gerais. Talvez, por exemplo, a delinqüência juvenil seja uma função de restrições estruturais, e alguns tipos de supervisão sejam indicadores destas restrições e outros não.

É importante entender o que se alcança através deste processo. Em vez de testar rotineiramente uma hipótese fixa, relacionando supervisão a delinqüência, você ganha uma compreensão mais abrangente da natureza da associação. Mas este feito só é significativo se virmos o objetivo da ciência como compreensão crescente, em vez de somente construir teorias e testar hipóteses.

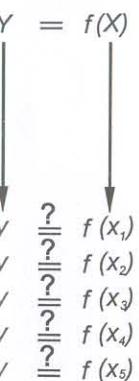
A implicação principal dos comentários anteriores é que a medição e a associação estão fortemente interligadas. A medição de uma variável faz pouco sentido fora dos contextos empíricos e teóricos das associações a serem testadas. Perguntado "Como devo medir classe social?", o cientista experiente responderá "Qual é seu propósito ao medi-la?" A maneira adequada de medir uma variável depende muito das variáveis que serão associadas a ela. Outro exemplo deve esclarecer isto.

Recentemente, surgiu uma controvérsia na sociologia religiosa sobre a relação entre religiosidade e preconceito. Um livro de Charles Y. Glock e Rodney Stark, *Christian Beliefs and Anti-Semitism*,² relatou dados empíricos indicando que membros de igrejas cristãs com crenças ortodoxas tinham maior probabilidade de serem anti-semitas do que membros menos ortodoxos. Estas descobertas acirraram consideráveis discussões nas igrejas e resultaram em outras pesquisas sobre o mesmo tópico por outros pesquisadores.

Um outro projeto de pesquisa chegou à conclusão oposta à de Glock e Stark, relatando que, quando a ortodoxia crescia, diminuía o preconceito. Mas análise mais minuciosa mostrou que as medidas de ortodoxia, nesta pesquisa, se basearam na aceitação de declarações de questionário que refletiam as doutrinas cristãs tradicionais de que "Todos os homens são irmãos" e "Ame o seu próximo". Não foi surpreendente que respondentes que aceitaram as declarações baseadas nessas doutrinas pareceram menos preconceituosos do que os que as rejeitaram. Normalmente, estas descobertas seriam (e foram) questionadas por causa de "contaminação". As duas variáveis examinadas (ortodoxia religiosa e tolerância) na verdade mediam qualidades idênticas ou similares. Chamar um conjunto de indicadores de "ortodoxia" e outro de "tolerância" não mostra que preconceito diminui com o aumento da ortodoxia em qualquer sentido geral.

As discussões deste capítulo sugerem uma reação um pouco diferente às descobertas da pesquisa. Perguntando *como* ortodoxia e preconceito estão associados entre si, em vez de perguntar se eles estão associados, concluímos que a ortodoxia medida pelos indicadores de Glock e Stark (crença em Deus, Jesus, milagres, e coisas assim) está positivamente associada

FIGURA 13-2
Intercambiabilidade
de índices



ao preconceito, enquanto a ortodoxia medida como compromisso com as normas de amor fraterno e igualdade está negativamente associada a preconceito. Ambas conclusões estão empiricamente corretas, apesar de nenhuma delas responder à pergunta geral, se religião e preconceito estão relacionados entre si.

É claro que o passo final é avaliar a utilidade relativa das conclusões. O achado de que ortodoxia e preconceito estão negativamente associados provavelmente seria descartado como tautológico ou trivial. (É claro que a medida da ortodoxia em termos de amor fraterno e de igualdade pode ser muito útil em algum outro contexto.)

Resumo

O objetivo deste capítulo foi prover uma perspectiva saudável sobre os objetivos da medição e da associação na ciência. Tentei proporcionar-lhe, a este respeito, melhor compreensão da empresa científica. O capítulo começou revendo o discurso tradicional de manual sobre o método científico, indicando que esta visão não reflete com exatidão a pesquisa científica na prática. No lugar da perspectiva tradicional, tentei oferecer um modelo alternativo, que acredito ser mais útil nas atividades de pesquisa reais.

Minha motivação para abordar essa questão vem da observação das dificuldades de pesquisadores inexperientes, que aceitam a perspectiva tradicional como um quadro verdadeiro de como cientistas "reais" trabalham, resultando em profundo desapontamento com a própria pesquisa. Não quero dizer que a pesquisa científica é descuidada ou "não-científica", apenas que ela não é rotineira. O cientista é bem diferente do técnico.

Notas

¹ LAZARFELD, Paul. Problems in Methodology. In: MERTON, Robert K. (Ed.). *Sociology Today*. New York: Basic Books, 1959. p.375.

² GLOCK, Charles Y., STARK, Rodney. *Christian Beliefs and Anti-Semitism*. New York: Harper & Row, 1967.

Leituras Adicionais

BABBIE, Earl. *Observing Ourselves: Essays in Social Research*. Belmont, CA: Wadsworth, 1986.

LAZARFELD, Paul. Problems in Methodology. In: MERTON, Robert K. (Ed.). *Sociology Today*. New York: Basic Books, 1959.

LAZARFELD, Paul F., PASANELLA, Ann K., ROSENBERG, Morris (Ed.). *Continuities in the Language of Social Research*. New York: Free Press, 1972. seção I.

Capítulo 14

Construindo e Compreendendo Tabelas

A maioria das análises de *survey* caem na rubrica geral da análise multivariada; a Parte 4 concentra-se nas variedades da análise multivariada. O termo se refere ao exame simultâneo de diversas variáveis. A análise das associações simultâneas entre idade, escolaridade e preconceito é um exemplo de análise multivariada.

A análise multivariada não é uma forma específica de análise; técnicas específicas para fazer análise multivariada incluem análise fatorial, análise do menor espaço, correlação múltipla, regressão múltipla e análise de trajetória [*path analysis*], entre outras. A lógica básica da análise multivariada pode ser mais bem-estudada através de tabelas simples, chamadas tabelas de contingência ou tabulações cruzadas. Portanto, este capítulo debruça-se sobre a construção e a compreensão dessas tabelas. Além disso, a análise multivariada não pode ser totalmente compreendida sem a compreensão firme de dois modos analíticos ainda mais fundamentais, análise univariada e análise bivariada. Portanto, o grosso deste capítulo discute estes modos analíticos.

Análise Univariada

Análise univariada é o exame da distribuição de casos de apenas uma variável de cada vez. Começaremos nossa discussão com a lógica e os formatos para a análise dos dados univariados.

Distribuições

O formato mais básico para apresentar dados univariados é relatar todos os casos individuais, isto é, listar o atributo de cada caso estudado na variável em questão. Suponha que você esteja interessado nas idades de executivos pesquisados num estudo sobre práticas empresariais. A maneira mais direta de relatar as idades dos executivos seria listá-las: 63, 57, 49, 62, 80, 72, 55 etc. Apesar desta lista dar ao leitor os dados mais detalhados, ela seria complicada demais para a maioria dos propósitos. Poderíamos distribuir os dados de maneira mais administrável sem perder qualquer detalhe, relatando que cinco executivos têm 38 anos de idade, sete têm 39, dezoito têm 40, e assim por diante. Tal formato impediria a duplicação de dados nesta variável.

Um formato ainda mais administrável, com alguma perda de detalhe, seria relatar as idades dos executivos como *marginais*, que são *distribuições de frequências* de dados agrupados: 246 executivos com menos de 45 anos, 517 entre 45 e 50 anos, e assim por diante. Neste caso, o leitor teria menos dados para examinar e interpretar mas não seria capaz de reproduzir completamente as idades originais dos executivos. Por exemplo, o leitor não teria como saber quantos deles tinham 41 anos.

O exemplo anterior apresentou os marginais na forma de números brutos. Uma forma alternativa seria o uso de *porcentagens*. Por exemplo, você pode relatar que x% dos executivos têm menos de 45 anos, y% estão entre 45 e 50 etc. (veja a Tabela 14-1). Frequentemente, você precisa decidir qual será a *base* a partir da qual computar a porcentagem, ou seja, o número que representa 100%. Nos exemplos mais diretos, a base é o número total de casos estudados. Entretanto, temos problemas sempre que faltarem dados para alguns casos. Por exemplo, suponha um *survey* no qual os respondentes devem informar a idade. Se alguns não responderem essa pergunta, há duas alternativas. Primeira, basear as porcentagens sobre o número total de respondentes, relatando os que não respondem como porcentagens do total. Segunda, usar o número de pessoas que responderam a pergunta como a base sobre a qual computar as porcentagens. Deve-se relatar o número dos que não responderam, mas estes não figuram nas porcentagens.

TABELA 14-1
Uma ilustração de análise univariada

Idades de Executivos de Empresas (Hipotéticas)	
Abaixo de 35	9%
36-45	21
46-55	45
56-65	19
66 ou mais	6
	100 % = (433)
Sem dados =	(18)

A escolha da base depende inteiramente do propósito da análise. Se este for comparar a distribuição etária da amostra com dados comparáveis da população da qual foi tirada a amostra, provavelmente deve-se omitir da computação os "sem resposta". A melhor estimativa da distribuição etária de todos os respondentes é a distribuição dos que responderam a pergunta. Como "sem resposta" não é uma categoria etária significativa, sua presença entre as categorias de base confunde a comparação dos números da amostra e da população. (Veja o exemplo da Tabela 14-1.)

Tendência Central

Além de informar os marginais, você pode apresentar os dados na forma de *médias* resumidas ou medidas de *tendência central*. Suas opções são a *moda* (o atributo mais freqüente, agrupado ou não agrupado), a *média aritmética* e a *mediana* (o atributo do meio na distribuição ordenada dos atributos observados). Vejamos como as três medidas são calculadas.

Suponha que você está fazendo um estudo-piloto que entreviste adolescentes, cujas idades, na sua amostra, variam de 13 a 19 anos.

Idade	Número
13	3
14	4
15	6
16	8
17	4
18	3
19	3

Tendo listado as idades dos 31 respondentes, qual você diria é a idade deles em geral, ou “na média”? Há três diferentes maneiras de responder esta pergunta.

A medida mais fácil de calcular é a *moda*, o valor mais freqüente. Como se vê na lista, há mais pessoas com 16 anos (oito) do que com qualquer outra idade, de forma que a idade modal é 16, como indicado na Figura 14-1.

A Figura 14-1 também mostra o cálculo da média, que envolve três passos: (1) Multiplique cada idade pelo número de entrevistados com esta idade, (2) some os resultados destas multiplicações e (3) divida o total pelo número de entrevistados. Como se vê, a idade média neste exemplo é 15,87.

A *mediana* representa o valor “do meio”; metade está acima dele e metade abaixo. Se tivéssemos a *exata* idade de cada entrevistado (por exemplo, 17 anos e 124 dias), poderíamos pôr os 31 entrevistados em ordem de idade, e a mediana para todo o grupo seria a idade do entrevistado do meio. Mas não sabemos as idades exatas; nossos dados são “dados agrupados”, no sentido em que três pessoas que não têm precisamente a mesma idade foram agrupadas na categoria “13 anos de idade”, por exemplo.

FIGURA 14-1
Três “Médias”

Idade	Número
13	☺☺☺
14	☺☺☺☺
15	☺☺☺☺☺☺
16	☺☺☺☺☺☺☺☺
17	☺☺☺☺
18	☺☺☺
19	☺☺☺

Modo = 16
Mais freqüente

Idade	Número
13	☺☺☺
14	☺☺☺☺
15	☺☺☺☺☺☺
16	☺☺☺☺☺☺☺☺
17	☺☺☺☺
18	☺☺☺
19	☺☺☺

13 x 3 = 39
14 x 4 = 56
15 x 6 = 90
16 x 8 = 128
17 x 4 = 68
18 x 3 = 54
19 x 3 = 54

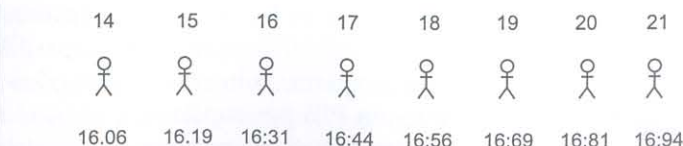
Média = 15.87
Média Aritmética

$$\frac{492}{31} = 15.87$$

(Total) (Casos)

Idade	Número
13	☺☺☺ 1-3
14	☺☺☺☺ 4-7
15	☺☺☺☺☺☺ 8-13
16	☺☺☺☺☺☺☺☺ 14-21
17	☺☺☺☺ 22-25
18	☺☺☺☺ 26-28
19	☺☺☺☺ 29-31

Mediana = 16.31
Ponto Médio



A Figura 14-1 exemplifica a lógica do cálculo da mediana para dados agrupados. Como há 31 entrevistados, o entrevistado “do meio” seria o número 16 se eles forem ordenados por idade; quinze seriam mais jovens e quinze mais velhos. Na parte de baixo da Figura 14-1, você vê que a pessoa do meio é uma das oito com 16 anos de idade. Na ampliação do grupo, vemos que o número 16 é o terceiro da esquerda para a direita.

Não sabemos as idades exatas dos entrevistados deste grupo; a convenção estatística em tais casos é presumir que eles estejam distribuídos igualmente ao longo da extensão do grupo. Neste caso, as *possíveis* idades dos entrevistados variam de 16 anos e zero dias a 16 anos e 364 dias. Falando de maneira mais rigorosa, a amplitude é de 364/365 anos. Na prática, basta chamar isso de um ano.

Se os oito entrevistados deste grupo estivessem igualmente distribuídos de um extremo ao outro, estariam com um oitavo de ano de diferença — um intervalo de 0,125-ano. O exemplo mostra que, se colocarmos o primeiro entrevistado na metade do intervalo a partir do limite inferior e acrescentarmos um intervalo completo à idade de cada entrevistado seguinte, o último entrevistado estará à metade de um intervalo de distância do limite superior.

Portanto, o que fizemos foi calcular hipoteticamente as idades exatas dos oito entrevistados, supondo que suas idades estejam distribuídas igualmente. Feito isso, anotamos a idade do entrevistado do meio (16,31) e a usamos como a idade mediana para o grupo.

Sempre que o número total de entrevistados for par, não haverá caso no meio, é claro. Nesta situação, você calcula a média dos dois valores que o cercam. Suponha, por exemplo, que houvesse mais um adolescente de 19 anos. O ponto do meio então cairia entre os números 16 e 17. A média, neste caso, seria calculada como $(16,31 + 16,44)/2 = 16,38$.

Na literatura de pesquisas, se mostram as médias e as medianas. Sempre que as médias forem mostradas, lembre que elas são susceptíveis ao efeito de valores extremos — poucos números muito grandes ou muito pequenos. Por exemplo, o Produto Interno Bruto (PIB) *per capita* de 1988 para os Estados Unidos foi de US\$17.500, em contraste com US\$400 para o Sri Lanka, US\$5.080 para a Irlanda, e US\$11.910 para a Austrália. Já o pequeno reino petrolífero dos Emirados Árabes Unidos teve um PIB *per capita* de US\$14.410, apesar de a maioria dos residentes daquele país viverem na pobreza, como se pode inferir

do fato de que sua taxa de mortalidade infantil foi quase quatro vezes maior que a dos Estados Unidos em 1988. O alto PIB *per capita* dos Emirados Árabes reflete a enorme petro-riqueza de poucos.¹

Dispersão

As medidas de tendência central oferecem a vantagem de reduzir os dados brutos a uma forma mais manejável. Um único número (ou atributo) pode representar todos os dados detalhados coletados sobre uma variável. Esta vantagem tem um custo, porque o leitor não poderá reconstruir os dados originais a partir da média, ou mediana. Esta desvantagem pode ser mais ou menos compensada pela informação sobre a *dispersão* das respostas. A medida mais simples da dispersão é a *amplitude* — a distância que separa o valor mais alto do mais baixo. Assim, além de informar que nossos entrevistados têm uma idade média de 15,87 anos, também podemos informar que suas idades variam de 13 a 19. Uma medida um pouco mais sofisticada de dispersão é o *desvio padrão*. A lógica dessa medida foi discutida no Capítulo 5, ao descrever o erro padrão da distribuição amostral.

Há várias outras medidas da dispersão, como por exemplo o *desvio quartílico* (ou *amplitude semi-interquartilica*). Um quartil representa um quarto dos casos estudados. Por exemplo, se estivermos estudando QI, o primeiro quartil seria composto pelos indivíduos com os QIs mais altos, o quarto quartil com aqueles com os QIs mais baixos. O desvio quartílico é a metade da distância entre o fim do primeiro quartil e o começo do quarto quartil. Se o quartil mais alto tiver QIs variando de 120 a 150, e o mais baixo tiver valores indo de 60 a 90, o desvio quartílico será $(120-90)/2 = 15$.

Veja que, já que os quatro quartis são de mesmo tamanho, o primeiro e segundo quartis têm metade das pessoas com os valores mais altos, e o terceiro e quarto, a metade das pessoas com os valores mais baixos. Assim, o ponto médio entre o segundo e o terceiro quartis é o valor mediano.

Variáveis Contínuas e Discretas

Os cálculos descritos acima não são apropriados para todas as variáveis. Para entender por que, precisamos distinguir dois tipos de variáveis: *contínuas* e *discretas*. Idade é uma variável contínua, de razão; ela aumenta firmemente em frações minúsculas em vez de pular de categoria para categoria como

acontece com variáveis discretas, como sexo ou posto militar. Se estivermos analisando variáveis discretas — uma variável nominal ou ordinal, por exemplo —, não se aplicam algumas das técnicas descritas acima. Falando estritamente, medianas e médias devem ser calculadas apenas para dados de intervalo e de razão, respectivamente. Se, por exemplo, a variável em questão for sexo, números brutos ou marginais de porcentagens são análises adequadas e úteis. Calcular a moda é uma análise legítima, apesar de não muito reveladora, mas médias, medianas e resumos de dispersão não seriam apropriados. Algumas vezes, os pesquisadores podem aprender algo de valor violando regras como estas, mas é preciso cautela ao fazê-lo.

Descrições de Subgrupos

Análises univariadas servem para *descrever* a amostra do *survey* e, por extensão, a população da qual foi extraída. Análises bivariadas e multivariadas objetivam primariamente temas explicativos. Mas, antes de entrarmos na explicação, consideremos o caso intermediário da descrição de subgrupos.

Freqüentemente, você quer descrever subconjuntos da sua amostra. Numa análise univariada direta, você pode querer apresentar a distribuição das respostas a uma questão relacionada a direitos iguais para homens e mulheres. Na exploração detalhada das respostas, faz sentido examinar separadamente as respostas de homens e mulheres da amostra. Ao examinar as atitudes com relação à Ku Klux Klan, faz sentido descrever separadamente os respondentes brancos e negros, assim como os respondentes de diferentes regiões do país.

Ao computar e apresentar descrições estratificadas, siga os mesmos passos descritos na seção sobre análise univariada; a diferença é que os passos são dados independentemente para cada subgrupo relevante. Por exemplo, todos os homens na amostra são tratados como uma amostra total, representando 100%, e as distribuições de respostas ou as medidas de tendência central são computadas para eles. O mesmo é feito para as mulheres. Assim, você poderia informar, por exemplo, que 75% das mulheres aprovam a igualdade entre os sexos e que 63% dos homens também aprovam. Cada grupo é submetido a uma análise univariada simples. Distribuições de frequência para subgrupos são geralmente referidas como marginais *estratificados*.

Em algumas situações, você apresenta os marginais estratificados ou outras análises de subgrupo apenas por razões descritivas. O relatório de dados do censo geralmente tem esse propósito. O valor médio de unidades de moradia em diferentes bairros pode ser apresentado por razões descritivas. Assim, você pode anotar o valor médio de uma casa para qualquer bairro.

Mas a razão mais freqüente da descrição de subgrupos é comparar. No caso da igualdade sexual, você estaria interessado em determinar se as mulheres têm *probabilidade maior* que os homens de aprovar a proposta. Além disso, na maioria dos casos, comparações não são motivadas por curiosidade ociosa. Tipicamente, a comparação se baseia na expectativa de que a variável de estratificação terá alguma forma de efeito causal sobre a variável de descrição. Ser homem ou mulher afetará a atitude com relação à igualdade entre os sexos. Igualmente, ser branco ou negro afetará atitudes com relação à Ku Klux Klan. Sempre que a análise for motivada por tais expectativas, entramos no domínio da explicação em vez da descrição.

Antes de entrarmos na lógica da análise causal, bivariada, vamos considerar outro exemplo de comparações de subgrupos que nos permitirá abordar alguns assuntos de formatação de tabelas.

Fundindo Categorias de Respostas

“Exemplos de manual” de tabelas são freqüentemente mais simples do que as que em geral encontramos publicadas em relatórios de pesquisas ou em nossas próprias análises de dados; esta seção e a próxima abordam dois problemas comuns e sugerem soluções.

Começamos voltando à Tabela 14-2, que mostra os dados coletados numa pesquisa multinacional de opinião pública feita pelo *New York Times*, *CBS News* e o *Herald Tribune* em 1985, sobre atitudes com relação às Nações Unidas. A pergunta da Tabela 14-2 trata de atitudes gerais sobre como as Nações Unidas davam conta de seu trabalho.

Comparemos como as pessoas das cinco nações na Tabela 14-2 apóiam o trabalho feito pela ONU. Ao inspecionarmos a tabela, descobrimos que há números demais para permitir ver qualquer padrão significativo.

Parte do problema é a relativamente pequena porcentagem de entrevistados que selecionaram as duas categorias de

respostas extremas, trabalho *muito bom* e trabalho *muito ruim*. Embora seja tentador ler apenas a segunda linha da tabela, que diz “bom trabalho”, seria incorreto fazê-lo. Observando apenas a segunda linha, concluiríamos que a Alemanha Ocidental e os Estados Unidos foram os mais positivos (46%) sobre o desempenho das Nações Unidas, seguidos de perto pela França (45%), com a Inglaterra menos positiva (39%) do que os outros três e o Japão (11%) o menos positivo de todos.

Este procedimento não é correto, porque ignora os respondentes que deram a resposta mais positiva, “muito bom trabalho”. Numa situação como essa, devemos combinar, ou fundir, os dois extremos da amplitude de variação. Isto é, combinar o “muito bom” com o “bom” e o “muito ruim” com o “ruim”. Se você fizer isso na análise dos seus próprios dados, é aconselhável somar as freqüências brutas e recomputar as porcentagens para as categorias combinadas; ao analisar uma tabela publicada como a Tabela 14-2, você pode simplesmente somar as porcentagens, como mostrado na Tabela 14-3.

TABELA 14-2

Atitude com relação às Nações Unidas: “Como a ONU está resolvendo os problemas que ela tem que enfrentar?”

	Alemanha Ocidental	Inglaterra	França	Japão	Estados Unidos
Muito bom	2 %	7 %	2 %	1 %	5 %
Bom	46	39	45	11	46
Ruim	21	28	22	43	27
Muito ruim	6	9	3	5	13
Não sei	26	17	28	41	10

FONTE: 5-Nation Survey Finds Hope for U.N. *New York Times*, p.6, 26 June 1985.

TABELA 14-3
Fundindo categorias extremas

	Alemanha Ocidental	Inglaterra	França	Japão	Estados Unidos
Trabalho bom ou melhor	48 %	46 %	47 %	12 %	51 %
Trabalho ruim ou pior	27	37	25	48	40
Não sei	26	17	28	41	10

As categorias fundidas na Tabela 14-3 nos permitem ler facilmente as porcentagens nacionais, mostrando os que dizem que a ONU está fazendo um trabalho pelo menos bom. Os Estados Unidos agora são os mais positivos; Alemanha, Inglaterra e França são pouco menos positivas e são quase indistinguíveis uma da outra; o Japão fica sozinho em sua avaliação bastante negativa. Apesar das conclusões a serem tiradas agora não diferirem radicalmente das que poderíamos ter tirado a partir da leitura da segunda linha da Tabela 14-2, vemos que a Inglaterra agora parece relativamente mais positiva.

Quero poupá-lo de um erro comum. Suponha que você tenha lido apressadamente a segunda coluna da Tabela 14-2 e notado que a Inglaterra tem uma avaliação um pouco pior do trabalho das ONU do que os Estados Unidos, a Alemanha e a França. Você pode se sentir obrigado a elaborar uma explicação para isso, talvez criando uma engenhosa teoria psico-histórica sobre o doloroso declínio da outrora digna e poderosa Inglaterra. Então, tendo alardeado sua “teoria”, alguém pode mostrar-lhe que uma leitura apropriada dos dados mostra que a Inglaterra não é menos positiva em sua avaliação do que as outras nações européias. Este *não é* um risco hipotético. Ele ocorre com freqüência, mas é possível evitá-lo fundindo categorias de respostas onde for o caso.

Lidando com os “Não Sei”

As Tabelas 14-2 e 14-3 ilustram outro problema comum na análise dos dados de *survey*. Geralmente, é bom dar às pessoas a opção de dizer “não sei” ou “sem opinião”, quando se pede suas opiniões sobre certos assuntos, mas o que fazer com estas respostas ao analisar os dados?

Repare a grande amplitude de variação nas porcentagens nacionais dizendo “não sei” na Tabela 14-3 — de apenas 10% nos Estados Unidos a 41% no Japão. Porcentagens substanciais respondendo “não sei” podem confundir os resultados de uma tabela. Por exemplo, será que os japoneses têm uma probabilidade tão menor de dizer que a ONU está fazendo um bom trabalho somente porque um grande número de japoneses não expressou sua opinião?

Há um modo fácil de calcular as porcentagens excluindo os “não sei”. Observe a primeira coluna de porcentagens na Tabela 14-3 — as respostas dos alemães ocidentais. Note que 26% dos respondentes disseram que não sabiam. Isso significa que os que disseram “trabalho bom” ou “ruim”, juntos, representam 74% (100 menos 26) do total. Se dividirmos os 48% que disseram “trabalho bom ou melhor” por 0,74 (a proporção que expressou uma opinião), podemos dizer que 63% “dos com opinião” disseram que as Nações Unidas estavam fazendo um trabalho bom ou muito bom ($48\%/0,74 = 63\%$). A Tabela 14-4 apresenta toda a tabela, excluídos os “não sei”.

Repare que estes novos dados oferecem uma interpretação um pouco diferente da das tabelas anteriores. Especificamente, agora parece que a França e a Alemanha Ocidental são os mais positivos nas avaliações das Nações Unidas, com os Estados Unidos e a Inglaterra um pouco abaixo. Apesar de o Japão continuar o mais negativo, ele passou de 12% positivos para 20%.

Neste ponto, tendo visto três versões dos dados, você pode estar se perguntando qual é a versão *certa*. A resposta depende do seu propósito na análise e interpretação dos dados. Por exemplo, se não for essencial para você distinguir entre “muito bom” e “bom”, faz sentido combiná-los porque assim será mais fácil ler a tabela.

É mais difícil decidir, no abstrato, se vamos excluir ou incluir os “não sei”. O fato de uma grande porcentagem de japoneses não ter opinião pode ser uma descoberta muito importante, se você quiser descobrir se as pessoas conhecem o trabalho das Nações Unidas. Por outro lado, se você quiser saber como elas votariam numa questão, é melhor excluir os “não sei”, baseado no suposto de que elas não votariam ou dividiriam seus votos entre os dois lados da questão.

De qualquer modo, a *verdade* contida nos dados é que uma certa porcentagem disse que não sabia e que os demais dividem suas opiniões na proporção relatada. Frequentemente, é certo informar seus dados em ambas as formas — com e sem os “não sei” —, para que os leitores possam tirar as próprias conclusões.

Agora vamos ajustar ligeiramente a lógica do nosso exame das comparações de subgrupos e virar nossa discussão na direção das análises bivariadas.

TABELA 14-4
Omitindo os “não sei”

	Alemanha Ocidental	Inglaterra	França	Japão	Estados Unidos
Trabalho bom ou melhor	63 %	55 %	65 %	20 %	57 %
Trabalho ruim ou pior	36	45	35	81	44

Análise Bivariada

A análise bivariada explicativa é, basicamente, a mesma coisa que descrição de subgrupos, com certas restrições especiais. Nas descrições de subgrupos, você tem total liberdade para escolher a variável de estratificação que quiser e descrever cada subgrupo nos termos de qualquer outra variável. No exemplo da igualdade sexual, você pode descrever homens e mulheres separadamente nos termos das porcentagens de homens e mulheres que aprovam ou desaprovam (veja Tabela 14-5), ou pode descrever separadamente os que aprovam e os que desaprovam em termos das porcentagens de homens e mulheres (veja Tabela 14-6).

Qualquer uma das tabelas é uma representação legítima de descrições de subgrupos. Entretanto, os dados apresentados nas duas tabelas são lidos diferentemente. Na Tabela 14-5, vemos que 63% dos homens da amostra aprovam a igualdade sexual, comparados a 75% das mulheres. Na Tabela 14-6, vemos que, dos que concordam, 46% são homens, enquanto, dos que discordam, 60% são homens. Ou podemos notar que 37% dos

400 homens discordam, comparados a 25% de 400 mulheres (Tabela 14-5), e, dos que concordam, 54% são mulheres, enquanto, dos que discordam, 40% são mulheres (Tabela 14-6).

TABELA 14-5

"Você concorda ou discorda da proposição de que homens e mulheres devem ser tratados igualmente em todos os aspectos?"

	Homens	Mulheres
Concordam	63%	75%
Discordam	37	25
	100%	100%
	(400)*	(400)

* Os números mostrados em parênteses representam a base para a porcentagem. Neste exemplo, há 400 homens ao todo, 63% (252 homens) dos quais "concordam". 37% (148 homens) "discordam".

TABELA 14-6

Direção oposta da porcentagem

	Concordam	Discordam
Homens	46%	60%
Mulheres	54	40
	100%	100%
	100% =(552)	(248)

Mas, numa análise bivariada explicativa, só a Tabela 14-5 faria sentido. As razões são:

1. Geralmente, mulheres têm status inferior na sociedade americana; portanto, elas apoiariam mais a proposta de igualdade entre os sexos.
2. O sexo do respondente afeta sua resposta ao item do questionário; mulheres têm maior probabilidade de aprovar do que os homens.

3. Se os respondentes homens e mulheres forem descritos separadamente em termos de suas respostas, uma porcentagem maior de mulheres do que de homens aprovaria.

Seguindo esta lógica, a Tabela 14-5 divide a amostra em dois grupos, homens e mulheres, depois descreve as atitudes dos dois grupos separadamente. As porcentagens expressando aprovação nos dois grupos são comparadas, e vemos que mulheres realmente têm maior probabilidade de aprovar do que homens.

Se a Tabela 14-6 fosse apresentada como análise bivariada explicativa, a lógica da tabela seria a seguinte. As atitudes relacionadas à igualdade sexual afetam o sexo da pessoa que tem estas atitudes. Aprovar a igualdade sexual tende a tornar a pessoa mais mulher do que homem. É claro que este raciocínio é absurdo. O sexo dos respondentes é predeterminado bem antes se formarem atitudes sobre igualdade sexual. Diferentes atitudes sobre igualdade sexual não podem ter efeito no fato de a pessoa com esta atitude ser homem ou mulher.

Mas a Tabela 14-6 seria legítima do ponto de vista da descrição de subgrupos e mesmo para fins de *predição*. Se soubermos a atitude de um respondente sobre igualdade sexual e quisermos predizer se ele é homem ou mulher, a Tabela 14-6 seria uma boa base sobre a qual fundamentar tal predição. Se soubermos que ele aprova a igualdade sexual, prediremos que o respondente é mulher. (*Nota:* Se fizermos várias predições independentes como esta, erraremos 46% das vezes.) Se soubermos que o respondente desaprova a igualdade sexual, adivinharemos que ele é homem (e erraremos 40% das vezes em testes repetidos).

Porém, para propósito de explicação, só a Tabela 14-5 é legítima. Para isso, é preciso conhecer a lógica das variáveis *independentes* e *dependentes*. Basicamente, você tenta explicar valores da variável dependente baseado nos valores da variável independente. Neste sentido, você raciocina que a variável independente causa a variável dependente (tipicamente, de forma probabilística). No exemplo anterior, as atitudes sobre igualdade sexual formam a variável dependente, e os sexos dos respondentes, a variável independente. Logo, sexo causa atitudes sobre igualdade sexual.

Algumas vezes, é difícil e até arbitrário determinar qual variável é dependente e qual é independente, mas algumas diretrizes ajudam. Para começar, sempre que houver uma clara

ordem temporal relacionando as duas variáveis, a variável cujos valores são determinados antes é sempre a variável independente; aquela cujos valores são determinados depois é sempre a variável dependente. A noção de causalidade retroativa no tempo é ilógica. Como os sexos dos respondentes são determinados antes de suas atitudes sobre igualdade, sexo deve ser a variável independente.

Uma implicação disso é que duas variáveis ocorrendo simultaneamente não podem ser ligadas causalmente. O sexo e a raça de uma pessoa não podem ser analisados, *per se*, de forma explicativa. É claro que a raça dos pais e o sexo da criança poderiam ser logicamente analisados desta maneira, se você acreditar, por exemplo, que pais negros têm maior probabilidade de ter filhos homens do que pais brancos. Neste caso, poderíamos determinar uma ordem temporal das variáveis. (Contudo, parece não haver evidência lógica ou empírica apoiando esta especulação.)

Em muitos casos, nenhuma ordem temporal clara parece relacionar as duas variáveis. Por exemplo, se você quiser examinar a relação causal entre escolaridade e preconceito, a ordem temporal das duas variáveis é um pouco mais ambígua do que no exemplo do sexo e atitudes. Você pode dizer que, aumentando a escolaridade, tornamos a pessoa menos preconceituosa; então, escolaridade seria a variável independente, e preconceito, a variável dependente. Por outro lado, pode ser que o preconceito afete a quantidade de educação que uma pessoa procura ou recebe. Você pode dizer que a educação seria anátema para uma pessoa muito preconceituosa e/ou uma pessoa muito preconceituosa teria maior probabilidade de ser expulsa da escola. Portanto, pode-se justificar o argumento que faz do preconceito a variável independente e da escolaridade a variável dependente.

Em situações onde a ordem temporal das variáveis não é clara, a designação de variáveis independentes e dependentes precisa ser feita e apresentada em bases lógicas. Frequentemente, o argumento não é bastante forte para satisfazer todos os leitores. Em outras situações, podemos achar que as duas variáveis se afetam de forma cíclica. Por exemplo, para alguns respondentes, escolaridade afeta o preconceito, enquanto para outros é o contrário; para outros, pode haver uma dinâmica ainda mais complexa: uma pessoa sem preconceito pode ser levada a buscar mais educação e esta reduzir ainda mais seu preconceito.

Seja como for, toda tabela bivariada explicativa implicitamente designa uma variável independente e uma dependente. Se você achar que nenhuma ordem temporal conecta as variáveis, precisa designar arbitrariamente uma ordem quase-temporal ao construir a tabela. A seguinte discussão presume que se tenha designado uma variável como dependente, desconsiderando os fundamentos desta designação.

Construção de Tabelas

A construção de tabelas bivariadas explicativas segue estes passos:

1. A amostra é dividida em valores ou categorias da variável dependente.
2. Cada subgrupo é descrito em termos dos valores ou categorias das variáveis dependentes.
3. Finalmente, a tabela é lida comparando-se os subgrupos da variável independente em termos de um valor da variável dependente.

Lembre nossa análise anterior sobre sexo e atitudes com relação à igualdade sexual. Seguindo os passos acima, sexo sendo designado como variável independente e atitudes sobre igualdade sexual, como variável dependente, procedemos assim:

1. A amostra é dividida em homens e mulheres.
2. Cada subgrupo sexual é descrito em termos de aprovação ou desaprovação da igualdade sexual.
3. Homens e mulheres são comparados em termos das porcentagens que aprovam a igualdade sexual.

A Figura 14-2 nos dá uma representação gráfica dos passos a dar na criação de tabelas de porcentagem. O número de casos foi reduzido para simplificar.

Um problema que costuma confundir pesquisadores experientes merece um comentário. A tabela deve ser percentuada “para baixo” ou “através”? Deve uma coluna de porcentagens

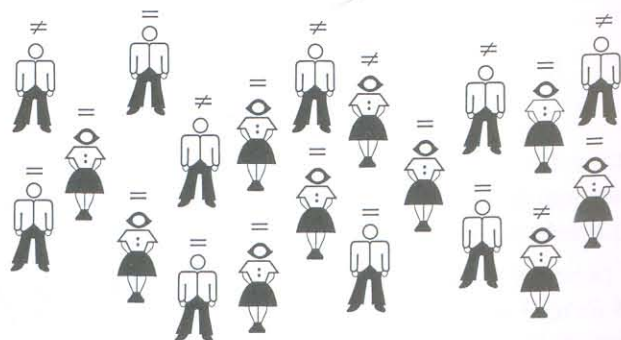
totalizar 100% ou o deve a linha de porcentagens? A resposta é completamente arbitrária. Neste livro, tendi a padronizar o procedimento calculando a porcentagem para baixo, de modo que as colunas de porcentagens totalizassem 100%, mas esta convenção é só uma questão de gosto e hábito pessoais.

Entretanto, podemos derivar dessa questão geral uma diretriz útil. Se uma tabela for percentuada para baixo, deve ser lida horizontalmente. Se é percentuada através, deve ser lida verticalmente. Usando a Tabela 14-5 como exemplo, vemos que foi percentuada para baixo, de modo que as porcentagens em cada coluna somem 100%. A tabela é interpretada lendo através: 63% dos homens aprovam, comparado a 75% das mulheres.

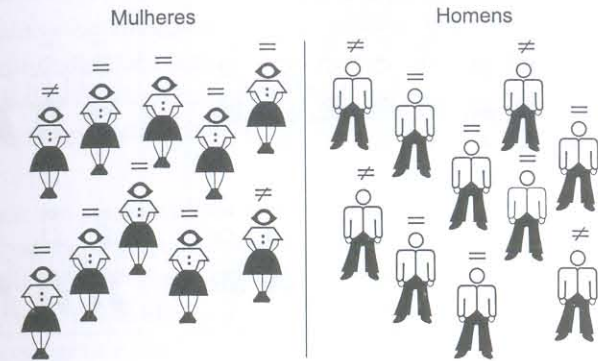
A obediência completa a esta regra de construção e interpretação de tabelas ajuda a evitar um erro comum. Muitos pesquisadores inexperientes lêem a Tabela 14-5 assim: "63% dos homens concordam com igualdade sexual enquanto 37% discordam. Portanto, os homens têm maior probabilidade de concordarem." Essa interpretação é incorreta. Embora seja verdade que os homens têm maior probabilidade de aprovar a igualdade sexual do que *desaprovar*, esse achado só tem significado como descrição das atitudes dos homens. A observação mais importante é que os homens têm uma probabilidade menor de aprovar *do que as mulheres*. Como a tabela é percentuada na vertical, deve ser lida horizontalmente.

FIGURA 14-2
Percentuando uma tabela

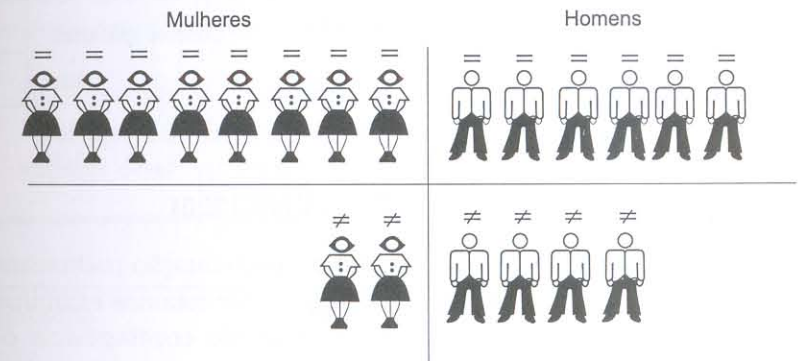
A. Alguns homens e mulheres que ou são a favor (sinal de igual) da igualdade sexual, ou são contra (≠)



B. Separar homens e mulheres (a variável independente)



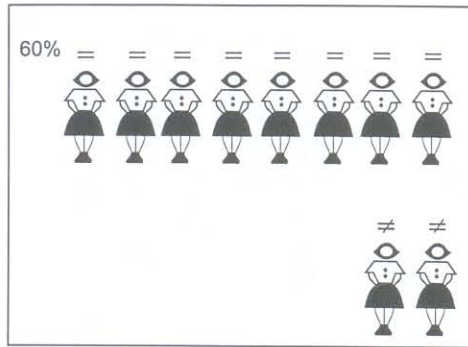
C. Em cada grupo do gênero, separar os que são a favor da igualdade sexual dos que são contra (variável dependente)



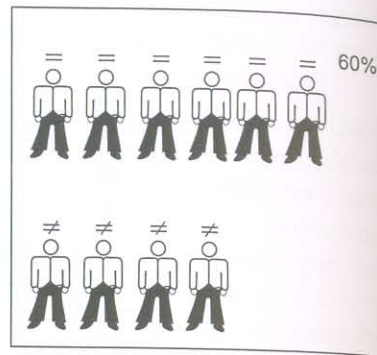
D. Conte os números em cada célula da tabela

	Mulheres	Homens
8		
2		

E. Qual a porcentagem das mulheres é a favor da igualdade?



F. Qual a porcentagem dos homens é a favor da igualdade?



G. Conclusões

Enquanto a maioria dos homens e mulheres são a favor da igualdade sexual, mulheres tem maior chances de o ser.

Portanto sexo parece ser uma das causas da atitude em relação a igualdade sexual.

	Mulheres	Homens
A favor da igualdade	80 %	60 %
Contra a igualdade	20 %	40 %
Total	100 %	100 %

Formatos de Tabelas Bivariadas

Não há formato de apresentação padronizado para o tipo de tabelas de porcentagens que estamos examinando (frequentemente chamadas tabelas de contingência ou tabulações cruzadas). Por isso, encontramos uma grande variedade de formatos na literatura de pesquisa. Se as tabelas puderem ser lidas e interpretadas facilmente, provavelmente não há razão para forçar a padronização. Contudo, certas diretrizes devem ser seguidas na apresentação da maioria dos dados tabulares. (Veja a Tabela 14-7 como exemplo de uma boa tabela.)

1. Tabelas devem ter cabeçalhos ou títulos descrevendo seu conteúdo.

2. O conteúdo original das variáveis deve ser apresentado claramente — na própria tabela, se for possível, ou no texto com uma paráfrase na tabela. Isso é especialmente crítico quando uma variável é descrita a partir de respostas a uma pergunta atitudinal, porque o significado das

respostas depende muito da construção da frase usada no item de questionário.

3. Os valores ou categorias de cada variável devem ser claramente indicados. Categorias complexas de resposta devem ser abreviadas, mas o significado deve estar claro na tabela, e a resposta completa deve ser relatada no texto.

4. Quando são apresentadas porcentagens na tabela, deve-se indicar a base sobre a qual foram computadas. É redundante apresentar todos os números brutos para cada categoria de porcentagem, já que eles podem ser reconstruídos a partir das porcentagens e das bases. Além disso, a apresentação de números e de porcentagens frequentemente torna a tabela confusa e mais difícil de ser lida.

5. Se alguns respondentes forem omitidos da tabela, devido a dados faltantes (por exemplo, "sem resposta"), seus números devem ser indicados na tabela.

TABELA 14-7

"Você aprova ou desaprova a proposição geral de que homens e mulheres devem ser tratados igualmente em todo os aspectos?"

	Homens	Mulheres
Aprovam	63%	75%
Desaprovam	37	25
	100%	100%
	(400)	(400)
Sem resposta =	(12)	(5)

Análise Multivariada

A lógica da análise multivariada é o assunto de capítulos posteriores deste livro, especialmente o Capítulo 15. Entretanto, neste ponto será útil discutir rapidamente a construção de tabelas multivariadas, isto é, tabelas construídas a partir de várias variáveis.

Tabelas multivariadas podem ser construídas com base numa descrição mais complicada de subgrupo, seguindo essencialmente os mesmos passos descritos para as tabelas bivariadas. Porém, em vez de uma variável independente e uma dependente, teremos mais de uma variável independente. Em vez de explicar a variável dependente com base numa única variável independente, buscaremos uma explicação baseada em mais de uma variável independente.

Voltando ao exemplo das atitudes sobre igualdade sexual, suponha que você acredite que idade também afeta tais atitudes, com os jovens tendo maior probabilidade de aprovar a igualdade sexual do que os mais velhos. O primeiro passo na construção de tabelas é dividir a amostra total em subgrupos baseados nos vários valores das duas variáveis independentes simultaneamente: homem jovem, homem velho, mulher jovem, mulher velha. Depois, os vários subgrupos são descritos nos termos da variável dependente, e se fazem as comparações. A Tabela 14-8 exemplifica uma tabela hipotética multivariada.

Seguindo a convenção, a tabela foi percentuada para baixo e, por isso, deve ser lida na horizontal. A interpretação leva a várias conclusões.

1. Entre homens e mulheres, os mais jovens dão mais apoio à igualdade sexual do que os mais velhos. Entre as mulheres, 90% das com menos de 30 e 60% das com 30 ou mais aprovam.
2. Dentro de cada grupo de idade, as mulheres apóiam a igualdade sexual mais que os homens. Entre os respondentes abaixo de 30 anos, 90% das mulheres apóiam, comparadas a 78% dos homens. Entre os de 30 ou mais, 60% das mulheres apóiam, contra 48% dos homens.
3. Como medido na tabela, idade parece ter maior efeito sobre atitudes do que sexo. Para homens e mulheres, o efeito da idade pode ser resumido como uma diferença de 30 pontos percentuais. Dentro de cada grupo de idade, a diferença em pontos percentuais entre homens e mulheres é de 12.
4. Idade e sexo têm efeitos independentes sobre as atitudes. Dentro de cada valor de uma variável independente, diferentes valores da segunda afetarão as atitudes.

5. Igualmente, as duas variáveis independentes têm um efeito cumulativo sobre atitude. Mulheres jovens são as que mais apóiam, e homens velhos os que apóiam menos.

TABELA 14-8
"Você aprova ou desaprova a proposição geral de que homens e mulheres devem ser tratados igualmente em todos os aspectos?"

	Mulheres		Homens	
	Abaixo de 30	30 e acima	Abaixo de 30	30 e acima
Aprovam	90%	60%	78%	48%
Desaprovam	10	40	22	52
	100%	100%	100%	100%
	(200)	(200)	(200)	(200)
Sem resposta =	(2)	(3)	(10)	(2)

O Capítulo 15 sobre o modelo de elaboração examina com mais detalhes a lógica da análise multivariada. Mas, antes de concluir esta seção, vejamos um formato alternativo para apresentar dados.

Cada tabela apresentada neste capítulo é ineficiente de algum modo. Como a variável dependente — atitude sobre igualdade sexual — é dicotômica (tem dois valores), conhecer um valor permite reconstruir o outro facilmente. Assim, se sabemos que 90% das mulheres com menos de 30 anos aprovam a igualdade sexual, sabemos automaticamente que 10% desaprovam (supondo que removemos os "não sei" e outras respostas desse tipo). Portanto, é desnecessário relatar a porcentagem dos que desaprovam. Por isso, a Tabela 14-8 poderia ser apresentada no formato alternativo da Tabela 14-9.

Nesta, as porcentagens aprovando a igualdade sexual são relatadas nas células representando as interseções das duas variáveis independentes. Os números entre parênteses abaixo de cada porcentagem representam os números de casos sobre os quais as porcentagens são baseadas. Por exemplo, sabemos que havia 200 mulheres abaixo de 30 anos na amostra e que 90%

delas aprovavam a igualdade sexual. Além disso, que 180 das 200 mulheres aprovavam e que as outras 20 (10%) desaprovavam. É mais fácil ler a Tabela 14-9 do que a 14-8, sem sacrificar qualquer detalhe.

TABELA 14-9

"Você concorda ou discorda da proposição geral de que homens e mulheres devem ser tratados igualmente em todos os aspectos?"

Porcentagem que Concorda	Mulheres	Homens
Abaixo de 30	90 (200)	78 (200)
30 e acima	60 (200)	48 (200)

Resumo

Este capítulo introduziu a lógica e a mecânica da construção de tabelas. À primeira vista, tabelas parecem muito simples para merecer discussão extensa. Na verdade, elas são bastante complexas. Por isso, são freqüentemente malconstruídas e mal-interpretadas.

Tentei apresentar a lógica da construção e interpretação de tabelas, começando com a análise univariada, passando para a descrição de subgrupos e depois discutindo a análise explicativa bivariada e a análise multivariada. Os capítulos seguintes dependem muito da compreensão das tabelas de contingência, portanto é essencial que você as entenda bem.

Vale a pena repetir os passos essenciais da construção e interpretação das tabelas explicativas.

1. Divida a amostra em subgrupos baseados nos valores da(s) variável(eis) independente(s).

2. Descreva cada grupo com base nos valores da variável dependente.

3. Compare os subgrupos da variável independente em relação ao valor da variável dependente.

Finalmente, memorize esta regra básica: percentue para baixo e leia através, ou percentue através e leia para baixo.

Nota

¹ Dados extraídos de HAUB, Carl, KENT, Mary Mederios. *1988 World Population Data Sheet* (a poster). Washington, DC: Population Reference Bureau.

Leituras Adicionais

COLE, Stephen. *The Sociological Method*. Chicago: Markham, 1972.

DAVIS, James. *Elementary Survey Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

LABOVITZ, Sanford, HAGEDORN, Robert. *Introduction to Social Research*. New York: McGraw-Hill, 1971.

ZEISEL, Hans. *Say It with Figures*. New York: Harper & Row, 1957.

Capítulo 15

O Modelo de Elaboração

Este capítulo estuda aquilo que, em análise de *survey*, é chamado de “modelo de elaboração”, “método de interpretação”, “a escola de Columbia” ou “método de Lazarsfeld”. Essa nomenclatura variada deriva do fato de que este método tem por fim a *elaboração* de uma relação empírica entre variáveis, para interpretar essa relação da maneira desenvolvida por Paul Lazarsfeld na Universidade de Columbia.

O modelo de elaboração é usado para tornar inteligível a relação entre duas variáveis através da introdução simultânea de variáveis adicionais. Foi desenvolvido primeiramente através de tabelas de contingência, mas em capítulos posteriores veremos como pode ser usado com outras técnicas estatísticas.

Acredito firmemente que o modelo de elaboração oferece uma visão mais nítida da lógica da análise de *survey* do que qualquer outro método disponível. Especialmente quando usado com tabelas de contingência, este método reproduz os processos lógicos da análise científica. Ademais, compreendendo completamente o uso do modelo de elaboração com tabelas de contingência, aumentamos nossa capacidade de usar e compreender técnicas estatísticas mais sofisticadas.

História do Modelo de Elaboração

Rever as origens históricas do modelo de elaboração é particularmente instrutivo para a apreciação realista da pesquisa científica na prática. Durante a II Guerra Mundial, Samuel Stouffer organizou e chefou um grupo especial de pesquisa social dentro do Exército dos Estados Unidos. Durante toda a

guerra, esse grupo fez um grande número e variedade de *surveys* entre os militares americanos. Embora os objetivos das pesquisas variassem, o foco estava, em geral, nos fatores que afetavam a efetividade dos soldados em combate.

Diversas pesquisas examinavam a questão do moral entre os militares. Como se acreditava que o moral afetava a efetividade em combate, sua melhora presumivelmente aumentaria a efetividade do esforço de guerra. Stouffer e seu grupo de pesquisa, portanto, procuraram desvendar algumas das variáveis que afetavam o moral. Em parte, o grupo buscou confirmar empiricamente algumas proposições comumente aceitas, inclusive as seguintes:

1. Promoções certamente afetavam o moral dos soldados, e os que serviam em unidades com baixas taxas de promoções teriam moral relativamente baixo.
2. Dadas a discriminação e a segregação social no Sul do país, soldados negros em treinamento nas bases do Norte teriam um moral mais alto do que os que treinavam no Sul.
3. Soldados com mais escolaridade provavelmente se ressentiriam mais de serem recrutados como soldados rasos do que os com menos escolaridade.

Essas proposições tinham lógica, e o saber comum sustentava que eram empiricamente verdadeiras. Stouffer decidiu testá-las empiricamente; para sua surpresa, nenhuma foi confirmada.

Primeiro, os soldados que serviam na Polícia do Exército, onde as promoções eram as mais lentas, tinham menos queixas sobre o sistema de promoções do que os que serviam na Aviação do Exército, onde as promoções eram as mais rápidas do Exército. Este achado surgiu das respostas a uma pergunta sobre se o soldado acreditava que o sistema de promoções era justo em geral. Segundo, o moral geral dos soldados negros nos campos de treinamento do Norte e nos do Sul diferia pouco, se é que diferia. Terceiro, soldados menos escolarizados tendiam mais a se ressentir de terem sido recrutados do que os de maior escolaridade.

Frente a esses dados, muitos pesquisadores, sem dúvida, teriam tentado escondê-los, por sugerirem pouca capacidade científica. Outros teriam feito testes de significância estatística e tentariam publicar os resultados. Stouffer, em vez disso, perguntou por que tinha obtido tais dados.

Ele encontrou a resposta nos conceitos de "grupo de referência" e de "privação relativa". Stouffer sugeriu que os soldados avaliavam suas posições na vida não de acordo com padrões absolutos e objetivos, mas com base em sua posição relativamente aos outros em seu redor. As pessoas com quem se comparavam eram seu grupo de referência, e sentiam privação relativa se não se comparassem favoravelmente com elas.

Com esses conceitos de grupo de referência e privação relativa, Stouffer encontrou uma resposta para as anomalias em seus dados. Com relação à promoção, sugeriu que os soldados avaliavam a justiça do sistema de promoções com base nas próprias experiências relativas a outros em torno deles. Na Polícia do Exército, onde as promoções eram poucas e lentas, menos soldados conheciam um colega menos qualificado que tivesse sido promovido antes deles. Na Aviação do Exército, a rápida taxa de promoções significava que muitos soldados sabiam de colegas menos qualificados que tinham sido promovidos mais rapidamente do que seria justo. Assim, ironicamente, os policiais do exército diziam que o sistema de promoção era geralmente justo, e os da Aviação diziam que não era.

Uma explicação semelhante parece adequada no caso dos soldados negros. Em vez de comparar as condições gerais do Norte com as do Sul, eles comparavam seu próprio status, como soldados negros, com o status de civis negros em seu redor. No Sul, onde a discriminação estava na pior fase, ser soldado protegia um pouco os negros das normas culturais adversas nas comunidades ao redor. Enquanto civis negros eram fortemente discriminados e negados à auto-estima, bons empregos etc., soldados negros tinham um status ligeiramente melhor. No Norte, todavia, muitos civis negros que os soldados encontravam tinham empregos bem pagos; com menor discriminação, ser soldado não melhorava o status de um negro na comunidade.

Finalmente, grupo de referência e privação relativa pareciam explicar a anomalia de recrutas mais escolarizados aceitando seu recrutamento com mais boa vontade do que os menos escolarizados. Stouffer raciocinou assim:¹

1. Os amigos de um soldado terão, em geral, o mesmo nível educacional que ele.
2. Soldados com menor escolaridade terão maior probabilidade de estar em ocupações semiqualficadas

de produção em série ou de agricultura do que os mais bem-educados.

3. Durante a guerra, muitas indústrias de produção em série e unidades agrícolas foram declaradas vitais para o interesse nacional; agricultores e trabalhadores de produção em série nessas indústrias eram dispensados do serviço militar.

4. Soldados com pouca escolaridade tinham maior probabilidade de ter amigos isentos de serviço militar do que soldados com maior escolaridade.

5. Recrutas com baixa escolaridade tinham maior probabilidade de se sentirem discriminados do que os com maior escolaridade, ao se compararem com os amigos.

Essas foram as explicações que Stouffer sugeriu para desvendar o mistério dos três achados anômalos. Mas como isso não fazia parte de um projeto de estudo pré-planejado, ele careceu de dados empíricos para testá-las. Ainda assim, a exposição lógica de Stouffer forneceu a base para o desenvolvimento posterior do modelo de elaboração: o entendimento da relação entre duas variáveis através da introdução controlada de outras variáveis.

TABELA 15-1

Resumo dos dados de Stouffer sobre escolaridade e aceitação do recrutamento

	Alta Esc.	Baixa Esc.
Não deveriam ser dispensados	18%	70%
Deveriam ser dispensados	12%	30%
	100%	100%
	(1761)	(1876)

As Tabelas 15-1, 15-2, 15-3 e 15-4 estão modificadas com permissão da Macmillan Company, a partir de *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"*, por Robert K. Merton e Paul F. Lazarsfeld. Copyright 1950 by The Free Press, a Corporation.

TABELA 15-2

Relação hipotética entre escolaridade e dispensa de amigos

	Alta Esc.	Baixa Esc.
Sim	19%	79%
Não	81	21
	100%	100%
	(1761)	(1876)

Amigos dispensados?

O desenvolvimento formal do modelo de elaboração foi feito por Paul Lazarsfeld e seus associados na Universidade de Columbia. Na revisão metodológica dos estudos de Stouffer sobre o Exército, Lazarsfeld e Patricia Kendall apresentaram tabelas hipotéticas que teriam provado a argumentação de Stouffer sobre a relação entre nível educacional e aceitação do recrutamento, caso os dados empíricos estivessem disponíveis.²

Kendall e Lazarsfeld começaram com os dados de Stouffer que mostravam a associação negativa entre escolaridade e aceitação do recrutamento (ver Tabela 15-1). Seguindo a explicação de Stouffer, Kendall e Lazarsfeld criaram uma tabela hipotética, compatível com os dados empíricos, para mostrar que escolaridade se relacionava com o soldado ter ou não amigos dispensados. Na Tabela 15-2; repare que 19% dos com alta escolaridade relataram ter amigos que foram dispensados, comparados com 79% dos com baixa escolaridade.

A explicação de Stouffer supunha, em seguida, que soldados com amigos dispensados teriam maior probabilidade de se ressentir com seu próprio recrutamento do que os que não tinham amigos dispensados. A Tabela 15-3 apresenta os dados hipotéticos de Kendall e Lazarsfeld que apoiariam esta suposição.

Os dados hipotéticos das Tabelas 15-2 e 15-3 confirmam as ligações especificadas por Stouffer em sua explicação. Primeiro, soldados com baixa escolaridade tinham maior

probabilidade de terem amigos dispensados do que os com mais escolaridade. Segundo, ter amigos dispensados faz que um soldado tenda a pensar que ele também devia ter sido dispensado. Stouffer tinha sugerido que essas duas relações esclareceriam a relação original entre escolaridade e aceitação do recrutamento. Kendall e Lazarsfeld criaram a tabela hipotética que confirmaria essa última explicação (ver Tabela 15-4).

Lembre que o achado original foi que recrutas com maior escolaridade tinham mais probabilidade de aceitar como justo seu recrutamento do que os de escolaridade mais baixa. Na Tabela 15-4, entretanto, notamos que escolaridade não tem efeito na aceitação do recrutamento nos que relatam terem tido amigos dispensados. Nos dois grupos de escolaridade, 63% dizem que não deviam ter sido dispensados. De modo semelhante, escolaridade não tem efeito significativo na aceitação de recrutamento nos que relataram não terem tido amigos dispensados; 94 e 95% dos dois grupos de escolaridade dizem que eles não deviam ter sido dispensados.

TABELA 15-3

Relação hipotética entre a dispensa de amigos e a aceitação do próprio recrutamento

	Amigos dispensados?	
	Sim	Não
Não deveriam ter sido dispensados	63%	94%
Deveriam ter sido dispensados	37	6
	100%	100%
	(1819)	(1818)

TABELA 15-4
Dados hipotéticos relacionando escolaridade à aceitação do recrutamento, pelo fator de ter amigos dispensados

	Amigos dispensados		Nenhum amigo dispensado	
	Alta Esc.	Baixa Esc.	Alta Esc.	Baixa Esc.
Não deveriam ter sido dispensados	63%	63%	94%	95%
Deveriam ter sido dispensados	37	37	6	5
	100%	100%	100%	100%
	(335)	(1484)	(1426)	(392)

Por outro lado, entre os de alta escolaridade, a aceitação do recrutamento é fortemente relacionada a terem amigos dispensados ou não: 63% versus 94%. O mesmo vale para os menos escolarizados. Os dados hipotéticos da Tabela 15-4, portanto, apóiam o argumento de Stouffer de que escolaridade afetava a aceitação do recrutamento apenas pela intermediação de ter tido amigos dispensados. Recrutas de alta escolaridade tinham menor probabilidade de terem amigos dispensados e, em virtude disso, tinham maior probabilidade de aceitarem seu recrutamento como justo. Recrutas com menor escolaridade tinham maior probabilidade de terem amigos dispensados e, por isso, menor probabilidade de aceitarem seu próprio recrutamento.

É importante reconhecer que nem a explicação de Stouffer nem os dados hipotéticos negaram a realidade da relação original. À medida que a escolaridade aumentava, aumentava também a aceitação do próprio recrutamento. A natureza dessa relação empírica, entretanto, foi interpretada pela introdução de uma terceira variável. Essa variável, a dispensa de amigos, não negava a relação original; ela apenas esclareceu o mecanismo pelo qual ocorreu a relação original. É este, pois, o coração do modelo de elaboração e da análise multivariada.

Tendo observado uma relação empírica entre duas variáveis, busca-se entender a natureza dessa relação através dos efeitos produzidos pela introdução de outras variáveis. Mecanicamente, isso se faz, primeiro, dividindo a amostra em subconjuntos, com base na variável de controle ou de teste. Por exemplo, ter amigos dispensados ou não é a variável de

controle do nosso exemplo; a amostra é dividida entre os que têm e os que não têm amigos dispensados. A relação entre as duas variáveis originais é recomputada separadamente para cada sub-amostra. As tabelas produzidas assim são chamadas de tabelas parciais, e as relações encontradas nas tabelas parciais são chamadas de relações parciais. As relações parciais são a seguir comparadas com a relação inicial descoberta na amostra total.

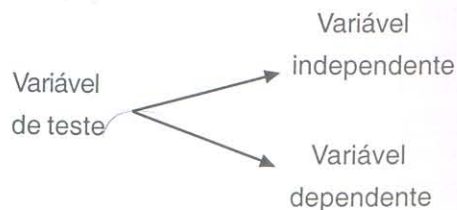
O Paradigma de Elaboração

Esta seção dá diretrizes para entender uma análise de elaboração. Para começar, deve-se saber se a variável de teste é antecedente (anterior no tempo) às outras duas variáveis ou se intervém entre elas; essas diferenças sugerem diferentes relações lógicas no modelo multivariado. Se a variável de teste é interveniente, como é o caso da escolaridade, dispensa de amigos e aceitação do recrutamento, então as relações da Figura 15-1 são pertinentes. A lógica desta relação multivariada é a seguinte: a variável independente (escolaridade) afeta a variável de teste interveniente (amigos dispensados ou não), que, por sua vez, afeta a variável dependente (aceitar o recrutamento).

FIGURA 15-1



FIGURA 15-2



Se a variável de teste é antecedente tanto à variável independente quanto à dependente, uma relação multivariada muito diferente se aplica (ver Figura 15-2). Nessa segunda situação, a variável de teste afeta tanto a variável independente quanto a dependente.³ Por causa de suas relações individuais com a variável de teste, as variáveis independente e dependente se

relacionam empiricamente, mas não há nenhuma conexão causal entre elas. Sua relação empírica é somente um produto de suas relações com a variável de teste. (Exemplos posteriores esclarecerão melhor este ponto.)

A Tabela 15-5 é um guia para o entendimento da análise de elaboração. As duas colunas na tabela indicam se a variável de teste é antecedente ou interveniente no sentido descrito acima. No lado esquerdo da tabela, indica-se a natureza das relações parciais comparadas com a relação original entre as variáveis independente e dependente. As notações técnicas designadas a cada caso são dadas no corpo da tabela.

Replicação

Quando as relações parciais são essencialmente as mesmas que a relação original, chamamos o resultado de *replicação*, não importando se a variável de teste é antecedente ou interveniente. O significado é essencialmente o mesmo que ditaria o senso comum. A relação original foi replicada sob condições de teste. Se, em nosso exemplo, escolaridade afetasse aceitação do recrutamento tanto entre os que tiveram amigos dispensados e os que não tiveram, diríamos que a relação original tinha sido replicada. Contudo, este achado não confirmaria a explicação dada por Stouffer à relação original. Ter amigos dispensados ou não constituiria o mecanismo pelo qual a escolaridade afetava a aceitação do recrutamento.

TABELA 15-5
O paradigma da elaboração

Relações Parciais Comparadas c/ Original	Variável de Teste	
	Antecedente	Interveniente
Mesma relação	Replicação	
Menor ou nenhuma	Explicação	Interpretação
Dividida*	Especificação	

* Um parcial é o mesmo ou maior, enquanto o outro é menor ou nenhum.

Os pesquisadores usam rotineiramente o modelo de elaboração na esperança de replicar seus achados entre subconjuntos da amostra. Se se descobre, por exemplo, uma relação entre escolaridade e preconceito, pode-se introduzir variáveis de teste como idade, região do país, raça, religião etc. para testar a estabilidade da relação original. Se a relação se replicar para velhos e jovens, para pessoas de diversas partes do país etc., pode-se concluir que a relação original é genuína e geral.

Explicação

Explicação é o termo usado para descrever uma *relação espúria*, isto é, uma relação original que é “anulada” pela introdução de uma variável de teste. Duas condições são necessárias para que isto ocorra. A variável de teste deve ser antecedente tanto à variável independente quanto à dependente, e as relações parciais devem ser zero ou significativamente menores do que a encontrada no original. Três exemplos devem esclarecer o ponto.

Primeiro, há uma relação empírica entre o número de cegonhas em diferentes áreas e as taxas de nascimento naquelas áreas. Quanto mais cegonhas numa área, mais nascimentos. A relação empírica pode levar a supor que o número de cegonhas afeta a taxa de nascimentos. Uma variável de teste antecedente explica essa relação. Áreas rurais têm mais cegonhas e mais nascimentos que áreas urbanas. Dentro das áreas rurais, não há nenhuma relação entre o número de cegonhas e a taxa de nascimentos, e essa relação também não existe nas áreas urbanas.

Segundo, há uma relação positiva entre o número de caminhões de bombeiros que correm para atender a um incêndio e as dimensões dos danos causados. Se mais caminhões respondem, há mais danos causados. Pode-se supor, com isso, que os próprios caminhões de bombeiros causam os danos. Uma variável de teste antecedente, contudo — o tamanho do incêndio —, explica a relação original. Incêndios maiores causam mais danos que os menores, e mais caminhões respondem aos maiores do que aos menores. Considerar apenas incêndios grandes faria desaparecer (ou talvez reverter) a relação original; o mesmo ocorreria se se considerasse apenas os incêndios menores.

Por fim, há uma relação empírica entre a região do país em que os professores de escolas de medicina fizeram seus cursos de medicina e suas atitudes com relação ao Medicare.⁴ Para simplificar, examinemos apenas o Leste e o Sul. Dos

professores que freqüentaram escolas de medicina do Leste, 78% disseram aprovar o Medicare, comparados com 59% dos que freqüentaram escolas sulistas de medicina. Este achado faz sentido em virtude de o Sul, em geral, ser mais resistente a tais programas que o Leste e de o treinamento na escola de Medicina presumivelmente afetar as atitudes do médico. Esta relação, entretanto, é anulada pela introdução de uma variável de teste antecedente: a região do país em que o professor foi criado.

Dos professores criados no Leste, 89% freqüentaram escolas de medicina no Leste, e 11%, escolas de medicina no Sul. Dos criados no Sul, 53% freqüentaram escolas de medicina no Leste, e 47%, no Sul. Além disso, a área em que os professores foram criados se relaciona com as atitudes para com o Medicare. Dos criados no Leste, 84% aprovaram o Medicare, comparados com 49% dos criados no Sul.

A Tabela 15-6 apresenta a relação de três variáveis entre a região em que os professores foram criados, a região do treinamento médico e as atitudes com relação ao Medicare. Os professores de medicina criados no Leste têm grande probabilidade de aprovar o Medicare, não importa onde tenham estudado. Do mesmo modo, os criados no Sul tinham menor probabilidade de aprovar o Medicare, mas, de novo, a região em que foram treinados tinha pouco ou nenhum efeito. Esses dados indicam que a relação original entre região do treinamento médico e atitudes com relação ao Medicare era espúria. Deveu-se apenas ao efeito coincidente da região de origem tanto na região do treinamento médico quanto nas atitudes para com o Medicare. Quando a região de origem é *mantida constante*, como na Tabela 15-6, a relação original desaparece nas relações parciais.

TABELA 15-6

Região de origem, região de treinamento e atitudes para com o Medicare

Percentual dos que aprovam Medicare	Região de Criação	
	Leste	Sul
Região do Treinamento Médico	84	50
	Sul	80
		47

FONTE — BABBIE, Earl R. *Science and Morality in Medicine*. Berkeley: University of California Press, 1970. Ver especialmente p.181.

Interpretação

A interpretação é semelhante à explicação, exceto pela posição temporal da variável de teste e das implicações que se seguem dessa diferença. O exemplo de escolaridade, amigos dispensados e aceitação do recrutamento é uma excelente ilustração da interpretação. Nos termos do modelo de elaboração, o efeito da escolaridade na aceitação do recrutamento não é anulado: continua a ser uma relação genuína. Realmente, as diferenças de escolaridade *causam* uma aceitação diferencial do recrutamento. A variável interveniente, dispensa dos amigos, somente ajuda a interpretar o mecanismo pelo qual ocorre a relação.

Um ponto importante deve ser notado. Podia-se ter começado a análise observando que ter amigos dispensados tornava os recrutas menos dispostos a aceitar como justo o próprio recrutamento. Tentando entender melhor este achado original, podia-se ter introduzido escolaridade como variável de teste antecedente. Se isso tivesse sido feito, todavia, ter-se-ia achado que a relação entre ter amigos dispensados e aceitar o recrutamento foi *replicada* entre os soldados de maior escolaridade e os menos escolarizados (ver Tabela 15-4). Teria também sido notado que os soldados de maior escolaridade tinham menor probabilidade de ter amigos dispensados, mas a relação original não teria sido anulada.

Como exemplo final de interpretação, pesquisadores observaram, no passado, que crianças cujas mães trabalham fora têm maior probabilidade de se tornarem delinqüentes que crianças cujas mães não trabalham fora. Esta relação pode ser interpretada introduzindo "supervisão" como variável de teste. Entre as crianças supervisionadas, as taxas de delinqüência não são afetadas pela mãe trabalhar fora ou não. O mesmo vale para as crianças não supervisionadas. É a relação entre mães que trabalham e falta de supervisão que produziu a relação original.

Especificação

Algumas vezes, o modelo de elaboração produz relações parciais que diferem significativamente entre si. Por exemplo, uma relação parcial pode parecer muito com a relação original entre duas variáveis, enquanto a segunda relação parcial se aproxima de zero. Chama-se isto de *especificação* no paradigma de elaboração. Especificaram-se as condições nas quais ocorre a relação original.

TABELA 15-7

Classe social e envolvimento médio com a igreja entre mulheres episcopais

	Níveis de Classe Social				
	Baixo		Alto		
	0	1	2	3	4
Envolvimento Médio	0,63	0,58	0,49	0,48	0,45

FORTE - GLOCK, Charles Y. et al. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. p.85.

TABELA 15-8

Classes sociais e a ocupação de cargos em organizações seculares

	Níveis de Classe Social				
	Baixo		Alto		
	0	1	2	3	4
Porcentagem dos detentores de cargos em organizações Seculares	46	47	54	60	83

FORTE - GLOCK, Charles Y. et al. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. p.92.

Num estudo das fontes do envolvimento religioso, Glock e seus associados descobriram que, entre membros da Igreja Episcopal, o envolvimento diminuía com o aumento de classe social.⁵ A Tabela 15-7 examina níveis médios de envolvimento na igreja encontrados entre paroquianas de níveis sociais diversos. Glock interpretou este achado no contexto de outros na análise, e concluiu que o envolvimento nas atividades da igreja provê uma forma alternativa de gratificação para pessoas que não a encontram na sociedade secular. Esta interpretação explica por que mulheres eram mais religiosas que homens, por que idosos eram mais religiosos que jovens, e assim por diante. Glock raciocinou que pessoas de classe social mais baixa (medida por renda e escolaridade) tinham menos oportunidade de obter auto-estima na sociedade secular do que pessoas de classe social mais alta. Para exemplificar, ele notou que classe social se relacionava fortemente com o fato de a mulher ter algum cargo numa organização secular (ver Tabela 15-8).

Glock raciocinou então que, se classe social se relacionava com envolvimento na igreja apenas em virtude do fato de as mulheres de classe mais baixa não terem oportunidade de gratificação na sociedade secular, a relação original não deveria se manter para mulheres que recebiam gratificação social. Como indicador da recepção de gratificação na sociedade secular, usou a variável ocupação de cargos em organizações seculares. Nos termos deste teste, classe social não deveria se relacionar com envolvimento com a igreja entre mulheres que tinham tais cargos (ver Tabela 15-9).

TABELA 15-9

Envolvimento na igreja por classe social e ocupação de cargos seculares

	Níveis de Classe Social				
	Baixo				Alto
Principal envolvimento da Igreja	0	1	2	3	4
Com cargos	0,46	0,53	0,46	0,46	0,46
Sem cargos	0,62	0,55	0,47	0,46	0,40

FORTE- GLOCK, Charles Y. et al. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. p.92.

A Tabela 15-9 dá um exemplo de especificação. Entre as mulheres que tinham cargos em organizações seculares, não há essencialmente qualquer relação entre classe social e envolvimento com a igreja. Com efeito, a tabela especifica as condições nas quais a relação original se mantém, como a falta de gratificação na sociedade secular.

O termo *especificação* é usado no paradigma da elaboração independentemente da variável de teste ser interveniente ou antecedente. O significado é o mesmo em qualquer caso: especificam-se as condições particulares nas quais a relação original se mantém.

Refinamentos do Paradigma

As seções anteriores apresentam a lógica primária do modelo da elaboração tal como desenvolvido por Lazarsfeld e colegas. Morris Rosenberg fez uma excelente apresentação desse paradigma, indo além dele para sugerir outras variações.⁶ Em vez de resenhar os comentários de Rosenberg, pode ser útil agora considerar as variações logicamente possíveis. Alguns desses comentários podem ser encontrados no livro de Rosenberg; outros foram sugeridos por ele.

Primeiro, o paradigma básico supõe uma relação original entre duas variáveis. Um modelo mais abrangente, contudo, poderia diferenciar proveitosamente relações positivas e negativas. Além do mais, Rosenberg sugere a aplicação do modelo de elaboração a uma relação original *zero*, com a possibilidade de surgirem relações nas parciais. Ele cita como exemplo um estudo de filiação a sindicatos e atitudes sobre haver judeus na diretoria do sindicato.⁷ A análise inicial indicava que tempo de filiação não se relacionava com a atitude. Os que pertenciam ao sindicato há menos de quatro anos estavam tão dispostos a aceitar judeus na diretoria quanto os que lá estavam há mais de quatro anos. A *idade* dos membros, entretanto, *suprimia* a relação entre tempo de filiação e atitudes sobre judeus. Em geral, os membros mais jovens eram mais favoráveis aos judeus do que os mais velhos. Ao mesmo tempo, é claro, os mais jovens eram menos passíveis de pertencer ao sindicato por tanto tempo quanto os mais velhos. Dentro de grupos de idade específicos, contudo, os filiados ao sindicato há mais tempo eram os que mais apoiavam a presença de judeus. Idade, neste caso, era uma *variável supressora*, escondendo a relação entre tempo de filiação ao sindicato e atitudes sobre judeus.

Segundo, o paradigma básico considera os casos de as parciais serem tão fortes ou mais fracas que a relação original, mas não oferece diretrizes para decidir o que constitui uma diferença significativa entre a relação original e as parciais. Todo pesquisador que usa o modelo de elaboração tem freqüentemente que decidir arbitrariamente sobre se uma parcial é significativamente mais fraca que a original. Este dilema, portanto, sugere uma outra dimensão do paradigma.

Terceiro, a limitação do paradigma básico a parciais iguais ou mais fracas que a relação original negligencia duas outras possibilidades. Uma relação parcial pode ser *mais forte* que a original, ou uma relação parcial pode ser o inverso da original — negativa, enquanto a original era positiva.

Rosenberg dá um exemplo hipotético desta situação sugerindo que o pesquisador pode descobrir respondentes da classe trabalhadora que apóiam mais o movimento dos Direitos Civis que os de classe média.⁸ Sugere ainda que *raça* pode ser uma *variável de distorção* neste caso, distorcendo a verdadeira relação entre classe e atitudes. Presumivelmente, respondentes negros apoiariam mais o movimento do que brancos, mas também negros estariam sobre-representados entre os trabalhadores e sub-representados na classe média. Todavia, negros de classe média poderiam apoiar mais o movimento do que trabalhadores negros, e a mesma relação poderia ser encontrada entre brancos. Mantendo constante a *raça*, então, concluiríamos que o apoio ao movimento dos Direitos Civis seria maior na classe média do que na classe operária.

Essas novas dimensões complicam mais ainda a noção de especificação. Se uma parcial é igual à original, enquanto a outra parcial é ainda mais forte que a original, qual seria a reação? Houve a especificação de uma condição na qual a relação original se mantém, mas também se especificou uma outra condição na qual a relação se mantém ainda mais.

Por fim, o paradigma básico enfoca principalmente as variáveis de teste dicotômicas. O modelo de elaboração não é tão limitado, nem na teoria nem no uso, mas o paradigma básico fica mais complicado quando a variável de teste divide a amostra em três ou quatro subamostras e se complica ainda mais quando se usa mais de uma variável de teste ao mesmo tempo.

Esses comentários não são feitos para apontar defeitos no paradigma de elaboração. Ao contrário, a intenção é ressaltar que o modelo de elaboração não é um algoritmo simples, ou um conjunto de procedimentos através dos quais a análise se faz. O modelo de elaboração é um dispositivo lógico usado para assisti-lo na compreensão dos dados. Uma boa compreensão do modelo de elaboração facilitará uma análise de *survey* sofisticada. No entanto, ele não sugere quais variáveis devem ser introduzidas como controles, nem sugere conclusões definitivas sobre a natureza dos resultados da elaboração. Para isso, deve-se confiar na própria engenhosidade, a qual só aparece depois de extensa experiência. Ao apontar as supersimplificações no paradigma de elaboração básico, tentei mostrar que o modelo fornece apenas um quadro de referências lógico. A análise sofisticada será muito mais

complexa do que os exemplos usados para ilustrar o paradigma básico.

Ao mesmo tempo, o paradigma de elaboração é um quadro de referências lógico muito *poderoso*. Se você entender bem o modelo básico, terá melhores condições para entender outras técnicas, como correlações, regressões, análises fatoriais etc. O próximo capítulo procura situar técnicas como correlações parciais e regressões parciais no contexto do modelo de elaboração.

Elaboração e Hipótese *Ex Post Facto*

Devemos fazer uma observação sobre o poder do modelo de elaboração em conexão com uma infeliz "vaca sagrada" das normas tradicionais da pesquisa científica. O leitor de textos metodológicos encontrará inúmeras referências à falácia da *hipotetização ex post facto*. As intenções desses avisos são boas, mas o pesquisador inexperiente pode se deixar enganar.

Observar uma relação empírica entre duas variáveis e depois sugerir uma razão para ela é, às vezes, chamado *hipotetização ex post facto* — gerar uma hipótese ligando duas variáveis depois de sua relação já ser conhecida. Lembre-se da discussão anterior de que todas as hipóteses devem ser desconfirmáveis. A menos que se possa especificar achados empíricos que desprovem sua hipótese, ela é essencialmente inútil. Raciocina-se, portanto, que uma vez tendo *observado* uma relação entre duas variáveis, qualquer hipótese relativa a essa relação não pode ser falseada.

Essa é uma avaliação justa apenas em casos em que se revestem observações empíricas com hipóteses enganosas após o fato. Tendo observado que mulheres são mais religiosas que homens, não se deve simplesmente afirmar que mulheres são mais religiosas por causa de alguma dinâmica geral do comportamento social e, a partir daí, apoiar seu argumento nessa observação inicial.

O efeito infeliz da injunção contra a hipotetização *ex post facto* é inibir boas e honestas hipóteses pós fato. Pesquisadores inexperientes são freqüentemente levados a crer que devem formular todas as suas hipóteses antes de examinar os dados, mesmo que isso signifique fazer uma porção de hipóteses malformuladas. Além disso, são levados a ignorar qualquer relação empiricamente observada que não confirme alguma hipótese anterior.

Certamente, poucos pesquisadores gostariam que Sam Stouffer ocultasse seus achados anômalos a respeito do moral dos soldados no Exército. Stouffer viu observações empíricas peculiares e procurou hipóteses sobre as razões dos seus achados. Seu raciocínio provou ser valioso para futuros pesquisadores.

Mas devemos assinalar que, embora qualquer pessoa possa gerar hipóteses para explicar relações empíricas observadas num conjunto de dados, o modelo de elaboração fornece as ferramentas lógicas para testar essas hipóteses dentro do mesmo conjunto de dados. Um bom exemplo é a discussão anterior sobre classe social e envolvimento na igreja. Glock explicou a relação original em termos da teoria da privação social. Se ele tivesse parado ali, seus comentários teriam sido interessantes, mas pouco persuasivos. Ele, contudo, foi além e observou que, se as hipóteses estavam corretas, a relação entre classe social e envolvimento na igreja deveria desaparecer para as mulheres que tinham outra gratificação na sociedade secular, especificamente as que ocupavam cargos numa organização secular. Essa relação foi então submetida a um teste empírico. Se a nova hipótese não tivesse sido confirmada pelos dados, Glock teria sido forçado a reconsiderar.

Esses comentários devem mostrar que a análise de dados é um processo contínuo, que exige toda a engenhosidade e perseverança que se puder reunir. O cenário de um pesquisador cuidadosamente expondo as hipóteses e testando-as de modo ritualístico resulta apenas em pesquisa ritualística.

Se você estiver preocupado com a menor força das provas *ex post facto* em comparação com as tradicionais, deixe-me repetir que “prova científica” é uma contradição em termos. Nada é provado *cientificamente*. Hipóteses, explicações, teorias ou palpites podem escapar de tentativas de desconfirmção, mas nenhuma pode ser provada de modo absoluto. A aceitação de uma hipótese, portanto, é uma função do grau em que ela foi testada e não desconfirmada. Nenhuma hipótese pode ser considerada sólida com base num único teste, não importa se gerada antes ou depois da observação de dados empíricos. Com isso em mente, você não deve se negar alguns dos caminhos mais proveitosos da análise de dados. Tente sempre conseguir uma compreensão honesta dos dados, desenvolva teorias significativas para o entendimento mais geral, e não se preocupe com a maneira como atingiu esta compreensão.

Resumo

Neste capítulo, examinamos a lógica fundamental da análise de *survey* — o modelo de elaboração. Embora tivéssemos limitado as análises a tabelas de porcentagens, veremos no capítulo seguinte que esta lógica se aplica igualmente nos casos de análises estatísticas mais complexas.

A lógica da elaboração é a seguinte:

1. Começamos com a observação de uma relação empírica entre duas variáveis, digamos, X e Y.
2. Buscamos entender a natureza dessa relação através do processo de manter constantes outras variáveis. Isso permite comparar a relação original entre X e Y e as *relações parciais* encontradas entre subconjuntos baseados na variável de controle.
3. Se as relações parciais forem essencialmente as mesmas que a original, dizemos que a relação original foi *replicada*, e concluimos que a relação entre X e Y é genuína e generalizável.
4. Se apenas uma das relações parciais é essencialmente a mesma (ou mais forte) que a original, enquanto a outra relação parcial é essencialmente zero, chamamos o resultado de *especificação*: especificaram-se as condições nas quais X causa Y.
5. Se a relação original desaparece nas parciais e se a variável de controle é antecedente (anterior) a X e Y, o resultado é chamado de *explicação*, significando que anulamos uma relação espúria (não-genuína).
6. Se a relação original desaparece nas parciais, e se a variável de controle intervém cronologicamente entre X e Y, este resultado é chamado de *interpretação*, significando que descobrimos os meios pelos quais X causa Y.

Notas

¹ STOUFFER, Samuel A. et al. *The American Soldier*. Princeton, N.J.: Princeton University Press, 1949. v.1. p.122 et seq., esp. p.127.

² KENDALL, Patricia L., LAZARSELD, Paul F. Problems of Social Analysis. In: MERTON, Robert K., LAZARSELD, Paul F. (Ed.). *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"*. New York: Free Press, 1950. p.133-196.

³ Note, é claro, que os termos "variável independente" e "variável dependente" são, a rigor, usados incorretamente no diagrama. De fato, temos uma variável independente (a variável de teste) e duas variáveis dependentes. A terminologia incorreta foi usada apenas para dar continuidade ao exemplo anterior.

⁴ N.T. Medicare: sistema de seguro-saúde. BABBIE, Earl R. *Science and Morality in Medicine*. Berkeley: University of California Press, 1970. Ver especialmente p.181.

⁵ GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. p.92.

⁶ ROSENBERG, Morris. *The Logic of Survey Analysis*. New York: Basic Books, 1968.

⁷ Ibidem. p.88-89.

⁸ Ibidem. p.94-95.

Leituras Adicionais

GLOCK, Charles Y. (Ed.). *Survey Research in the Social Sciences*. New York: Russel Sage Foundation, 1967. cap.I.

HIRSCHI, Travis, SELVIN, Hanan. *Principles of Survey Analysis*. New York: Free Press, 1973.

HYMAN, Herbert. *Survey Design and Analysis*. New York: Free Press, 1955.

LAZARSELD, Paul F., PASANELLA, Ann K., ROSENBERG, Morris (Ed.). *Continuities in "The Language of Social Research"*. New York: Free Press, 1972. seção II.

ROSENBERG, Morris. *The Logic of Survey Analysis*. New York: Basic Books, 1968.

STOUFFER, Samuel A. *Social Research to Test Ideas*. New York: Free Press, 1962.

Capítulo 16

Estatística Social

Muitas pessoas se deixam intimidar pela pesquisa empírica por não ficarem confortáveis com a matemática e a estatística. Muitos relatórios de pesquisa são cheios de cálculos não especificados. O papel da estatística em pesquisa de *survey* é muito importante, mas é igualmente importante ver esse papel na perspectiva adequada. A pesquisa empírica é, antes de mais nada, uma operação lógica, e não uma operação matemática. A matemática é apenas uma linguagem conveniente e eficaz para descrever as operações lógicas inerentes à boa análise de dados. A estatística, um ramo da matemática aplicada, é especialmente adequada para várias análises de pesquisa.

Neste capítulo, examinaremos dois tipos de estatística: a *descritiva* e a *inferencial*. A *estatística descritiva* é um meio de descrever dados em formas manejáveis. A *estatística inferencial*, por seu lado, ajuda a tirar conclusões de observações; tipicamente, envolve tirar conclusões sobre uma população a partir de uma amostra daquela população.

Estatística Descritiva

A estatística descritiva é um método de apresentar descrições quantitativas de modo manejável. Às vezes, deseja-se descrever variáveis isoladamente, outras vezes quer-se descrever as associações que ligam uma variável a outra. Vejamos alguns dos modos de fazer essas descrições.

Redução de Dados

A pesquisa científica freqüentemente envolve coletar grandes massas de dados. Suponha que pesquisamos 2.000 pessoas, fazendo 100 perguntas a cada uma — um estudo de grandeza comum. Com isso, teríamos 200.000 respostas. Ninguém poderia ler todas e extrair delas alguma conclusão significativa. Portanto, muita análise científica implica a *redução* de dados, de detalhes imaneáveis para resumos mais fáceis de trabalhar.

Para começar, vejamos a matriz de dados brutos criados por um projeto de pesquisa quantitativa. A Tabela 16-1 apresenta uma matriz de dados parciais. Cada fileira na matriz representa uma pessoa (ou outra unidade de análise), cada coluna uma variável e cada célula na matriz o atributo ou valor codificado que uma pessoa tem numa variável. A primeira coluna da Tabela 16-1 representa sexo. Suponha que “1” represente homem e “2”, mulher. Isso significa que as pessoas 1 e 2 são homens, a pessoa 3 é mulher, e assim por diante.

No caso de idade, o “3” da pessoa 1 pode significar de 30 a 39 anos de idade, e o “4” da pessoa 2, 40-49. Não importa como a idade tenha sido codificada (ver Capítulo 11), os números de código mostrados na coluna 2 da Tabela 16-1 descreveriam cada pessoa lá representada.

TABELA 16-1
Matriz de dados brutos parciais

	Sexo	Idade	Escolaridade	Renda	Profissão	Filiação Política	Orientação Política	Filiação Religiosa	Importância da religião
Pessoa 1	1	3	2	4	1	2	3	0	4
Pessoa 2	1	4	2	4	4	1	1	1	2
Pessoa 3	2	2	5	5	2	2	4	2	3
Pessoa 4	1	5	4	4	3	2	2	2	4
Pessoa 5	2	3	7	8	6	1	1	5	1
Pessoa 6	2	1	3	3	5	3	5	1	1

Note que os dados já foram reduzidos quando criamos uma matriz de dados como a da Tabela 16-1. Se a idade foi codificada como acima, a resposta específica “33 anos de idade” já foi reduzida à categoria 30-39. As pessoas entrevistadas podem ter dado 60 ou 70 idades diferentes, agora reduzidas a seis ou sete categorias.

O Capítulo 14 discutiu algumas formas de resumir dados univariados: medidas de tendência central, como a moda, a mediana e a média, e medidas de dispersão, como amplitude e desvio padrão. Também é possível resumir a associação entre variáveis.

Medidas de Associação

A associação entre duas variáveis quaisquer pode ser representada por uma matriz de dados produzida pelas distribuições de freqüência conjuntas das duas variáveis. A Tabela 16-2 apresenta uma matriz assim. Ela dá toda a informação necessária para determinar a natureza e a extensão da relação entre escolaridade e preconceito.

TABELA 16-2
Dados brutos hipotéticos sobre escolaridade e preconceito

Preconceito	Nível de Escolaridade				
	Nenhum	Primário	Secundário	Superior	Pós-Graduado
Alto	23	34	156	67	16
Médio	11	21	123	102	23
Baixo	6	12	95	164	77

Note, por exemplo, que 23 pessoas (a) não têm nenhuma escolaridade e (b) tiveram um escore alto de preconceito; 77 pessoas (a) tinham curso de pós-graduação e (b) tiveram baixo escore de preconceito. Como a matriz de dados brutos na Tabela 16-1, esta matriz dá mais informação do que o necessário à compreensão. Ao examiná-la com cuidado, vê-se que, à medida que a escolaridade aumenta de “nenhum” para “pós-graduado”, há uma tendência geral de diminuição do preconceito, mas não é possível ter nada mais além dessa impressão geral. Várias estatísticas descritivas permitem resumir esta matriz de dados. A seleção da medida adequada depende, inicialmente, da natureza das duas variáveis.

Vejamos algumas alternativas para resumir a associação entre duas variáveis. Se quiser mais informação sobre esses tópicos, consulte um manual de estatística social.¹ As medidas de associação que vamos discutir baseiam-se no modelo de *redução proporcional de erro* (RPE). Para ver como esse modelo

funciona, suponha que eu lhe peça para adivinhar atributos de respondentes em alguma variável, por exemplo, se responderam sim ou não a algum item do questionário.

Primeiro, suponha que você conhece a distribuição geral de respostas na amostra total, digamos, 60% sim e 40% não. Você cometeria o menor número de erros se sempre usasse a resposta *modal* (a mais freqüente): o sim.

Segundo, suponha que você também conhece a relação empírica entre a primeira variável e alguma outra, digamos, sexo. Toda vez que eu perguntar se algum respondente respondeu sim ou não, lhe direi também se é homem ou mulher. Se as duas variáveis são relacionadas, você deve cometer menos erros da segunda vez. É possível, portanto, calcular o RPE sabendo a relação entre as duas variáveis: quanto maior a relação, maior a redução de erro.

O modelo básico de RPE é modificado ligeiramente para levar em conta níveis diferentes de medida: nominal, ordinal ou intervalo. As seções seguintes consideram cada nível de medição e apresentam uma medida de associação adequada a cada uma. As medidas discutidas são apenas três de muitas medidas apropriadas.

Variáveis Nominais. Se as duas variáveis consistem em dados nominais (por exemplo, sexo, filiação religiosa, raça), o *lambda* (λ) seria uma medida adequada. O *lambda* se baseia na capacidade de prever valores de uma das variáveis, isto é, o RPE obtido do conhecimento dos valores de outra variável. Vejamos um exemplo hipotético simples da lógica e método do *lambda*. A Tabela 16-3 apresenta dados hipotéticos relacionando sexo a situação empregatícia. No geral, vemos que 1.100 pessoas estão empregadas e 900 desempregadas. Se você fosse predizer se as pessoas estão empregadas sabendo apenas a distribuição geral dessa variável, predizer sempre “empregado” levaria a menos erros do que sempre predizer “desempregado”. Mas isso resultaria em 900 erros em 2.000 previsões.

TABELA 16-3

Dados hipotéticos relacionando sexo e emprego

	Homens	Mulheres	Total
Empregados	900	200	1.100
Desempregados	100	800	900
Total	1.000	1.000	2.000

Suponha que você teve acesso aos dados da Tabela 16-3 e ouvisse o sexo da pessoa antes de predizer a situação de emprego. A estratégia mudaria. Para cada homem a predição seria “empregado” e para cada mulher seria “desempregada”. Neste caso, você faria 300 erros — os 100 homens desempregados e as 200 mulheres empregadas — ou 600 erros menos do que cometeria se não soubesse o sexo da pessoa.

Lambda, então, representa a redução de erros como proporção dos erros que teriam sido cometidos com base na distribuição geral. Neste exemplo, *lambda* seria igual a 0,67, ou seja, os 600 erros a menos divididos pelos 900 erros totais baseados apenas no emprego. Neste exemplo, *lambda* mede a associação estatística entre sexo e emprego.

Se sexo e emprego fossem estatisticamente independentes, encontraríamos a mesma distribuição de situação de emprego para homens e mulheres. Neste caso, saber o sexo dos respondentes não afetaria o número de erros cometidos ao se predizer situação de emprego, e o *lambda* resultante seria zero. Se, por outro lado, todos os homens estivessem empregados e todas as mulheres desempregadas, saber o sexo da pessoa evitaria erros na predição de situação de emprego. Especificamente, você cometeria 900 erros a menos (em 900), de forma que *lambda* seria 1,0, representando uma associação estatística perfeita.

Lambda é apenas uma de várias medidas de associação para a análise de duas variáveis nominais.

Variáveis Ordinais. Se as variáveis relacionadas são ordinais (por exemplo, classe social, religiosidade, alienação), *gamma* (γ) é uma medida de associação adequada. Como *lambda*, *gamma* se baseia na capacidade de adivinhar valores de uma variável sabendo valores de outra. Em vez de valores exatos, todavia, *gamma* se baseia no arranjo ordinal de valores. Para qualquer *par* de casos, você adivinha que a ordenação de uma variável vai corresponder, positiva ou negativamente, à ordenação da outra. Por exemplo, se você suspeita que religiosidade se relaciona positivamente com conservadorismo político e se a Pessoa A é mais religiosa que a Pessoa B, você adivinha que ela é também mais conservadora que B. *Gamma* se baseia no número de comparações emparelhadas que se ajustam a este padrão versus as que o contradizem.

TABELA 16-4
Dados hipotéticos relacionando classe social e preconceito

Preconceito	Classe Baixa	Classe Média	Classe Alta
Baixo	200	400	700
Médio	500	900	400
Alto	800	300	100

A Tabela 16-4 apresenta dados hipotéticos relacionando classe social e preconceito. A natureza geral da relação entre essas duas variáveis é que, à medida que aumenta a classe social, diminui o preconceito. Há uma associação negativa entre classe social e preconceito.

Gamma é calculado de duas quantidades: (1) o número de pares que têm a mesma posição de ordem nas duas variáveis e (2) o número de pares com posições de ordem opostas nas duas variáveis. O número de pares com a mesma posição é calculado multiplicando-se a freqüência de cada célula na tabela pela soma de todas as células que aparecem abaixo e à direita dela e depois somando-se todos os produtos. Na Tabela 16-4, o número de pares com a mesma posição seria calculado da seguinte forma: $200(900 + 300 + 400 + 100) + 500(300 + 100) + 400(400 + 100) + 900(100)$ ou $340.000 + 200.000 + 200.000 + 90.000 = 830.000$.

O número de pares com posições opostas nas duas variáveis é calculado multiplicando-se a freqüência de cada célula na tabela pela soma de todas as células que aparecem abaixo e à esquerda dela e depois somando-se todos os produtos. Na Tabela 16-4, o número de pares com posições opostas seria calculado assim: $700(500 + 800 + 900 + 300) + 400(800 + 300) + 400(500 + 800) + 900(800)$ ou $1.750.000 + 440.000 + 520.000 + 720.000 = 3.430.000$.

Gamma é calculado a partir dos números de pares com a mesma posição e de pares com posições opostas, da seguinte maneira:

$$\text{Gamma} = \frac{\text{igual - oposto}}{\text{igual} + \text{oposto}}$$

Neste exemplo, gamma é igual a $(830.000 - 3.430.000)$ dividido por $(830.000 + 3.430.000)$, ou $-0,61$. O sinal negativo indica a associação negativa sugerida pela inspeção inicial da

tabela. Classe social e preconceito, neste exemplo hipotético, são associados negativamente. A expressão numérica de gamma indica que 61% a mais dos pares examinados tinham posições opostas do que posições iguais.

Note que, enquanto os valores de lambda variam de 0 a 1, os valores de gamma variam de -1 a $+1$, representando a direção tanto quanto a magnitude da associação. Como as variáveis nominais não têm estrutura ordinal, não faz sentido falar da direção da relação. (Um lambda negativo indicaria que você cometeu mais erros na predição dos valores de uma variável sabendo os valores da segunda do que os que cometeu ignorando a segunda; isso não é logicamente possível.)

Considere o seguinte exemplo do uso de gamma na pesquisa de *survey* contemporânea. Para estudar até que ponto as viúvas santificavam seus maridos falecidos, Helena Znaniecki Lopata aplicou um questionário a uma amostra probabilística de 301 viúvas na área de Chicago.² Em parte, o questionário pedia a cada uma caracterizar seu marido falecido em termos da *escala de diferenciação semântica* mostrada na Tabela 16-5. Pediu-se a cada uma descrever o cônjuge falecido fazendo um círculo num número para cada par de características opostas. Note que a série de números conectando cada par de características é uma medida ordinal.

Depois, Lopata queria descobrir até que ponto as várias medidas se relacionavam uma com a outra e escolheu gamma como medida da associação. A Tabela 16-6 mostra como ela apresentou os resultados. O formato mostrado é chamado de *matriz de correlação*. Para cada par de medidas, Lopata calculou o gamma. "Bom" e "Útil", por exemplo, se relacionam um com o outro num gamma igual a 0,79. A matriz é uma forma conveniente de apresentar as intercorrelações entre várias variáveis e é encontrada com freqüência na literatura de pesquisa. Neste caso, vê-se que todas as variáveis se relacionam fortemente umas com as outras, embora para alguns pares a relação seja mais forte do que para outros.

Gamma é apenas uma de várias medidas de associação para variáveis ordinais.

TABELA 16-5
Escala de diferenciação semântica

		Característica								
Extremo Positivo									Extremo Negativo	
Bom	1	2	3	4	5	6	7	Mau		
Útil	1	2	3	4	5	6	7	Inútil		
Honesto	1	2	3	4	5	6	7	Desonesto		
Superior	1	2	3	4	5	6	7	Inferior		
Gentil	1	2	3	4	5	6	7	Cruel		
Amigável	1	2	3	4	5	6	7	Hostil		
Caloroso	1	2	3	4	5	6	7	Frio		

TABELA 16-6
Associações gamma entre os itens de diferenciação semântica da escala de santificação

	Útil	Honesto	Superior	Gentil	Amigável	Caloroso
Bom	0,79	0,88	0,80	0,90	0,79	0,83
Útil	—	0,84	0,71	0,77	0,68	0,72
Honesto		—	0,83	0,89	0,79	0,82
Superior			—	0,78	0,60	0,73
Gentil				—	0,88	0,90
Amigável					—	0,90

FONTE - LOPATA, Helena Znaniecki. Widowhood and Husband Sanctification. *Journal of Marriage and the Family*, p.439-450, maio 1991.

Variáveis de Intervalo ou de Razão. Se estão sendo associadas variáveis de intervalo ou de razão (por exemplo, idade, renda, média de notas escolares), uma medida adequada de associação é a *correlação produto-momento* de Pearson (r). A derivação e a computação desta medida de associação são por demais complexos para os propósitos deste livro, portanto faremos apenas alguns comentários gerais.

Como gamma e lambda, r se baseia em adivinhar o valor de uma variável conhecendo-se a outra. Para variáveis de intervalo ou de razão contínuas, entretanto, é pouco provável que se possa adivinhar o valor *preciso* da variável. Por outro lado, apenas predizer o arranjo ordinal de valores para as duas variáveis deixaria de tirar vantagem do maior teor de informação transmitido por uma variável de intervalo ou razão. Em certo sentido, r reflete o *quão exata* foi a adivinhação do valor de uma variável, com base no conhecimento do valor da outra.

Para entender a lógica de r , considere como você pode, hipoteticamente, adivinhar os valores das unidades de análise numa variável qualquer. Com variáveis nominais, vimos que se pode sempre apostar no valor modal. Para dados de razão ou intervalo, você sempre minimiza os erros apostando no valor médio da variável. Embora esta prática produza nenhuma ou poucas estimativas perfeitas, a extensão dos erros será minimizada.

No cálculo de lambda, vimos o número de erros produzidos por apostar sempre no valor modal. No caso de r , os erros são medidos em termos da soma do quadrado das diferenças entre o valor real e a média. Esta soma é chamada de *variação total*. Para entender este conceito, devemos expandir o âmbito de nossa investigação. Retornaremos a esta discussão quando examinarmos a *análise de regressão*, no Capítulo 17. Por enquanto, mudaremos nosso foco da estatística descritiva para a estatística inferencial.

Estatística Inferencial

A maioria dos *surveys* examina dados coletados de amostras extraídas de populações maiores. Uma amostra de pessoas pode ser entrevistada num *survey*; uma amostra de registros de divórcio pode ser codificada e analisada; uma amostra de jornais pode ser examinada através de análise de conteúdo. Os pesquisadores raramente estudam amostras apenas para descrevê-las *per se*; na maioria dos casos, o objetivo final é fazer afirmações sobre a população maior da qual a amostra foi extraída. Frequentemente, portanto, será necessário tratar os achados univariados e multivariados amostrais como a base para *inferências* a respeito de alguma população.

Esta seção examina as medidas estatísticas usadas para fazer tais inferências e suas bases lógicas. Começaremos com

dados univariados e passaremos para os multivariados na discussão de testes de significância estatística.

Inferências Univariadas

As primeiras seções do Capítulo 14 trataram dos métodos de apresentação de dados univariados. Cada medida de resumo pretendia oferecer um método para descrever a amostra estudada. Usaremos essas medidas para fazer afirmativas mais amplas sobre a população.

Se 50% de uma amostra de pessoas dizem que tiveram resfriados no ano anterior, 50% é também nossa melhor estimativa da proporção de resfriados na população total de onde foi retirada a amostra. (Esta estimativa supõe uma amostra aleatória simples, naturalmente.) É improvável, entretanto, que *precisamente* 50% da população tiveram resfriados durante o ano. Se foi seguido um desenho de amostragem rigoroso para seleção aleatória, seremos capazes de estimar a faixa de erro esperada quando aplicarmos o achado amostral à população.

O Capítulo 5, sobre a teoria da amostragem, tratou dos procedimentos para fazer tais estimativas, portanto faremos aqui apenas uma revisão. No caso de uma porcentagem, a quantidade

$$\sqrt{\frac{p \times q}{n}}$$

em que p é uma porcentagem, q é igual a $1 - p$ e n é o tamanho da amostra, chama-se *erro padrão*. Esta quantidade é muito importante na estimativa do erro amostral. Podemos ter 68% de confiança que o número para a população cairá dentro de mais ou menos um erro padrão do número para a amostra, 95% de confiança que cairá dentro de mais ou menos dois erros padrão ou 99,9% de confiança que cairá dentro de mais ou menos três erros padrão.

Qualquer afirmação sobre erro amostral, portanto, deve ter dois componentes essenciais: o *nível de confiança* (por exemplo, 95%) e o *intervalo de confiança* (por exemplo, $\pm 2,5\%$). Se 50% de uma amostra de 1.600 pessoas disseram ter tido resfriados durante o ano, podemos dizer que temos 95% de confiança que o número para a população vai estar entre 47,5% e 52,5%.

Neste exemplo, fomos além de apenas descrever a amostra, para fazer estimativas (inferências) sobre a população maior. Ao fazê-lo, devemos estar atentos às implicações de várias suposições.

Primeiro, a amostra deve ser tirada da população sobre a qual se está inferindo. Uma amostra tirada de uma lista telefônica não é base legítima para inferências estatísticas sobre a população de uma cidade, porque algumas pessoas não têm telefone e outras têm números não listados.

Segundo, a estatística inferencial supõe o uso de amostragem aleatória simples, o que virtualmente nunca é o caso em *surveys* por amostragem. A estatística supõe amostragem com reposição, que quase nunca é feita, mas talvez isso não seja um problema sério. Amostragem sistemática é usada mais que a aleatória; também isso não é problema sério, se feita corretamente. Amostragem estratificada, já que aumenta a representatividade, não é, claramente, um problema. Mas a amostragem por conglomerados pode ser problemática, porque as estimativas de erro amostral podem ser pequenas demais. É patente que amostragem de esquina não permite o uso de estatística inferencial. O cálculo do erro padrão também supõe que 100% da amostra será aproveitada; o problema aumenta de gravidade à medida que diminui esta taxa.

Terceiro, a estatística inferencial trata apenas do erro de amostragem. Não considera "erros não-amostrais". Assim, embora possa ser correto afirmar que entre 47,5% e 52,5% da população (com 95% de confiança) relataram ter tido resfriados no ano anterior, não podemos com tanta confiança estimar a porcentagem que realmente os teve. Alguns relatam resfriados quando tiveram outra doença; outros esquecem os resfriados que tiveram. Como os erros não-amostrais são provavelmente maiores que os erros de amostragem num desenho de amostra respeitável, é preciso muito cuidado ao generalizar de achados de amostra para a população.

Testes de Significância Estatística

Não há resposta científica à pergunta se uma associação entre duas variáveis é significativa, forte, importante, interessante ou digna de relato. Talvez o teste decisivo seja sua capacidade de persuadir sua audiência (presente e futura) da significância da associação. Mas os *testes de significância* podem ajudar nesse ponto. Como o nome sugere, a estatística paramétrica é a estatística que faz certas suposições sobre os parâmetros que descrevem a população da qual se extraiu a amostra.

Embora os *testes de significância estatística* sejam amplamente citados na literatura de ciências sociais, a lógica subjacente a eles é sutil e muitas vezes mal-entendida. Os testes de significância apóiam-se na mesma lógica de amostragem discutida neste livro. Para ajudar a entender essa lógica, retornemos ao conceito de erro amostral com relação a dados univariados.

Lembre que uma estatística amostral normalmente fornece a melhor estimativa do parâmetro de população correspondente, mas a estatística e o parâmetro raramente coincidem com exatidão. Desse modo, relatamos a probabilidade de que o parâmetro caia dentro de uma certa faixa (intervalo de confiança). O grau de incerteza dentro dessa faixa se deve a erro de amostragem normal. O corolário desta afirmação de probabilidade é o de que é *improvável* que o parâmetro caia fora da faixa especificada puramente como resultado de erro de amostragem. Assim, se estimamos que um parâmetro (com 99,9% de confiança) está entre 45% e 55%, dizemos, por implicação, que é *extremamente improvável* que o parâmetro seja na verdade, por exemplo, 90%, supondo que nosso único erro de estimativa se deva a amostragem normal. Essa é a lógica que fundamenta os testes de significância.

Uma Ilustração Gráfica. Vamos ilustrar esta lógica da significância estatística numa série de diagramas representando a seleção de amostras de uma população. Os elementos da lógica a ser ilustrada são:

1. Suposições com respeito à *independência* de duas variáveis no estudo da população.
2. Suposições com respeito à *representatividade* das amostras selecionadas através de procedimentos convencionais de amostragem probabilística.
3. A *distribuição conjunta* observada dos elementos da amostra em termos das duas variáveis.

A Figura 16-1 representa uma população hipotética de 256 pessoas, metade homens e metade mulheres. O diagrama indica como cada pessoa se sente sobre as mulheres gozarem de igualdade com os homens. No diagrama, os que favorecem a igualdade são representados por círculos abertos, enquanto os que se opõem a ela são representados por círculos sombreados.

A questão investigada é se há relação entre gênero e sentimentos sobre igualdade entre homens e mulheres. Mais especificamente, se a tendência das mulheres favorecerem a igualdade é maior do que a dos homens, já que, presumivelmente, elas se beneficiariam mais dessa igualdade. Examine a Figura 16-1 e tente ver a resposta a essa pergunta.

FIGURA 16-1

Uma população hipotética de homens e mulheres que são favoráveis ou se opõem à igualdade sexual



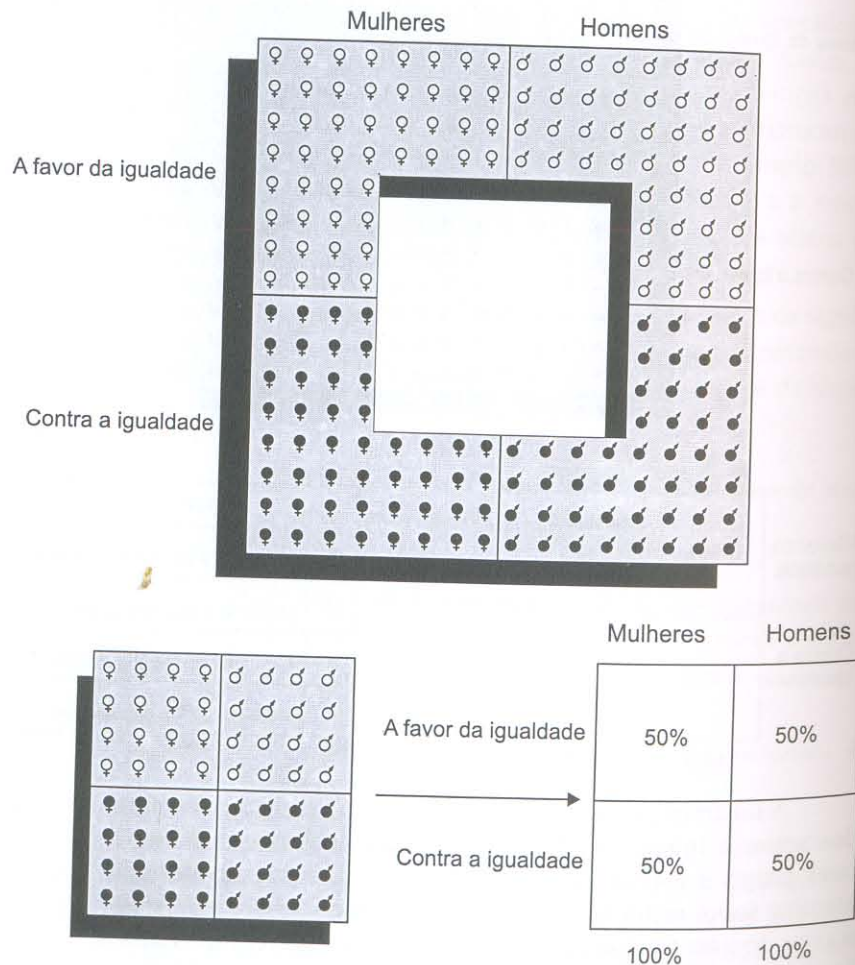
	Mulheres	Homens
A favor da igualdade	50%	50%
Contra a igualdade	50%	50%
	100%	100%

Legenda	
♀	Mulheres a favor da igualdade
♂	Homens a favor da igualdade
♀	Mulheres contra a igualdade
♂	Homens contra a igualdade

A ilustração indica que não há relação entre sexo e atitudes sobre a igualdade. Exatamente metade dos membros de cada grupo é a favor e metade é contra. Lembre a discussão anterior sobre redução proporcional de erro. Neste caso, saber o sexo de uma pessoa não reduziria os erros que cometeríamos ao tentar adivinhar a atitude dela com relação à igualdade. A tabela na parte inferior da Figura 16-1 fornece uma visão tabular do que se observa no diagrama.

A Figura 16-2 representa a seleção de uma amostra de um quarto da população hipotética. Na ilustração, um “quadrado” recortado no centro da população provê uma amostra representativa. Essa amostra tem dezesseis de cada tipo de pessoa. Metade são homens e metade são mulheres, e metade de cada grupo é a favor da igualdade, enquanto a outra metade se opõe a ela.

FIGURA 16-2
Uma amostra representativa



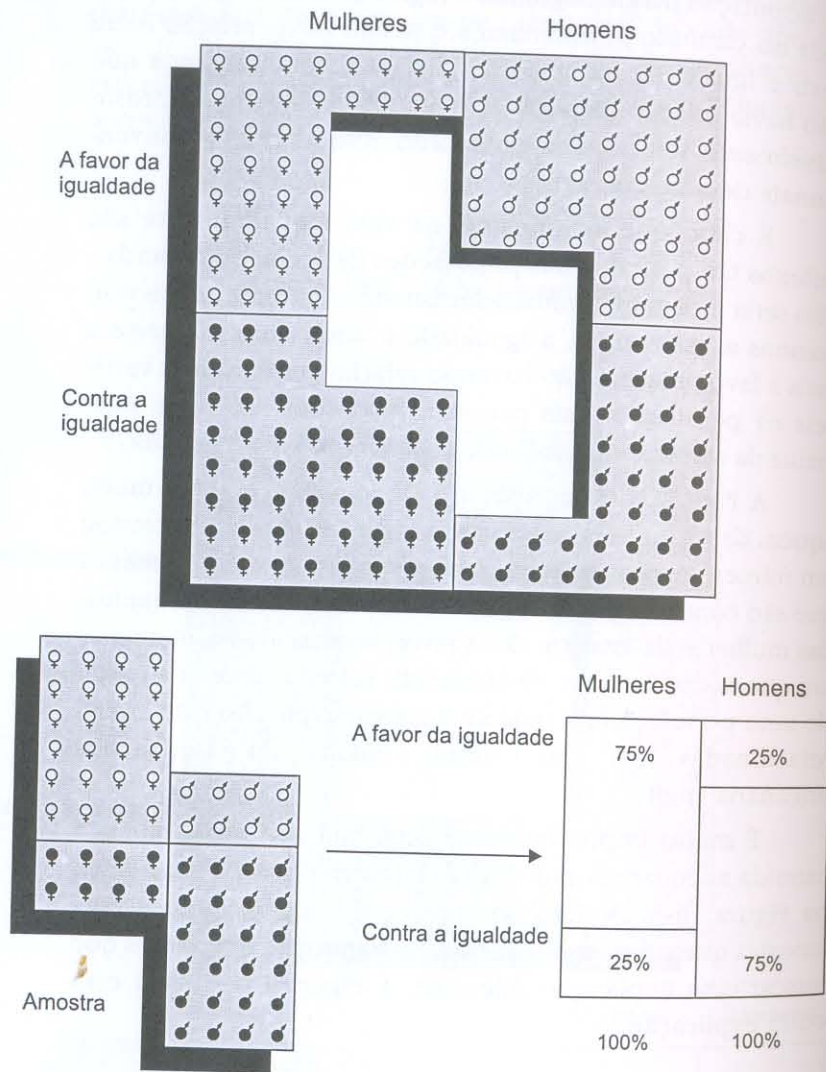
A amostra selecionada na Figura 16-2 permitiria tirar conclusões precisas sobre a relação entre sexo e igualdade na população maior. Seguindo a lógica de amostragem aprendida no Capítulo 5, notaríamos que não havia relação entre sexo e igualdade na amostra; portanto, concluiríamos que não havia relação na população maior (a amostra foi provavelmente selecionada de acordo com as regras convencionais de amostragem).

É claro que as amostras na vida real raramente são reflexos tão perfeitos das populações de onde foram tiradas. Não seria incomum termos selecionado, digamos, um ou dois homens a mais contra a igualdade e umas duas mulheres a mais a favor, mesmo não havendo relação entre as duas variáveis na população. Tais pequenas variações são parte integrante da amostra probabilística, conforme se viu no Capítulo 5.

A Figura 16-3 representa uma amostra que fica muito aquém da representação da população maior. Ela selecionou um número excessivo de mulheres que apóiam e de homens que são contra a igualdade. Como mostra a tabela, três quartos das mulheres da amostra são a favor, comparados com apenas um quarto dos homens. Se tivéssemos selecionado esta amostra de uma população na qual as duas variáveis não estivessem relacionadas uma com a outra, a análise da amostra nos enganaria muito.

É muito improvável que uma amostra probabilística extraída adequadamente seja tão imprecisa quanto a mostrada na Figura 16-3. De fato, se realmente selecionássemos uma amostra que desse os resultados da Figura 16-3, teríamos que buscar uma explicação diferente. A Figura 16-4 ilustra essa outra explicação.

FIGURA 16-3
Uma amostra não representativa

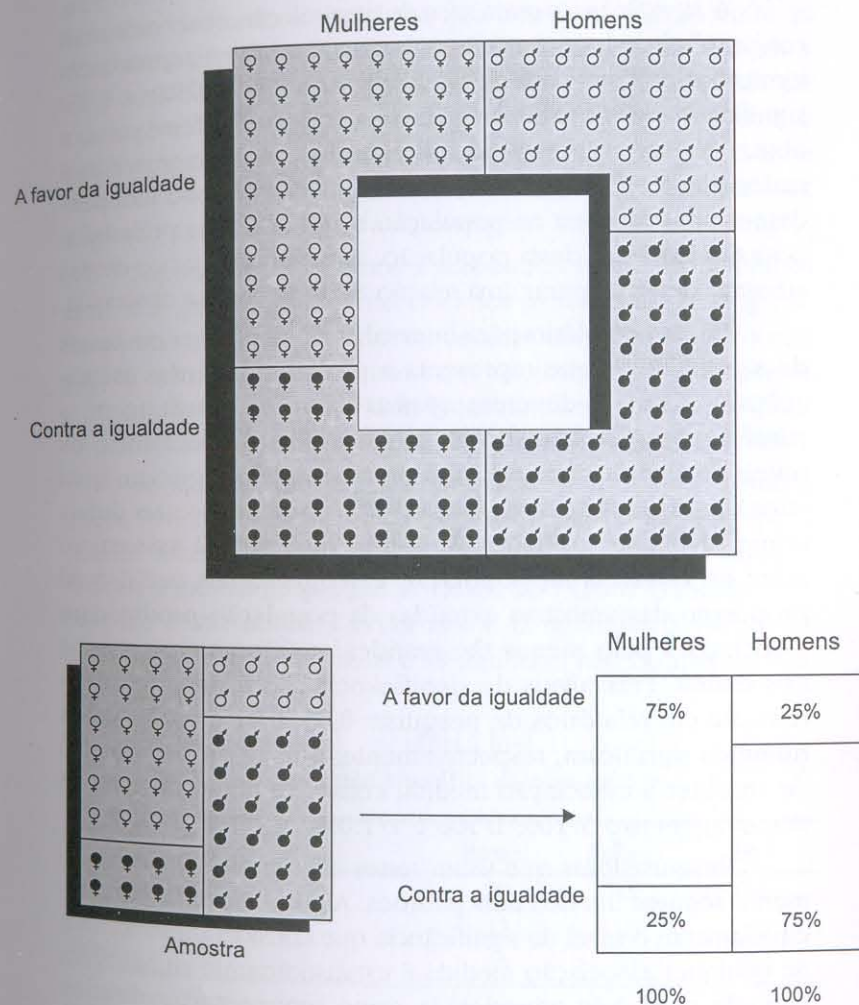


Note que a amostra selecionada na Figura 16-4 também mostra uma forte relação entre gênero e igualdade, mas desta vez a razão é diferente. Selecionamos uma amostra perfeitamente representativa, mas há de fato uma forte relação entre as duas variáveis na população. Nesta figura, há, na população, maior probabilidade das mulheres apoiarem a igualdade do que os homens, e a amostra reflete isso.

Na prática, nunca sabemos o que é verdadeiro para a população total; é por isso que extraímos amostras. Se selecionássemos uma amostra e descobríssemos a forte relação mostrada

nas Figuras 16-3 e 16-4, precisaríamos decidir se nosso achado refletiu a população de modo preciso ou foi apenas um resultado de erro de amostragem.

FIGURA 16-4
Uma amostra representativa de uma população na qual as variáveis se relacionam



A Lógica da Significância Estatística. A lógica fundamental dos testes de significação estatística, portanto, é a seguinte. Face a qualquer discrepância entre a suposta independência de variáveis numa população e a distribuição observada de elementos da amostra, podemos explicar a discrepância de um de dois modos: (1) podemos atribuí-la a uma amostra não representativa, ou

(2) podemos rejeitar a suposição de independência. A lógica e a estatística associadas com métodos de amostragem probabilística orientam sobre as probabilidades variadas de graus variados de não-representatividade (expressos como erro de amostragem). Ou seja, há uma *alta* probabilidade de um *pequeno* grau de não-representatividade e uma *baixa* probabilidade de um *grande* grau de não-representatividade.

A *significância estatística* de uma relação observada num conjunto de dados amostrais, portanto, é sempre expressa em termos de probabilidades. Significativo no nível 0,05 ($p < 0,05$) significa que a probabilidade de uma relação tão forte como a observada ser atribuível apenas a erro de amostragem não é maior que 5 em 100. Ou seja, se duas variáveis são independentes uma da outra na população e 100 amostras probabilísticas são extraídas desta população, não mais que cinco dessas amostras devem mostrar uma relação tão forte como a observada.

Há um corolário para intervalos de confiança em testes de significância, que representa a probabilidade das associações medidas se deverem *apenas* a erro de amostragem: o *nível de significância*. Como os intervalos de confiança, os níveis de significância derivam de um modelo lógico no qual várias amostras são extraídas de uma população. Ao determinar níveis de confiança, supomos que não há associação entre as variáveis na população e perguntamos então qual proporção das amostras extraídas da população produziriam associações pelo menos tão grandes quanto as encontradas nos dados. Três níveis de significância são usados frequentemente em relatórios de pesquisa: 0,05, 0,01 e 0,001. Esses números significam, respectivamente, que as probabilidades de se obter a associação medida como resultado de erro de amostragem são 5/100, 1/100 e 1/1.000.

Pesquisadores que usam testes de significância normalmente seguem um dos dois padrões. Alguns especificam antecipadamente o nível de significância que consideram suficiente. Se qualquer associação medida é estatisticamente significativa naquele nível, vão considerá-la como representação de uma associação genuína entre as duas variáveis. Em outras palavras, estão dispostos a descontinuar a possibilidade de ela resultar de erro de amostragem.

Outros pesquisadores preferem informar o nível específico de significância para cada associação, desconsiderando as convenções de 0,05, 0,01 e 0,001. Em vez de relatar que uma certa associação é significativa no nível 0,05, podem relatar a

significação no nível 0,023, indicando que as chances de ela ter resultado de erro de amostragem são 23 em 1000.

Qui-Quadrado. Qui-quadrado (χ^2) é um teste de significância muito usado nas ciências sociais. Baseia-se na *hipótese nula*, que é a suposição de não haver relação entre duas variáveis na população total. Dada a distribuição observada de valores nas duas variáveis separadas, computamos a distribuição conjunta que seria esperada se não houvesse relação entre as duas variáveis. O resultado desta computação é um conjunto de *frequências esperadas* para todas as células na tabela de contingência. Comparamos esta distribuição esperada com a distribuição de casos realmente encontrada nos dados da amostra e determinamos a probabilidade de a discrepância descoberta ter resultado apenas de erro de amostragem. Vejamos um exemplo.

Suponha que estamos interessados na possível relação entre ir à igreja e gênero, no caso dos membros de uma igreja. Para testar a relação, selecionamos aleatoriamente uma amostra de 100 membros. Descobrimos que na amostra 40 são homens e 60, mulheres, e que 70% da amostra afirmam ter ido à igreja na semana anterior, os restantes 30% dizendo que não o fizeram.

Se não houver relação entre sexo e comparecimento à igreja, 70% dos homens na amostra deveriam ter ido à igreja na semana anterior e 30%, não. As mulheres deveriam ter comparecido na mesma proporção. A Tabela 16-7 (Parte I) mostra que, baseado neste modelo, 28 homens e 42 mulheres deveriam ter ido à igreja, com 12 homens e 18 mulheres não comparecendo.

TABELA 16-7
Ilustração hipotética de qui-quadrado

I. Frequências Esperadas	Homens	Mulheres	Total
Compareceram	28	42	70
Não compareceram	12	18	30
Total	40	60	100
II. Frequências Observadas	Homens	Mulheres	Total
Compareceram	20	50	70
Não compareceram	20	10	30
Total	40	60	100
III. (Observadas - Esperadas) ² ÷ Esperadas	Homens	Mulheres	
Compareceram	2.29	1.52	$\chi^2=12.70$
Não compareceram	5.33	3.56	$p<.001$

A Parte II da Tabela 16-7 mostra o comparecimento observado à igreja para a amostra de 100 membros. Observe que 20 homens afirmaram ter ido à igreja na semana anterior, enquanto outros 20 disseram que não foram. Das mulheres da amostra, 50 foram e 10 não. Comparando as frequências esperadas e observadas (Partes I e II), observamos que menos homens foram à igreja do que o esperado, no caso das mulheres ocorrendo o contrário.

Qui-quadrado é calculado assim. Para cada célula na tabela, o pesquisador (1) subtrai a frequência esperada para aquela célula da frequência observada, (2) eleva esta quantidade ao quadrado e (3) divide a diferença ao quadrado pela frequência esperada. Faz-se o mesmo para cada célula na tabela, e os resultados são agregados. (A Parte III da Tabela 16-7 apresenta as computações célula por célula.) O somatório final é o valor do qui-quadrado — 12,70 no exemplo atual.

Este valor é a discrepância geral entre a distribuição conjunta observada na amostra e a distribuição que devíamos esperar se as duas variáveis não se relacionassem entre si. Evidentemente, a simples descoberta de uma discrepância não prova que duas variáveis se relacionam. Um erro normal de amostragem pode produzir discrepâncias mesmo quando não há relação na população. Mas a grandeza do valor de qui-quadrado permite estimar a probabilidade disto haver acontecido.

Para determinar a significância estatística da relação observada, usamos um conjunto padrão de valores de qui-quadrado. Isto requer a computação dos *graus de liberdade*, que, para o qui-quadrado, são computados assim. O número de linhas na tabela de frequências observadas, menos um, é multiplicado pelo número de colunas, menos um, que pode ser escrito como $(l - 1)(c - 1)$. No exemplo, temos duas linhas e duas colunas (descontados os totais), portanto há um grau de liberdade.

Numa tabela de valores de qui-quadrado, vemos que, para um grau de liberdade e amostragem aleatória de uma população em que não há relação entre duas variáveis, 10% das vezes devemos esperar um qui-quadrado de pelo menos 2,7. Assim, se selecionarmos 100 amostras desta população, devemos esperar que cerca de 10 delas produzam um qui-quadrado igual ou maior que 2,7. Além disto, devemos esperar valores mínimos de qui-quadrado de pelo menos 6,6 em apenas 1% das amostras e valores de pelo menos 7,9 em apenas 0,5%

das amostras. Quanto maior o valor do qui-quadrado, menor a probabilidade de este valor poder ser atribuído apenas a erro de amostragem.

No nosso exemplo, o valor do qui-quadrado é 12,70. Se não houvesse relação entre sexo e comparecimento à igreja na população de membros, e tivesse sido selecionado e estudado um grande número de amostras, esperaríamos um qui-quadrado desta magnitude em menos da metade de 1% das amostras. A probabilidade de obter um qui-quadrado desta magnitude é menos de 0,001%, supondo o uso de amostragem aleatória e não haver relação na população. Relatamos este achado afirmando que a relação é estatisticamente significativa no nível 0,001. Já que é tão improvável que a relação observada possa ter resultado apenas do erro de amostragem, rejeitamos a hipótese nula e supomos haver uma relação entre as duas variáveis na população de membros daquela Igreja.

A maioria das medidas de associação pode ser testada para significância estatística. Tabelas padronizadas de valores nos permitem determinar se uma associação tem significância estatística e em que nível. Qualquer manual de estatística traz instruções para o uso destas tabelas, de modo que não há necessidade de o fazermos aqui.

Revisão. Testes de significância fornecem uma medida objetiva em relação às quais podemos estimar a significância de associações entre variáveis, que nos ajudam a excluir associações que possam não representar associações genuínas na população estudada. Quem usar ou ler relatórios de testes de significância estatística deve, contudo, se precaver contra vários perigos na interpretação.

Primeiro, estamos discutindo testes de significância *estatística*; não há testes objetivos de significância substantiva. Assim, podemos estar legitimamente convencidos que uma certa associação não se deve a erro de amostragem, mas também podemos afirmar sem medo de contradição que duas variáveis têm uma relação mínima entre si. Erro de amostragem é uma função inversa ao tamanho da amostra: quanto maior a amostra, menor o erro esperado. Uma correlação de 0,1 pode ser significativa (em algum nível) se for descoberta numa amostra grande, mas a mesma correlação entre as mesmas variáveis não seria significativa numa amostra menor. Isto faz sentido se entendermos a lógica básica dos testes de significância. Na amostra maior, há menos chance de a correlação ser apenas o produto de erro de amostragem. Nas duas amostras,

porém, a correlação poderia representar uma correlação muito fraca, de essencialmente zero.

A distinção entre significância estatística e substantiva talvez seja melhor ilustrada em casos de certeza absoluta de que diferenças observadas não podem resultar de erros de amostragem. Este é o caso quando observamos uma população toda. Imagine que consigamos saber a idade de todas as autoridades governamentais nos EUA e na URSS. Só para argumentar, suponha que a média de idade das autoridades americanas fosse 45 e a das autoridades soviéticas 46. Não há erro de amostragem, pois sabemos a idade de todos. Sabemos com certeza que as autoridades soviéticas são mais velhas que as americanas. Obviamente diremos que a diferença não é substantivamente significativa. Concluiremos que, nos dois países, as autoridades têm essencialmente a mesma idade.

Segundo, a menos que sejamos enganados por este exemplo hipotético, *não* devemos calcular significância estatística de relações observadas em dados colhidos de populações inteiras. Testes de significância estatística medem a probabilidade de relações entre variáveis serem produtos apenas de erro de amostragem. Se não há amostragem, não pode haver erro de amostragem.

Terceiro, testes de significância se baseiam nas mesmas suposições de amostragem usados na computação dos intervalos de confiança. Se tais suposições não se enquadram no desenho da amostra, testes de significância não têm legitimidade.

Como acontece na maioria dos temas deste livro, também neste ponto tenho meu viés pessoal, neste caso, contra testes de significância. Não objeto à lógica estatística destes testes, que é sólida. Mas preocupa-me que estes testes parecem enganar mais que iluminar. Minhas reservas principais são:

1. Testes de significância fazem suposições de amostragem que virtualmente nunca são satisfeitas por desenhos de amostragens reais.
2. Dependem da ausência de erros não-amostrais, suposição questionável na maioria das medições empíricas.
3. Na prática, são aplicados com demasiada frequência a medidas de associação cuja computação violou suposições

destas medidas (por exemplo, correlações produto-momento computadas a partir de dados ordinais).

4. Significância estatística é facilmente mal entendida como "força de associação" ou significância substantiva.

Ao mesmo tempo — talvez paradoxalmente — acho que testes de significância podem ser valiosos para o pesquisador como instrumentos úteis para a compreensão de dados. Apesar de os comentários acima sugerirem uma abordagem conservadora dos testes de significância, isto é, que só devem ser usados quando todas as suposições são satisfeitas, minha perspectiva geral é justamente a inversa. Encorajo o uso de qualquer técnica estatística (qualquer medida de associação ou teste de significância) em qualquer conjunto de dados, se isto ajudar a compreender os dados. Se calcular correlações produto-momento entre variáveis nominais e testar significância estatística no contexto de amostragens não controladas satisfizerem este critério, então encorajo estas atividades. Digo isto no espírito do que Hanan Selvin chamou de técnicas para garimpar dados. Vale tudo, contanto que, ao cabo, haja compreensão dos dados e do mundo social sob estudo.

Porém, o preço desta liberdade radical é abdicar de interpretações estatísticas estritas. Você não poderá mais basear a importância definitiva de seus achados apenas numa correlação significativa no nível 0,05. Qualquer que seja a rota da descoberta, os dados empíricos têm de ser apresentados de forma legítima e sua importância deve ser argumentada logicamente.

Resumo

Neste capítulo, trabalhamos com uma visão ampla e abrangente do mundo das estatísticas sociais. Não pretendi fazer do leitor um especialista em estatísticas, mas entender a lógica básica de algumas técnicas estatísticas mais usadas. Começamos fazendo uma distinção ampla entre estatísticas *descritivas* e *inferenciais*.

Estatísticas descritivas são usadas para resumir dados. Algumas resumem a distribuição de atributos numa só variável; outras, chamadas *medidas de associação*, resumem a associação entre variáveis.

Estatísticas inferenciais são usadas para estimar a generalizabilidade de achados obtidos pela análise de amostras às populações das quais foram extraídas. Algumas estatísticas inferenciais estimam as características da população para uma variável (em termos de *níveis e intervalos de confiança*). Outras, chamadas *testes de significância estatística*, estimam as relações entre variáveis da população.

Muitas medidas de associação se baseiam num modelo de *redução proporcional de erro* (RPE), baseado na comparação de (a) o número de erros que cometeríamos se tentássemos adivinhar os atributos de uma variável para cada caso estudado, se tudo o que soubéssemos fosse a distribuição dos atributos desta variável, e (b) o número de erros que cometeríamos se soubéssemos a distribuição conjunta geral e fôssemos informados em cada caso sobre o atributo de uma variável, cada vez que procurássemos adivinhar o atributo da outra.

Lambda é uma medida de associação para a análise de duas variáveis *nominais*. Oferece uma ilustração límpida do modelo RPE. Gamma é uma medida de associação para a análise de duas variáveis *ordinais*. A *correlação produto-momento* de Pearson é uma medida de associação para a análise de duas variáveis *intervalo* ou *de razão*.

Inferências sobre alguma característica de uma população, tal como o percentual de eleitores a favor do candidato A, devem indicar o *intervalo de confiança* (âmbito dentro do qual o valor é esperado, por exemplo, 45% a 55% a favor do candidato A) e uma indicação do *nível de confiança* (a probabilidade de que o valor esteja dentro deste âmbito, por exemplo, 95% de confiança). Cálculos de níveis e intervalos de confiança se baseiam na teoria da probabilidade e supõem técnicas convencionais de amostragem probabilística.

Inferências sobre a generalizabilidade para a população das associações descobertas entre variáveis numa amostra envolvem *testes de significância estatística*, que estimam a probabilidade de uma associação tão forte quanto a observada resultar de erro de amostragem normal, caso não haja associação entre as variáveis na população maior. Portanto, testes de significância estatística também se baseiam na teoria da probabilidade e supõem técnicas convencionais de amostragem probabilística. Significância estatística não deve ser confundida com significância *substantiva*; esta quer dizer que uma associação observada é forte, importante, significativa.

O nível de significância de uma associação observada é relatado como a probabilidade de ela ter resultado apenas de

erro de amostragem. Afirmar que uma associação é significativa no nível 0,05 é dizer que uma associação tão forte quanto a observada não poderia ser esperada como resultado de erro de amostragem mais do que 5 vezes em 100. Pesquisadores de *survey* tendem a usar um leque particular de níveis de significância em conexão com testes de significância estatística: 0,05, 0,01 e 0,001. Isto é apenas uma convenção.

Falando estritamente, testes de significância estatística fazem suposições sobre dados e métodos quase nunca completamente satisfeitos por *surveys* reais. Apesar disto, os testes podem ser úteis na análise e interpretação dos dados. Mas tome cuidado para não interpretar a "significância" do teste precisamente demais.

Notas

¹ Dois livros manuais de estatística social que podem ser úteis: JENDREK, Margaret Platt. *Through the Maze: Statistics with Computer Applications*. Belmont, CA: Wadsworth, 1985; BLALOCK, Hubert. *Social Statistics*. New York: McGraw-Hill, 1979.

² LOPATA, Helena Znaniecki. *Widowhood and Husband Sanctification*. *Journal of Marriage and the Family*, p.439-450, maio 1981.

Leituras Adicionais

BLALOCK, Hubert. *Social Statistics*. New York: McGraw-Hill, 1979.

HENKEL, Ramon E. *Tests of Significance*. Beverly Hills, CA: Sage, 1976.

JENDREK, Margaret Platt. *Through the Maze: Statistics with Computer Applications*. Belmont, CA: Wadsworth, 1985.

KISH, Leslie. *Chance, Statistics and Statisticians*. *Journal of the American Statistical Association*, v.73, n.361, p.1-6 mar. 1978.

MORRISON, Denton, HENKEL, Ramon (Ed.). *The Significance Test Controversy: a Reader*. Chicago: Aldine-Atherton, 1970.

SHARP, Vicki. *Statistics for the Social Sciences*. Boston: Little, Brown, 1979.

Capítulo 17

Técnicas Multivariadas Avançadas

A maior parte deste livro até agora focalizou formas um tanto rudimentares de manipulação de dados, como o uso de tabelas de contingência percentuadas. O modelo de análise de elaboração foi apresentado sob a forma de tabelas de contingência, assim como as técnicas estatísticas descritas no Capítulo 16.

Iremos agora um passo além e consideraremos brevemente alguns métodos mais complexos de análise e apresentação de dados. Onde for possível, cada técnica examinada será apresentada sob a perspectiva lógica do modelo de elaboração. Examinaremos as seguintes técnicas: análise de regressão, análise de trajetória, análise fatorial, análise de variância, análise discriminante e modelos log-lineares. Estas são apenas algumas de muitas técnicas disponíveis para o pesquisador de *survey*.

Meu propósito é *introduzi-lo* a estas técnicas. Você não concluirá o capítulo proficiente no uso delas, mas as reconhecerá ao encontrá-las em relatório de pesquisas. Se quiser aprender mais sobre elas, serão dadas algumas referências bibliográficas.

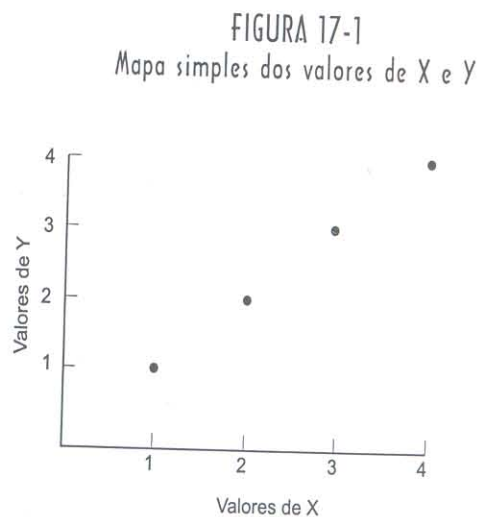
Análise de Regressão

Referi-me diversas vezes à fórmula geral para descrever a associação entre duas variáveis: $Y = f(X)$. A fórmula se lê "Y é uma função de X", significando que valores de Y podem ser explicados em termos de variações nos valores de X. Afirmando

de forma mais forte, podemos dizer que X causa Y , isto é, o valor de X determina o valor de Y . *Análise de Regressão* é um método para determinar a função específica relacionando Y com X . Há várias formas de análise de regressão, dependendo da complexidade das relações estudadas. Começemos com a mais simples: a *regressão linear*.

Regressão Linear

O modelo de regressão pode ser visto com clareza máxima no caso da associação linear perfeita entre duas variáveis. A Figura 17-1 mostra graficamente os valores conjuntos de X e Y num estudo hipotético. Mostra que, nos quatro casos, os valores de X e Y são idênticos. A relação entre as duas variáveis é descrita pela equação $X = Y$, chamada *equação de regressão*. Como os quatro pontos estão numa linha reta, podemos sobrepor-lhes esta linha. É a *linha de regressão*.

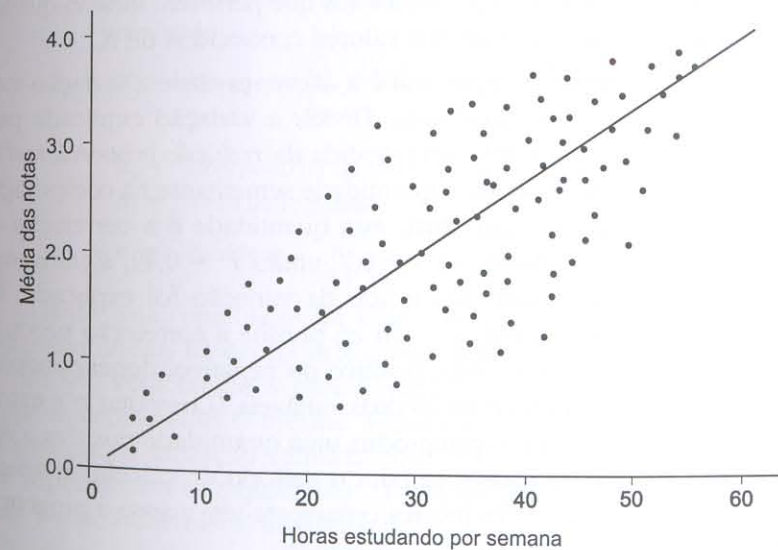


O modelo de regressão linear tem importantes usos descritivos. A linha de regressão mostra graficamente a associação entre X e Y , e a equação de regressão é uma forma eficiente de resumir esta associação. O modelo de regressão tem também valor inferencial. Na medida em que a equação de regressão descreve corretamente a associação *geral* entre as duas variáveis, pode ser usada para prever outros conjuntos de valores. Se soubermos, por exemplo, que um novo caso tem o valor 3,5 para X , podemos prever um valor de 3,5 também para Y . É claro

que, na prática, as pesquisas raramente se limitam a quatro casos, e as associações entre variáveis raramente são tão claras quanto a da Figura 17-1.

Um exemplo mais realista é o da Figura 17-2, que representa uma relação hipotética entre número de horas de estudo e notas de estudantes. Como no exemplo anterior, os valores de Y (notas) geralmente correspondem aos de X (horas de estudo); aumentando os valores de X , aumentam os de Y . Mas esta associação não é tão clara quanto a da Figura 17-1. Na Figura 17-2 não é possível sobrepor uma linha reta passando por todos os pontos do gráfico, mas podemos traçar uma linha aproximada, mostrando a melhor representação linear possível dos vários pontos. Esta linha foi traçada no gráfico.

FIGURA 17-2
Gráfico dos valores de duas variáveis com linha de regressão (hipotética)



Quem estudou geometria sabe que qualquer linha reta num gráfico pode ser representada por uma equação na forma $Y = a + bX$, onde X e Y são valores das duas variáveis. Nesta equação, a é igual ao valor de Y quando X é 0 e b representa a *inclinação* da linha. Se soubermos os valores de a e de b , podemos calcular uma estimativa de Y para cada valor de X .

Análise de regressão é uma técnica para estabelecer a equação de regressão representando a linha geométrica que

Regressão Múltipla

Freqüentemente, pesquisadores de *survey* descobrem que uma variável dependente é afetada simultaneamente por diversas variáveis independentes. A análise de *regressão múltipla* é um meio para analisar estas situações. Beverly Yerg usou esta análise ao estudar a eficácia de professores de educação física, manifestando suas expectativas através da seguinte equação de regressão múltipla:¹

$$F = b_0 + b_1/ + b_2X_1 + b_3X_2 + b_4X_3 + b_5X_4 + e$$

onde

F = Escore final do desempenho do aluno

$/$ = Escore inicial do desempenho do aluno

X_1 = Composto da prática de orientação e apoio

X_2 = Composto do domínio do professor sobre o conteúdo

X_3 = Composto do *feedback* específico relacionado à tarefa

X_4 = Composto da apresentação clara e concisa das tarefas

b = Peso da regressão

e = Resíduo

Note que, em lugar de uma única variável X numa regressão linear, há diversos X e também diversos b em vez de um só. Yerg também optou por representar a como b_0 nesta equação, com o mesmo significado discutido acima. Finalmente, a equação termina com um fator residual (e), que representa a raiz quadrada da variância em Y não explicada pelas variáveis X analisadas.

Começando com esta equação, Yerg calculou os valores de diversos b , para mostrar as contribuições relativas das várias variáveis independentes na determinação dos escores finais dos desempenhos estudantis. Ela também calculou o coeficiente de correlações múltiplas como indicador do quanto as seis variáveis previram os escores finais. Este procedimento segue a mesma lógica da correlação bivariada simples discutida antes e é tradicionalmente escrito como R maiúsculo. Neste caso, $R = 0,877$, significando que 77% da variância (R^2) nos escores finais se explicam pelas seis variáveis agindo em conjunto.

mais se aproxima da distribuição de pontos. Esta equação é valiosa tanto descritiva quanto inferencialmente. Primeiro, a equação de regressão faz uma descrição matemática da relação entre as variáveis. Segundo, a equação de regressão permite inferir os valores de Y quando temos os valores de X . Voltando à Figura 17-2, podemos estimar as notas dos alunos se soubermos quantas horas eles estudam por semana.

Para melhorar a exatidão das estimativas, construímos uma *linha de regressão*, na forma de uma equação de regressão, que permita estimar valores de uma variável a partir dos valores da outra. O formato geral desta equação é $Y' = a + b(X)$, onde a e b são valores computados, X é um valor numa variável e Y' é o valor *estimado* na outra. Os valores de a e de b são computados para minimizar as diferenças entre valores reais de Y e as estimativas correspondentes (Y') baseadas no valor conhecido de X . O somatório das diferenças ao quadrado entre valores reais e os estimados de Y é chamado *variação não explicada*, porque representa erros que persistem mesmo quando estimativas se baseiam nos valores conhecidos de X .

A *variação explicada* é a diferença entre a variação total e a variação não explicada. Dividir a variação explicada pela variação total produz uma medida da *redução proporcional do erro*, correspondendo a quantidade semelhante na computação do lambda. No caso atual, esta quantidade é a correlação ao quadrado, r^2 . Assim, se $r = 0,7$, então $r^2 = 0,49$, significando que aproximadamente metade da variação foi explicada. Na prática, computamos r e não r^2 , porque a correlação produto-momento pode ter sinal positivo ou negativo, dependendo da direção da relação entre as duas variáveis. (Computar r^2 e extrair a raiz quadrada sempre produz uma quantidade positiva.) Veja um manual de estatística para o método de calcular r , apesar de que a maioria dos leitores certamente tem acesso a programas de informática projetados para fazê-lo.

Esta discussão pode ter sido exageradamente estatística, mas a vida social é tão complexa que o modelo de regressão linear simples não basta para representar o verdadeiro estado das coisas. Como vimos no Capítulo 14, é possível, usando tabelas de porcentagem, analisar mais de duas variáveis. Aumentando o número de variáveis, estas tabelas se tornam mais complicadas e difíceis de ler. O modelo de regressão é uma boa alternativa para tais casos.

Regressão Parcial

Ao explicar o modelo de elaboração, demos atenção especial à relação entre duas variáveis, quando uma terceira variável teste se mantém constante. Assim, podemos examinar o efeito de educação sobre preconceito com idade se mantendo constante, testando o efeito independente de educação. Para isto, computamos a relação tabular entre educação e preconceito separadamente para cada grupo de idade.

Regressões parciais se baseiam no mesmo modelo lógico. A equação resumindo a relação entre variáveis é computada com base nas variáveis teste permanecendo constante. Como no caso do modelo de elaboração, o resultado pode então ser comparado com a relação não controlada entre as duas variáveis, para esclarecer mais a relação total.

Regressão Curvilínea

Até agora discutimos a associação entre variáveis representada por uma linha reta, embora em mais de duas dimensões. O modelo de regressão é ainda mais geral que nossa discussão até agora implica.

Se você conhece geometria, sabe que funções curvilíneas podem ser representadas por equações. Por exemplo, a equação $X^2 + Y^2 = 25$ descreve um círculo com raio 5. Elevar variáveis a grandezas maiores que 1 tem como efeito produzir curvas e não linhas retas. Na pesquisa empírica, não há razão para supor que a relação entre todas as variáveis será linear. Em alguns casos, a análise de regressão curvilínea pode propiciar melhor compreensão das relações empíricas que qualquer modelo linear.

Entretanto, lembre que uma linha de regressão tem duas funções. Descreve um conjunto de observações empíricas e oferece um modelo *geral* para fazer inferências sobre a relação entre duas variáveis na população representada pelas observações. Uma equação muito complexa produz uma linha errática passando por cada ponto individual. Neste sentido, descrevendo perfeitamente as observações empíricas. Mas sem garantia de que esta linha prediria adequadamente novas observações ou que representaria significativamente a relação entre duas variáveis. Assim, teria pouco ou nenhum valor inferencial.

Já discutimos antes a necessidade de equilibrar detalhe e utilidade na redução de dados. Pesquisadores procuram oferecer a representação mais fiel e mais simples de seus dados. Isto

se aplica também à análise de regressão. Dados devem ser apresentados da forma mais simples que melhor descreva os dados reais (portanto, regressões lineares são usadas mais freqüentemente). Análise de *regressão curvilínea* oferece ao pesquisador uma nova opção, mas sem resolver de todo o problema. Nada o faz.

Precauções com a Análise de Regressão

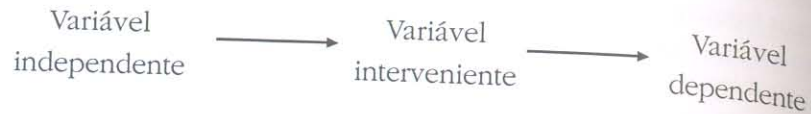
O uso da análise de regressão para inferências estatísticas se baseia nos mesmos supostos feitos da análise de correlação: amostragem aleatória simples, ausência de erros não amostrais e dados de intervalos contínuos. Já que a pesquisa científica social raramente satisfaz completamente estes supostos, deve-se ter cuidado na avaliação dos resultados nas análises de regressão.

Também as linhas de regressão — lineares ou curvilíneas — podem ser úteis para *interpolação* (estimando casos no meio dos observados), mas são menos confiáveis quando usadas para *extrapolação* (estimando casos além da faixa de observações). Reconhecer esta limitação das extrapolações é importante por dois motivos. Primeiro, você provavelmente encontrará equações de regressão que parecem fazer predições ilógicas. Uma equação ligando horas de estudo e médias escolares, por exemplo, parece sugerir que alunos que estudem 100 horas por semana devem tirar nota 15, quando o máximo é 10. Esta falha na capacidade preditiva não desqualifica a equação, mas dramatiza o fato de que sua aplicabilidade é limitada a uma certa faixa de valores. Segundo, pesquisadores por vezes vão além desta limitação, fazendo inferências fora da faixa de observações, merecendo críticas por agir assim.

Análise de Trajetória

Análise de trajetória é um modelo *causal* para entender relações entre variáveis. Baseia-se na análise de regressão, mas, comparada com outros meios, pode oferecer um quadro graficamente mais útil das relações entre diversas variáveis. Análise de trajetória supõe que os valores de uma variável são causados pelos valores de outra, de modo que é essencial distinguir entre variáveis independentes e dependentes. Evidentemente, este requisito não se aplica apenas à análise de trajetória, mas ela provê um modo original de exibir os resultados explicativos para interpretação.

Lembremos como o modelo de elaboração foi representado no Capítulo 15. A lógica da interpretação pode ser assim diagramada:



A lógica desta apresentação é que uma variável independente tem um impacto sobre uma variável interveniente que, por sua vez, tem um impacto sobre uma variável dependente. A análise de trajetória constrói padrões semelhantes de relações entre variáveis, mas o diagrama típico de trajetória contém muito mais variáveis do que as mostradas acima.

Além de diagramar uma rede de relações entre as variáveis, a análise de trajetória mostra as forças destas relações, calculadas a partir de uma análise de regressão que produz números análogos às relações parciais no modelo de elaboração. Estes *coeficientes de trajetória*, ou pesos beta como são também conhecidos, representam as forças das relações entre pares de variáveis, mantendo constantes os efeitos das demais variáveis no modelo.

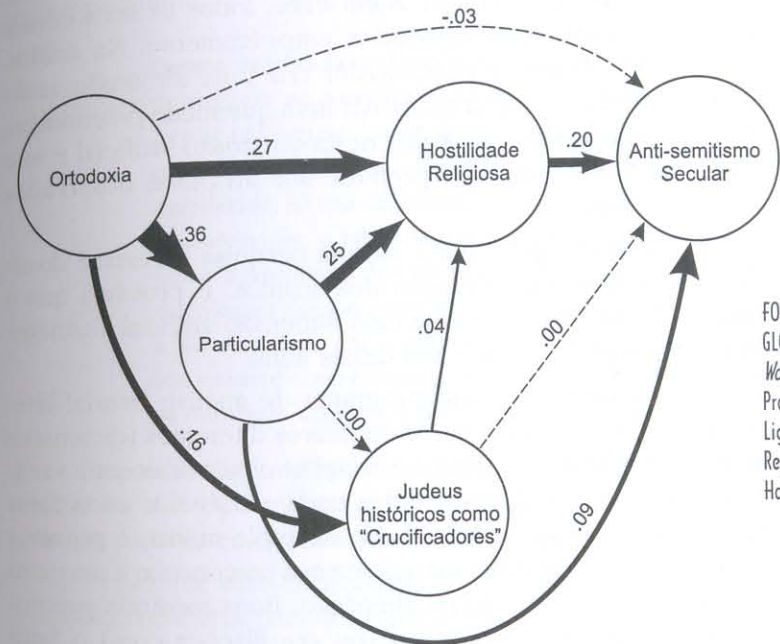
Por exemplo, a análise na Figura 17-3 focaliza as causas religiosas do anti-semitismo entre membros de igrejas cristãs. Da esquerda para a direita, as variáveis no diagrama são (1) ortodoxia, ou até que ponto as pessoas aceitam crenças convencionais sobre Deus, Jesus, os milagres bíblicos etc.; (2) particularismo, a crença de que a "única fé verdadeira" é a da pessoa; (3) aceitação do ponto de vista que os judeus crucificaram Jesus; (4) hostilidade religiosa para com os judeus contemporâneos, por exemplo acreditar que Deus os está castigando ou que serão condenados a menos que se convertam ao Cristianismo; e (5) anti-semitismo secular, como acreditar que judeus são desonestos nos negócios, desleais a seu país etc. De início, os pesquisadores que fizeram esta análise propuseram que o anti-semitismo secular era produzido movendo-se assim através das cinco variáveis: ortodoxia causava particularismo, que causava a visão dos judeus históricos como crucificadores, que causava hostilidade religiosa contra os judeus contemporâneos, que resultava finalmente em anti-semitismo secular.

O diagrama de trajetória conta uma história diferente. Por exemplo, os pesquisadores encontraram que a crença no papel histórico dos judeus como crucificadores de Jesus parece sem importância no processo. E, embora particularismo pareça ser parte do processo que resulta em anti-semitismo secular,

o diagrama também mostra que anti-semitismo é atingido mais diretamente através da ortodoxia e da hostilidade religiosa. Ortodoxia produz hostilidade religiosa mesmo sem particularismo e a hostilidade religiosa gera hostilidade secular.

FIGURA 17-3

Diagrama das fontes religiosas do anti-semitismo



FORNE - STARK, Rodney, FOSTER GLOCK, Charles Y., QUINLEY, Harold. *Wayward Shepherds - Prejudice of Protestant Clergy*. Copyright © Liga B'nai Brith Antidifamação. Reproduzido com permissão da Harper & Row.

Um último comentário sobre a análise de trajetória. Embora seja um modo ótimo de lidar com complexas cadeias causais e redes de variáveis, deve-se entender que a análise de trajetória por si mesma não mostra a ordem causal das variáveis. Também não foi gerada por computador. O pesquisador decidiu a estrutura de relações entre as variáveis e usou análise por computador para calcular os coeficientes de trajetória que se aplicam à estrutura decidida.

Análise Fatorial

Análise fatorial é uma abordagem da análise multivariada diferente da análise de regressão. Sua base estatística é tão mais complexa e tão diferente das discussões anteriores que pede

uma discussão geral. Análise fatorial é usada para descobrir padrões de variações nos valores de diversas variáveis, essencialmente pela geração de dimensões artificiais (fatores) que se correlacionam altamente com diversas das variáveis reais. Um computador deve ser usado para esta operação complexa.

Suponha que um arquivo de dados contenha diversos indicadores de preconceito de sujeitos de uma pesquisa. Cada item deve dar alguma indicação de preconceito, mas nenhum será um indicador perfeito. Além disto, todos os itens devem estar altamente correlacionados empiricamente. Na análise fatorial dos dados, o pesquisador cria uma dimensão artificial com alta correlação com cada item que mede preconceito. Cada sujeito recebe um valor nesta dimensão artificial e este valor deve ser um bom preditor dos atributos observados em cada item.

Suponha agora que a mesma pesquisa apresente diversos indicadores de alienação dos sujeitos. É provável que a análise fatorial também gere uma dimensão artificial altamente correlacionada com cada um destes itens.

O resultado de um programa de análise fatorial consiste em colunas representando fatores diferentes (dimensões artificiais) gerados a partir das relações observadas entre variáveis mais as correlações entre cada variável e cada fator (chamadas *cargas fatoriais*). No exemplo acima, é provável que um fator represente mais ou menos preconceito e um outro represente mais ou menos alienação. Itens medindo preconceito teriam altas cargas no (ou correlações com) o fator preconceito e baixas cargas com o fator alienação. Itens medindo alienação mostrariam o padrão oposto.

Na prática, porém, a análise fatorial não ocorre assim. Em vez disso, as variáveis são insumos do programa, resultando numa série de fatores com cargas fatoriais apropriadas. Determina-se então o significado de um determinado fator com base nas variáveis com alta carga nele. A geração de fatores não se refere ao significado das variáveis, apenas às suas associações empíricas. Dois critérios são considerados: (1) um fator deve explicar uma parte relativamente grande da variância encontrada nas variáveis, e (2) cada fator deve ser mais ou menos independente de cada um dos outros fatores. Vejamos um exemplo.

Muitos pesquisadores estudaram o problema da delinquência. Quando ele é visto com profundidade, descobre-se que há muitos tipos diferentes de delinquentes. Morris Forslund fez um *survey* com alunos de colégio numa cidadezinha de Wyoming, para criar uma tipologia da delinquência.² Seu

questionário pedia aos estudantes que dissessem se haviam cometido uma série de atos de delinquência, submetendo depois as respostas a uma análise fatorial, cujos resultados estão na Tabela 17-1. Os diversos atos de delinquência estão listados à esquerda e os números representam as cargas fatoriais nos quatro fatores construídos na análise. Forslund nomeou as quatro dimensões, e eu usei parênteses nos itens de cada fator que o levaram à escolha destes nomes. Forslund assim resumiu seus resultados:

Quatro padrões bem distintos de atos de delinquência se distinguem na amostra total. Em ordem de variância explicada, eles foram nomeados: 1) atos contra propriedade, incluindo vandalismo e furto; 2) incorrigibilidade; 3) uso de drogas e evasão escolar; 4) brigas. É interessante, e talvez surpreendente que vandalismo e furtos apareçam juntos no mesmo fator. Aparentemente, os colegiais que praticam atos contra propriedade tendem a se envolver em vandalismo e furtos. É também interessante observar que drogas, álcool e evasão escolar se encaixam no mesmo fator.³

Depois de determinar este padrão geral, Forslund fez nova análise fatorial separando rapazes e moças. Essencialmente os mesmos padrões emergiram em ambos os casos.

Penso que este exemplo mostra que a análise fatorial é um método eficiente para descobrir padrões predominantes num número grande de variáveis. Em vez de o pesquisador ser obrigado a comparar inúmeras correlações — simples, parciais e múltiplas — para descobrir padrões, pode-se usar a análise fatorial para isto. É um bom exemplo da utilidade dos computadores.

A análise fatorial também apresenta dados numa forma que pode ser interpretada pelo leitor ou pesquisador. O leitor consegue facilmente descobrir altas cargas de variáveis num fator, detectando grupos de variáveis, ou descobrir em que fatores uma variável tem ou não tem carga alta.

A análise fatorial também tem desvantagens. Primeiro, como observamos, fatores são gerados independentemente de qualquer significado substantivo. Frequentemente, pesquisadores encontram fatores produzindo cargas muito altas num grupo de variáveis substantivamente díspares. Por exemplo, podem encontrar que preconceito e religiosidade têm altas cargas positivas num fator, enquanto educação apresenta uma carga negativa igualmente alta. Certamente as três variáveis

são altamente correlacionadas, mas o que o fator representa? Com muita freqüência, pesquisadores inexperientes acabarão por chamar tais fatores de "falta de educação preconceito-religiosa" ou outra coisa semelhante sem sentido.

TABELA 17-1
Análise fatorial: atos de delinqüência (brancos)

Ato de Delinqüência	Atos contra propriedade		Drogas/Evasão	
	Fator I	Incorrigibilidade Fator II	escolar Fator III	Brigas Fator IV
Quebrar lâmpadas de poste	0,669	0,126	0,119	0,167
Quebrar janelas	0,637			
Quebrar cercas, varais de roupa	0,621	0,186	0,186	0,186
Furtar coisas (\$2 a \$50)	0,616	0,187	0,233	0,068
Tirar ar de pneus	0,587	0,243	0,054	0,156
Furtar coisas (acima de \$50)	0,548	-0,017	0,276	0,034
Jogar ovos, lixo etc.	0,526	0,339	-0,023	0,266
Furtar coisas (abaixo de \$2)	0,486	0,393	0,143	0,077
Furtos na escola	0,464	0,232	-0,002	0,027
Usar carro sem pedir ao dono	0,461	0,172	0,080	0,040
Pintar, pichar	0,451	0,237	0,071	0,250
Desobedecer aos pais	0,054	0,642	0,209	0,039
Estragar mesas, muros	0,236	0,550	-0,061	0,021
Xingar para se desferrar	0,134	0,537	0,045	0,100
Desobedecer ao professor, regente	0,240	0,497	0,223	0,195
Desafio aberto aos pais	0,232	0,458	0,305	0,058
Trotes telefônicos	0,373	0,446	0,029	0,13
Fumar maconha	0,054	0,064	0,755	-0,028
Usar outras drogas para se divertir	0,137	0,016	0,669	0,004
Assinar dispensa no boletim	0,246	0,249	0,395	0,189
Beber na ausência dos pais	0,049	0,247	0,358	0,175
Faltar à escola	0,101	0,252	0,319	0,181
Bater em alguém (briga)	0,309	0,088	0,181	0,843
Brigas vale-tudo	0,242	0,266	0,070	0,602
Variância	67,2	13,4	10,9	8,4

Segundo, muitas vezes a análise fatorial é criticada em bases filosóficas. Já comentamos que, para ter legitimidade, uma hipótese deve poder ser desconfirmada. Devemos poder especificar em que condições a hipótese seria desprovada. Não importa quais dados usamos, a análise fatorial produz solução em forma de fatores. Assim, se o pesquisador indagar "há padrões nestas variáveis?", a resposta será sempre sim. Tal não ocorre com todas as análises fatoriais, mas, no uso da análise fatorial para fins exploratórios, os fatores gerados não têm necessariamente significado.

Minha opinião sobre análise fatorial é a mesma que tenho sobre outros modos complexos de análise. Ela pode ser uma ferramenta muito útil para o pesquisador, e seu uso deve ser encorajado sempre que puder ajudar os pesquisadores a compreender um conjunto de dados. Mas não se esqueça que um instrumento é só isto, nunca uma solução mágica.

Análise de Variância

A *análise de variância* aplica a lógica da significância estatística, já discutida antes. Fundamentalmente, os casos estudados são combinados em grupos representando uma variável independente, e o grau de diferenciação entre os grupos é analisado em termos de uma variável dependente. O grau de diferenciação dos grupos é comparado com o padrão de distribuição aleatória: podemos esperar tais diferenças se alocarmos casos aleatoriamente aos diversos grupos?

Examinemos sucinta e rapidamente duas formas comuns de análise de variância, a saber, análise de variância de *mão única* e de *mão dupla*.

Análise de Variância de Mão Única

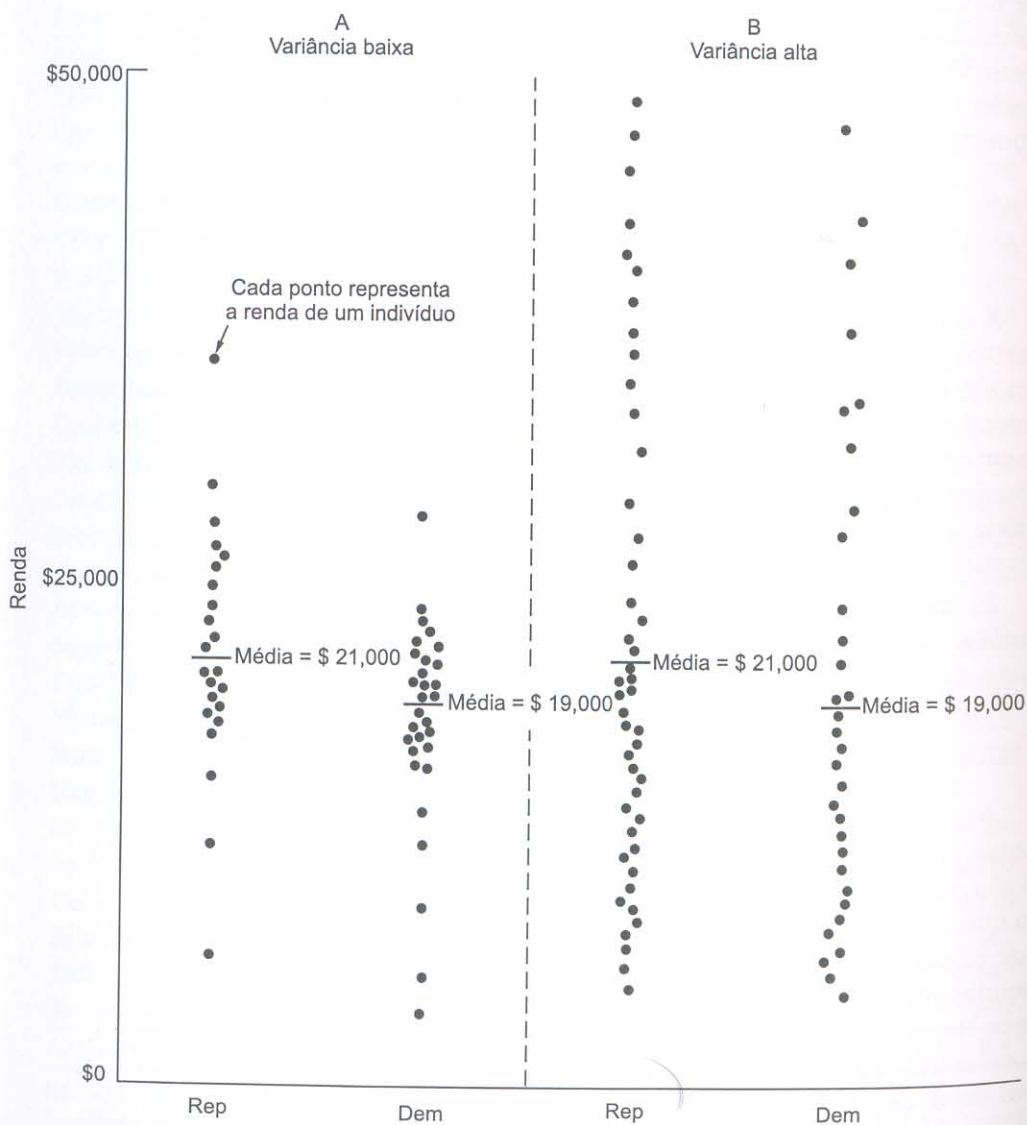
Suponha que queremos comparar níveis de renda de Republicanos e Democratas, para ver se os Republicanos de fato são mais ricos. Seleccionamos uma amostra de indivíduos e no questionário perguntamos (1) com qual partido político se identificam e (2) sua renda total no último ano. Calculamos as rendas médias ou medianas de cada grupo, descobrindo que os Republicanos na amostra têm renda média anual de \$21 mil e os Democratas de \$19 mil. Claramente, os Republicanos da amostra são mais ricos que os Democratas, mas esta

diferença é “significativa”? A diferença de \$2 mil seria provável, se tivéssemos criado dois grupos por seleção aleatória?

A análise de variância responde esta pergunta usando a *variância*. A variância de uma distribuição (de renda, por exemplo) mede em que medida os valores num conjunto de valores estão agrupados próximos da média ou se distribuem muito alto ou muito baixo em relação a ela.

FIGURA 17-4

Dois padrões de distribuição de renda de Republicanos e Democratas



A Figura 17-4 ilustra as duas possibilidades. Veja que, nas duas distribuições, os Republicanos têm renda média de \$21 mil e os Democratas de \$19 mil. Na Parte A, a maioria dos Republicanos tem rendas relativamente perto da média de \$21 mil, e a maioria dos Democratas tem renda perto da média de \$19 mil. Mas a Parte B mostra um quadro diverso. Apesar das médias dos grupos serem as mesmas da Parte A, tanto Republicanos quanto Democratas mostram rendas que variam de muito alta a muito baixa, com considerável justaposição nas distribuições dos partidos. Em termos técnicos, há um grau mais alto de variância na Parte B do que na Parte A.

Os dados na Parte A da Figura 17-4 dão a impressão de uma diferença genuína de renda entre Republicanos e Democratas. Não estaríamos tão certos com os dados da Parte B, no entanto; neste caso, parece mais provável que as variações normais produzidas por erro de amostragem aleatória poderiam ter produzido médias de \$21 mil e \$19 mil.

Na análise de variância, cálculos estatísticos, e não impressões, são usados na tomada de decisões. A diferença observada nas médias é expressa como múltiplos e frações padronizadas da variância observada. Já que a variância na Parte A é menor que na Parte B da Figura 17-4, \$2 mil representam uma diferença maior na Parte A que na Parte B. A diferença resultante nas médias — padronizada pela variância — é então verificada numa tabela estatística padrão mostrando a distribuição teórica destes valores, como vimos na discussão anterior sobre significância estatística. Enfim, concluímos que a diferença é significativa em algum nível de significância. Podemos descobrir, por exemplo, que erros de amostragem teriam produzido uma diferença tão grande quanto esta apenas uma vez em mil. Diríamos então que a diferença era “significativa no patamar de 0,001”.

Simplifiquei os cálculos deste exemplo em favor da lógica básica do procedimento. Na prática, atualmente, tais cálculos são realizados por programas de computador.

O caso mais simples da diferença entre duas médias é muitas vezes chamado de “teste t”. Com mais de dois grupos, os cálculos se tornam mais complexos, porque mais comparações são necessárias. Basicamente, é preciso comparar as médias dos grupos umas com as outras e examinar a variação dos valores de cada grupo. O resultado final da análise, como foi no caso mais simples, é expresso em termos de significância estatística — a probabilidade das diferenças observadas resultarem de erro de amostragem na seleção aleatória.

Análise de Variância de Mão Dupla

A análise de variância de mão única representa uma forma de análise bivariada (partido político e renda foram as duas variáveis do exemplo). Vimos porém que pesquisadores de *survey* muitas vezes fazem análises multivariadas. A análise de variância de mão dupla permite o exame simultâneo de mais de duas variáveis. Suponha, por exemplo, que desconfiamos que as diferenças de renda entre Republicanos e Democratas seja função de educação. Nossa hipótese é que os Republicanos são mais escolarizados que os Democratas, e que pessoas escolarizadas, independentemente de partido, ganham em média mais que pessoas com menos escolaridade. Uma análise de mão dupla faria a triagem dos efeitos das duas variáveis explicativas, de forma semelhante à da elaboração e às das correlações e regressões parciais já discutidas.

Análise Discriminante

Análise discriminante é uma versão interessante de várias técnicas examinadas neste capítulo. Sua lógica se assemelha à da regressão múltipla, exceto que a variável dependente pode ser nominal; lembre que a regressão requer variáveis de intervalo. Vejamos um exemplo.

A Figura 17-5 representa seis escritores, três dos quais escrevem a mão (caneta), enquanto os outros três redigem num computador. (Sei que há outras alternativas, mas este é um exemplo simples.) Nossa tarefa é verificar a diferença no método de redigir. Podemos prever se um determinado escritor usa caneta ou computador?

A Figura 17-6 explora duas variáveis que podemos logicamente esperar ter impacto na forma dos escritores redigirem. *Idade* pode fazer diferença, porque os escritores mais velhos se acostumaram a escrever a mão e podem ter dificuldades de se adaptar à nova tecnologia, enquanto os mais jovens cresceram usando computadores. *Renda* pode fazer diferença, porque computadores são mais caros que canetas. A Figura 17-6 projeta cada escritor no gráfico baseado em *idade* e *renda*. Veja se, a partir do gráfico, você consegue chegar a alguma conclusão que possa explicar a diferença no método de escrever.

A Figura 17-7 esclarece mais a conclusão que você tirou. Só renda parece ser um bom preditor, pelo menos no caso destes

seis escritores. Os que ganham \$30 mil ou menos usam caneta, e os com renda de \$55 mil ou mais usam computador.

No entanto, raramente a vida é tão simples, mesmo em exemplos simplificados. Vamos complicar um pouco as coisas. A Figura 17-8 apresenta os seis escritores numa configuração um pouco mais complicada em termos de idade e renda. Observe que não é possível traçar uma linha separando canetas e computadores, nem por idade nem por renda.

FIGURA 17-5

Seis escritores: três escrevem a mão e três com computadores

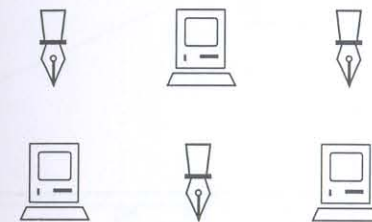


FIGURA 17-6

Projeção gráfica dos seis escritores em termos de idade e renda

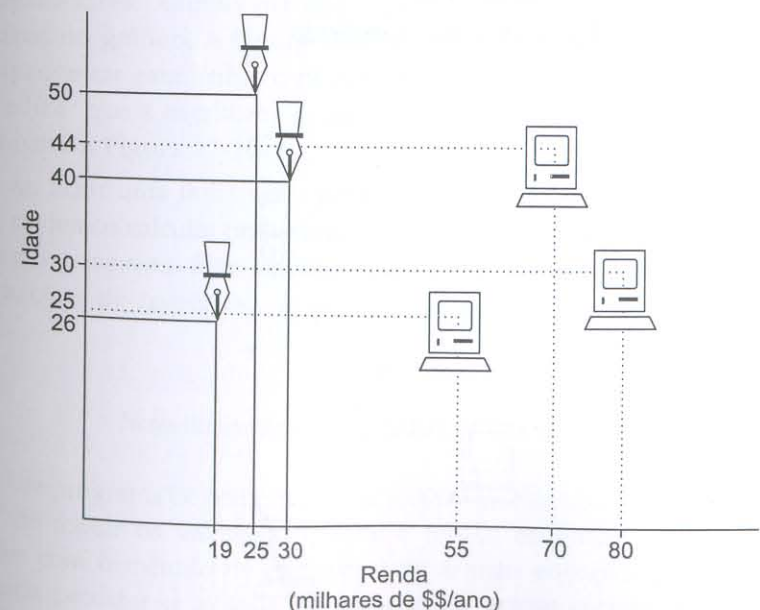


FIGURA 17-7

Apenas renda é suficiente para prever o método de escrever

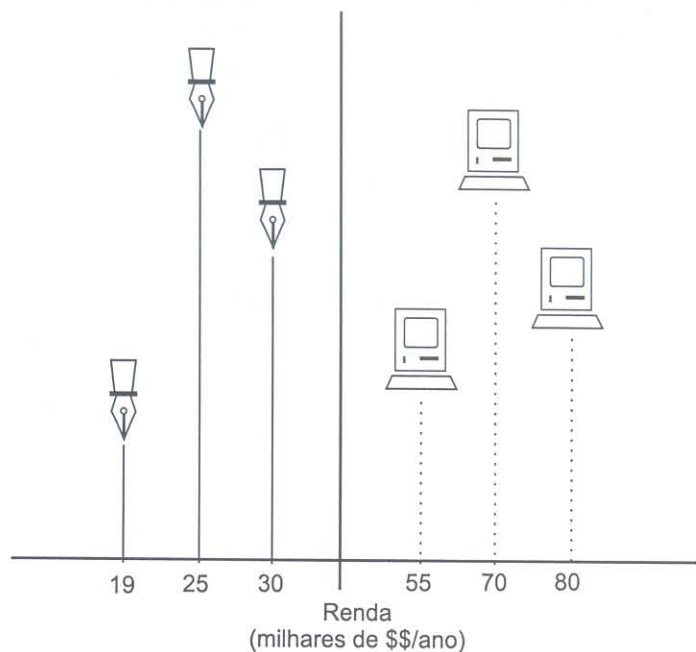


FIGURA 17-8

Padrão ligeiramente mais complicado

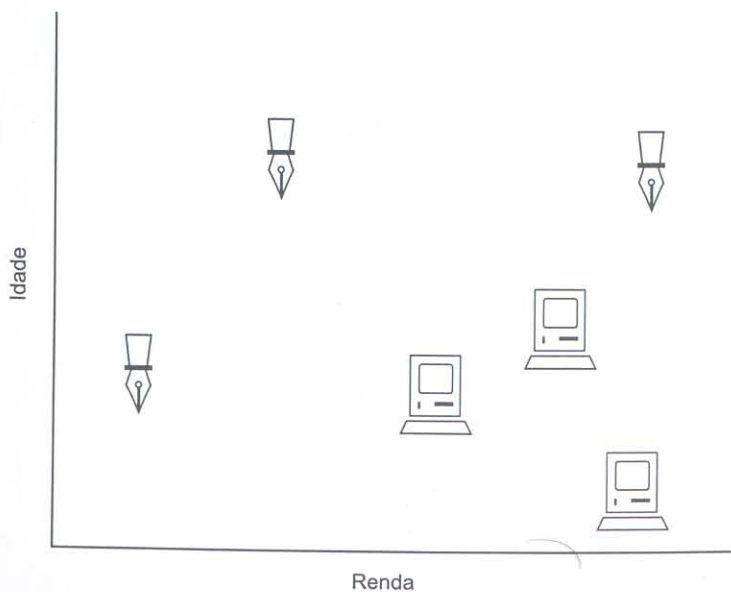
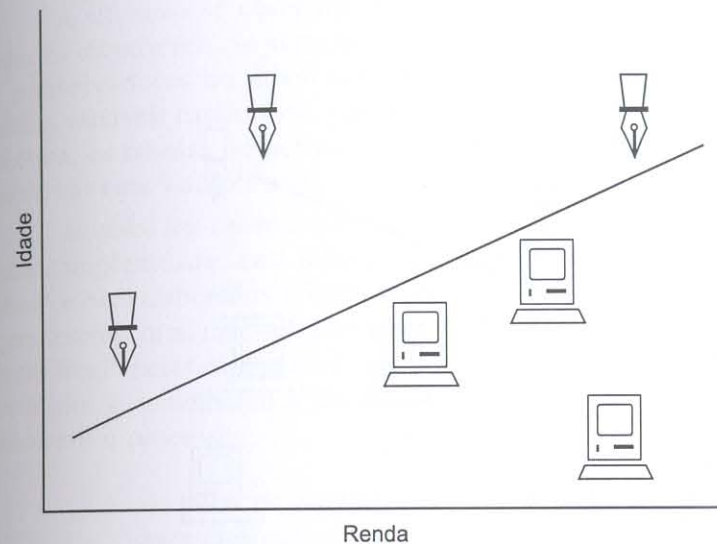


FIGURA 17-9

Separando canetas dos computadores



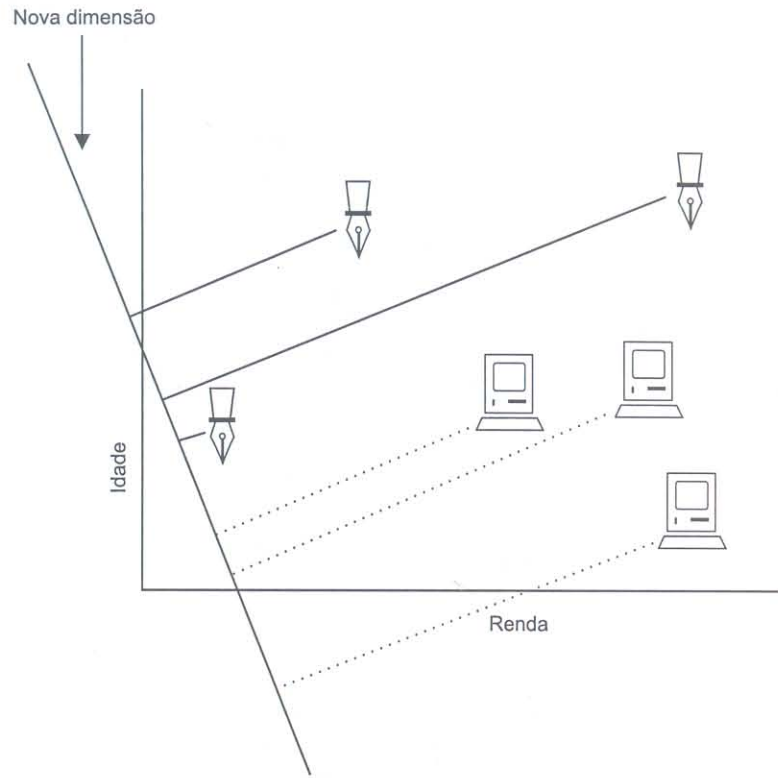
Se estudarmos com mais cuidado a Figura 17-8, descobriremos que *é possível* traçar uma linha separando canetas e computadores. Apenas ela não é perpendicular a nenhum dos eixos no gráfico. A Figura 17-9 mostra a linha que o faz. Para aproveitar esta linha, é necessário encontrar um modo de prever quais escritores ficam de que lado da linha. É o que mostra a Figura 17-10.

Ao fazer uma linha nova perpendicular à linha de separação, podemos calcular onde cada escritor se encaixará na nova dimensão composta. Este cálculo teria uma forma semelhante às equações de regressão. A equação seria mais ou menos assim:

$$\text{Nova dimensão} = a + (b \times \text{idade}) + (c \times \text{renda})$$

Um programa de computador de análise discriminante seria capaz de tomar os valores de idade e renda, examinar suas relações com o método de escrever e gerar uma equação que permitiria prever o método de redação de outros escritores, baseado em suas idades e rendas.

FIGURA 17-10
Projeção dos seis escritores em nova dimensão



Modelos Log-Lineares

Suponha que queremos saber se orientação política se relaciona com filiação partidária. Por exemplo, os liberais tendem, mais do que os conservadores, a ser Democratas? Dividindo nossa amostra em dois grupos — liberais e conservadores —, podemos calcular o percentual de Democratas em cada grupo. Se acharmos um maior percentual de Democratas entre os liberais, podemos concluir que de fato orientação política e filiação partidária são relacionadas.

Neste exemplo e nas análises de tabelas dos Capítulos 14 e 15, todas as variáveis dependentes analisadas eram dicotômicas, ou seja, tinham dois atributos. Mas, quando a variável dependente não é dicotômica, as coisas ficam mais complexas. Suponha que, além de Democratas e Republicanos, nossa amostra incluísse Independentes, Socialistas e Libertários. Não faz mais sentido examinar o percentual

de liberais e conservadores Democratas, como também não faz sentido examinar apenas os percentuais dos filiados a qualquer um dos outros grupos. Olhar para cada grupo separadamente resultaria em tabelas demais para a interpretação.

A situação se complica mais se a variável explicativa não for dicotômica. Se acrescentarmos “moderados” a liberais e conservadores ou examinarmos os efeitos interativos de outras variáveis explicativas, como raça ou religião, na equação política, as tabelas de porcentagem resultantes se tornariam incrivelmente complicadas.

Modelos log-lineares oferecem uma solução potencial a esta complexidade. Esta técnica, que envolve cálculos logarítmicos bem elaborados, é baseada na especificação de modelos que descrevem as inter-relações entre variáveis e na comparação entre frequências esperadas e observadas de células de tabela. (A lógica é semelhante à do *qui-quadrado*.) H. T. Reynolds descreve o processo:

No início da análise log-linear, como na maioria dos procedimentos estatísticos, o investigador propõe um modelo que ele acha que se ajusta aos dados. O modelo é um enunciado hipotético sobre como um conjunto de variáveis se inter-relacionam. Após escolher o modelo, ele estima as frequências esperadas numa amostra do tamanho certo *se o modelo fosse verdadeiro*. Em seguida, ele compara estas estimativas, F , com os valores esperados.⁴

Ao especificar os modelos a serem testados numa análise log-linear, o pesquisador considerará relações diretas entre a variável dependente e cada variável independente, relações entre pares de variáveis independentes e relações entre três variáveis (mais até, dependendo do número total de variáveis), como no modelo de elaboração. Consideremos um caso de três variáveis, tirado do exemplo anterior.

Podemos suspeitar que a filiação partidária de uma pessoa (“partido”) seja função de sua orientação política (“filosofia”) e raça. Os componentes deste modelo incluem então: (1) o efeito direto de filosofia em partido, (2) o efeito direto da raça em partido, (3) o efeito de raça em filosofia, (4) o efeito de raça na relação entre filosofia e partido (como no modelo de elaboração) e (5) o efeito de filosofia na relação raça e partido. Embora cada um destes componentes tenha *algum* poder explicativo, a análise log-linear provê um meio de identificar quais são mais importantes e quais, por questões práticas,

podem ser ignorados. Na análise log-linear os cálculos são muitos e complexos, mas os programas de informática podem fazê-los muito bem.

Parece crescer a popularidade da análise log-linear nas ciências sociais, embora seu uso ainda seja menor que o de outros métodos discutidos neste capítulo. É uma técnica relativamente nova e não são muitos os cientistas sociais treinados para usá-la.

Duas são as limitações principais da análise log-linear. Primeiro, sua lógica supõe certas premissas matemáticas que um determinado conjunto de dados pode não satisfazer (este é um tema complexo demais para abordar aqui). Segundo, como acontece com outras técnicas sumariantes, os resultados da análise log-linear não permitem a apreensão imediata e intuitiva que é possível em comparações de porcentagens ou de médias. Sendo assim, métodos log-lineares não seriam apropriados, mesmo se estatisticamente justificados, em casos onde a análise pode ser feita por simples tabelas percentuais, devendo seu uso ser reservado para situações complexas nas quais análises tabulares não são suficientemente fortes.

Completamos assim nossa discussão de algumas das técnicas analíticas mais comumente usadas por cientistas sociais. Apenas tocamos a superfície de cada uma delas e há muitas outras que nem foram abordadas. Minha intenção foi oferecer-lhe uma visão prévia de algumas das técnicas que podem depois ser estudadas com mais profundidade, além de familiarizá-lo com estas técnicas caso venha a encontrá-las ao ler relatos de pesquisas.

Resumo

Neste capítulo, sondamos mais profundamente as possibilidades de análises complexas e quantitativas de dados de pesquisas de *survey*. Não procurei treiná-lo no uso destas técnicas, apenas familiarizá-lo com elas. Talvez você queira aprender mais profundamente sobre alguma ou, pelo menos, ser capaz de lidar eficazmente com relatórios de pesquisa que usam estas técnicas. As técnicas descritas são apenas algumas das mais populares hoje em dia em pesquisas de *survey*, e penso que dão uma boa visão geral das possibilidades.

Análise de regressão representa as relações entre variáveis sob a forma de equações, que podem ser usadas para prever os valores de uma variável dependente com base nos valores de uma ou mais variáveis independentes.

A equação básica de regressão, para uma *regressão linear* simples, tem a forma $Y = a + bX$. Neste caso, Y é o valor (estimado) da variável dependente, a é um valor constante e b é outro valor numérico, que é multiplicado por X , o valor da variável independente.

Equações de regressão são computadas com base numa *linha de regressão*, a linha geométrica que representa, com a menor quantidade de discrepância, a localização real de pontos num gráfico. Análise de *regressão múltipla* resulta numa equação de regressão, que estima os valores de uma variável dependente a partir dos valores de diversas variáveis independentes. Análise de *regressão parcial* examina os efeitos de diversas variáveis independentes, com o efeito de cada uma delas expresso separadamente, quando os efeitos das demais são mantidos constantes.

A análise de *regressão curvilínea* permite que a linha de regressão mais bem ajustada não seja uma linha reta. Consegue-se a curvatura da linha de regressão permitindo que os valores das variáveis independentes sejam elevados a potências maiores que 1, por exemplo, ao quadrado, ao cubo etc.

Análise de trajetória é um método de apresentar graficamente as redes de relações causais entre diversas variáveis, ilustrando de forma gráfica as "trajetórias" principais das variáveis através das quais as variáveis independentes causam as dependentes. *Coefficientes de trajetória* são coeficientes de regressão padronizados que representam relações parciais entre variáveis.

Análise fatorial, só exequível com computador, é um método analítico para descobrir as dimensões gerais representadas por um conjunto de variáveis reais. Estas dimensões gerais, ou *fatores*, são dimensões hipotéticas calculadas que não são representadas perfeitamente por nenhuma das variáveis empíricas estudadas, mas que são fortemente associadas a grupos de variáveis empíricas. Uma *carga fatorial* indica o grau de associação entre uma variável empírica e um fator.

Análise de variância se baseia na comparação de variações entre e dentro de grupos e em determinar se diferenças entre grupos podem ter ocorrido razoavelmente em amostragem aleatória simples ou se é mais razoável concluir que representam uma relação genuína entre as variáveis envolvidas.

Análise discriminante procura explicar a variação de alguma variável dependente, encontrando uma dimensão composta hipotética que separa categorias da variável dependente. Resulta numa equação que atribui escores às pessoas com base na dimensão hipotética e permite predizer seus valores na dependente variável.

Modelos log-lineares oferecem um método para analisar relações complexas entre variáveis nominais com mais de dois atributos cada.

Notas

¹ Adaptado de YERG, Beverly J. Reflections on the Use of the RTE Model in Physical Education. *Research Quarterly for Exercise and Sport*, p.42, mar. 1981.

² FORSLUND, Morris A. Patterns of Delinquency Involvement: An Empirical Typology, apresentado no Encontro Anual da Western Association of Sociologists and Anthropologists, Lethbridge, Alberta, 8 de fevereiro, 1980.

³ Ibidem. p.4.

⁴ REYNOLDS, H. T. *Analysis of Nominal Data*. Beverly Hills, CA: Sage, 1977. p.76-77.

Leituras Adicionais

HENKEL, Ramon E. *Tests of Significance*. Beverly Hills, CA: Sage, 1976.

HILDEBRAND, David K., LAING, James D., ROSENTHAL, Howard. *Analysis of Ordinal Data*. Beverly Hills, CA: Sage, 1977.

IVERSEN, Gudmund R., NORPOTH, Helmut. *Analysis of Variance*. Beverly Hills, CA: Sage, 1976.

KIM, Jae-On, MUELLER, Charles W. *Introduction to Factor Analysis*. Beverly Hills, CA: Sage, 1978.

KNOKE, David, BURKE, Peter J. *Log-Linear Models*. Beverly Hills, CA: Sage, 1980.

LEWIS-BECK, Michael S. *Applied Regression: An Introduction*. Beverly Hills, CA: Sage, 1980.

REYNOLDS, H. T. *Analysis of Nominal Data*. Beverly Hills, CA: Sage, 1977.

Capítulo 18

O Relatório da Pesquisa de Survey

Até aqui, consideramos as atividades envolvidas na *feitura* da pesquisa. Agora, discutiremos uma questão frequentemente negligenciada: o relatório do *survey*. Se o *survey* não for comunicado apropriadamente, serão em vão os esforços feitos na sua *feitura*.

Para começar, bons relatórios científico-sociais exigem bom conhecimento da língua. Sempre que pedimos aos "números para falarem por si", eles tendem a permanecer mudos. Sempre que usamos terminologia ou construção indevidamente complexas, dificultamos a comunicação.

O relatório científico tem várias funções, e é bom lembrá-las. Primeira, o relatório comunica a uma audiência um corpo de dados e idéias específicas. Ele deve fornecê-los de maneira clara e com detalhes suficientes para permitir uma avaliação bem-informada. Segunda, o relatório científico deve ser visto como uma contribuição ao conhecimento científico geral. Humildemente, você deve reconsiderar o relatório da sua pesquisa como um acréscimo ao que é conhecido sobre o comportamento social. E, finalmente, o relatório deve estimular e orientar futuros inquéritos.

Algumas Considerações Básicas

Dentro destas diretrizes gerais, relatórios diferentes servem a diferentes fins. Um relatório apropriado para um fim pode ser inapropriado para outro. Esta seção lida com algumas das considerações básicas relativas à redação correta do relatório.

Audiência

Antes de rascunhar o relatório, você deve se perguntar quem o lerá. Normalmente, você deve distinguir entre cientistas e leigos. Se estiver escrevendo para os primeiros, você pode fazer certas suposições com relação aos conhecimentos deles e talvez possa resumir alguns pontos sem explicá-los detalhadamente. Igualmente, pode usar uma linguagem mais técnica da que usaria para uma audiência de leigos. Por exemplo, um pesquisador de *marketing*, ao escrever para um cliente, deve explicar as coisas com mais detalhes do que seria necessário para uma audiência de pesquisadores de *marketing*.

Ao mesmo tempo, lembre-se de que qualquer ciência é composta de facções ou cultos. Os termos e as suposições aceitáveis por seus colegas próximos podem confundir outros cientistas. Essa restrição se aplica tanto à substância quanto às técnicas.

Forma e Tamanho do Relatório

Estes comentários se aplicam a relatórios escritos e verbais. Entretanto, a natureza do relatório é influenciada pela forma que você usar.

Uma variedade de relatórios pode resultar de um projeto de pesquisa. Para começar, você pode querer preparar uma curta *nota de pesquisa* para publicação num periódico técnico ou acadêmico. Tais relatórios devem ter aproximadamente de uma a cinco páginas (digitadas em espaço duplo) e ser concisos e diretos. Você não conseguirá apresentar o estado do campo detalhadamente em pouco espaço, e suas notas metodológicas deverão ser breves. Basicamente, você deve dizer por que acha que uma nota tão pequena se justifica por suas descobertas, e quais foram estas descobertas.

Freqüentemente, é preciso fazer relatórios para os patrocinadores da pesquisa. Estes variam muito de tamanho. Na preparação do relatório, você deve considerar quem vai lê-lo — cientistas ou leigos — e a razão do patrocínio do projeto. É má política e falta de educação incomodar os patrocinadores com descobertas que não têm valor ou interesse para eles. Ao mesmo tempo, pode ser útil resumir o que a pesquisa acrescentou ao conhecimento científico básico (se é que ela acrescentou algo).

Textos de trabalho ou monografias são outra forma de relatório. Principalmente se você estiver envolvido num projeto grande e complexo, será útil obter comentários sobre sua análise e sua interpretação dos dados. Um texto de trabalho é uma apresentação experimental com um pedido implícito de comentários. Textos de trabalho também podem variar de tamanho, e podem apresentar todas as descobertas da pesquisa ou apenas parte delas. Sua reputação profissional não está em jogo num texto de trabalho, portanto você tem liberdade de apresentar interpretações experimentais que não é capaz de justificar, identificando-as como tal e pedindo avaliações.

Muitos projetos de pesquisa resultam em trabalhos a serem apresentados em encontros profissionais. Frequentemente, estes servem ao mesmo propósito dos textos de trabalho. Você pode apresentar descobertas e idéias aos colegas e pedir que as comentem. Apesar do tamanho dos trabalhos profissionais variar, dependendo da organização dos encontros, sugerimos que você fale pouco, ao invés de falar demais. Embora um texto de trabalho possa ir e vir através de várias conclusões tentativas, os participantes da conferência não devem ser forçados a sentar e ouvir o que já foi escrito. Ouvintes interessados podem sempre pedir mais detalhes mais tarde, e os desinteressados podem escapar graciosamente.

Provavelmente, o relatório mais comum é um artigo publicado num periódico acadêmico. Novamente, o tamanho varia; você deve examinar os tamanhos de artigos publicados previamente no periódico em questão. Entretanto, como orientação geral, 25 páginas digitadas é um bom tamanho. A seção deste capítulo sobre a organização do relatório se baseia primariamente na estrutura de um artigo de periódico, portanto não direi mais nada agora, exceto para sugerir que os trabalhos semestrais de estudantes devem ser escritos neste modelo. Como regra geral, um trabalho semestral que daria um bom artigo de periódico também seria um bom trabalho.

Um livro, é claro, representa a forma mais prestigiada de relatório de pesquisa. Tem todas as vantagens do trabalho escrito — tamanho e detalhes —, mas deve ser um documento mais bem-acabado. Uma vez que a publicação em livro de descobertas de pesquisa lhes dá uma aparência de mais substância e valor, você tem uma obrigação maior para com sua audiência. Apesar de esperar receber comentários dos colegas, talvez levando-o a rever suas idéias, você deve lembrar que leitores menos sofisticados podem ser levados a aceitar suas descobertas acriticamente.

Objetivo do Relatório

Já consideramos, acima, os diferentes propósitos da pesquisa de *survey*. Ao preparar o relatório, você deve lembrar estas diferenças.

Alguns relatórios podem focalizar primariamente a *exploração* de um tópico. São inerentes a este objetivo as conclusões experimentais e incompletas. Você deve indicar claramente à audiência o objetivo exploratório do estudo e apontar as limitações dele. Um aspecto importante de um relatório exploratório é sua utilidade em mostrar o caminho para pesquisas mais refinadas sobre o assunto.

A maioria dos estudos tem um objetivo *descritivo*, e os relatórios destes estudos terão um elemento descritivo. Você deve distinguir cuidadosamente as descrições que se aplicam apenas à amostra das que se aplicam, por inferência, à população. Se forem feitas descrições inferenciais, você deve dar à audiência alguma indicação da possível magnitude de erro delas.

Muitos relatórios têm um objetivo *explicativo*; você quer mostrar relações causais entre variáveis. Dependendo da audiência provável do relatório, você deve delinear cuidadosamente as regras da explicação que sustentam seus cálculos e conclusões; como no caso da descrição, dê aos leitores alguma orientação sobre a certeza relativa das suas conclusões.

Finalmente, alguns relatórios de pesquisa podem ter o objetivo de *propor ação*. Por exemplo, estudando as raízes do preconceito, você pode querer sugerir maneiras de reduzi-lo, baseado nas descobertas da pesquisa. Este objetivo frequentemente leva a problemas complicados, já que seus próprios valores e orientações podem interferir em suas propostas. É perfeitamente legítimo que suas propostas sejam motivadas por valores pessoais, mas você deve assegurar-se de que as ações específicas que propõe sejam garantidas pelos seus dados. Portanto, tenha cuidado especial em mostrar claramente a lógica pela qual você saiu dos dados empíricos em direção à ação proposta.

Organização do Relatório

Embora a organização de relatórios difira um pouco quanto à forma e ao propósito, é possível sugerir um formato genérico para se apresentar dados de pesquisa. Os comentários a seguir

se aplicam mais diretamente a um artigo de periódico, mas, com algumas modificações, também se aplicam à maioria dos formatos de relatórios de pesquisa.

Propósito e Resumo

Começar com uma introdução rápida ao propósito do estudo e às principais descobertas da análise sempre ajuda os leitores. Num artigo de periódico, isso pode ser feito na forma de um resumo.

Você pode achar difícil resumir, principalmente se sua análise envolveu considerável trabalho de detetive, com descobertas importantes revelando-se ao fim de deduções imaginativas e manipulações de dados. Neste caso, você pode querer conduzir os leitores através do mesmo processo excitante, narrando o processo de descoberta com suspense e surpresa. Se este formato de relatório fornecer um quadro preciso do processo da pesquisa, acredito que tenha considerável valor instrutivo. Mesmo assim, alguns leitores podem não estar interessados em acompanhar toda a narração da pesquisa, e desconhecer o propósito e as conclusões gerais antecipadamente pode dificultar para eles compreender o significado do estudo.

Um antigo ditado forense diz, "Diga-lhes o que você irá dizer-lhes, diga-lhes, e diga-lhes o que você lhes disse." Você se sairá bem seguindo este ditado na preparação dos relatórios da pesquisa.

Revisão da Literatura

Como todo relatório de pesquisa deve ser posto no contexto do corpo geral de conhecimento científico, você sempre deve indicar onde seu relatório se encaixa neste quadro. Após apresentar o propósito geral do estudo, você deve atualizar os leitores sobre as pesquisas anteriores na área, apontando para as concordâncias e discordâncias com os pesquisadores anteriores.

Em alguns casos, você pode querer questionar idéias previamente aceitas. Reveja cuidadosamente os estudos que levaram à aceitação dessas idéias e depois indique os fatores que não foram considerados ou falácias lógicas em estudos anteriores.

Quando quiser resolver uma discordância entre pesquisadores anteriores, organize sua revisão da literatura ao redor dos pontos de vista opostos. Resuma a pesquisa que apóia um lado, em seguida a que apóia o outro e sugira as razões do desacordo.

Até certo ponto, sua revisão da literatura serve a uma função bibliográfica para os leitores, indexando a pesquisa anterior sobre algum tema. Mas essa revisão pode ser excessiva, e você deve evitar escrever um parágrafo introdutório de três páginas mencionando todos os estudos prévios. Essa função bibliográfica abrangente pode ser melhor executada por uma bibliografia no fim do relatório; a revisão da literatura deve focalizar somente os estudos que tiverem relevância para o seu estudo.

Evitando o Plágio

Ao relatar o trabalho de outros, deixe claro quem disse o quê. É essencial evitar o plágio, que é o roubo de idéias e/ou palavras de outros — intencional ou acidentalmente — e a apresentação dessas idéias e palavras como se fossem suas. Como este problema é comum e algumas vezes pouco claro, vamos examiná-lo com mais detalhes.

Eis as principais regras com relação ao plágio:

- Você não pode usar as palavras exatas de outro autor sem pô-las entre aspas e sem indicar a fonte da citação, de forma que os leitores possam localizar a citação no contexto original. Como regra geral, dê a fonte de toda citação de oito ou mais palavras.
- Não é aceitável *editar* ou *parafrasear* as palavras de outra pessoa e apresentar a versão revista como seu próprio trabalho.
- Finalmente, não é aceitável apresentar *idéias* de outros como suas próprias idéias — mesmo usando palavras completamente diferentes para expressá-las.

Um exemplo explica melhor o que é e o que não é aceitável no uso do trabalho de outro. Considere o seguinte trecho original.

Sistemas são como bebês: quando você consegue um, você o tem. Eles não vão embora. Ao contrário, eles são de uma persistência notável. Eles não apenas persistem; eles crescem. E, à medida que crescem, ultrapassam limites. O potencial de crescimento de sistemas foi explorado de maneira experimental e preliminar por Parkinson, que

concluiu que sistemas administrativos mantêm uma média de crescimento de 5 a 6% ao ano, independentemente do trabalho a ser feito. Parkinson estava certo no que disse e devemos creditar-lhe toda a honra por iniciar um estudo sério sobre este importante tópico.¹

Vejamos algumas maneiras aceitáveis de usar o trabalho de Gall num texto.

Aceitável: John Gall, em seu trabalho sobre *Systemantics*, faz um paralelo bem-humorado entre sistemas e bebês: "Sistemas são como bebês: quando você consegue um, você o tem. Eles não vão embora. Ao contrário, eles são de uma persistência notável. Eles não apenas persistem; eles crescem."²

Aceitável: John Gall adverte que sistemas são como bebês. Crie um sistema e ele permanecerá. Pior ainda, Gall sugere, os sistemas continuam a crescer cada vez mais.³

Aceitável: Também é sugerido que sistemas têm uma tendência natural a persistir, e mesmo crescer e ir além dos limites (Gall, 1975, p.33).⁴

Eis, agora, alguns usos inaceitáveis do mesmo material, refletindo alguns erros comuns.

Inaceitável: Neste trabalho, quero ver algumas das características dos sistemas sociais que criamos em nossas organizações. Primeiro, sistemas são como bebês: quando você consegue um, você o tem. Eles não vão embora. Ao contrário, eles são de uma persistência notável. Eles não apenas persistem; eles crescem.⁵

Inaceitável: Neste trabalho, quero ver algumas das características dos sistemas sociais que criamos em nossas organizações. Primeiro, os sistemas são muito parecidos com crianças: quando você consegue uma, ela é sua. Elas não vão embora; elas persistem. Elas não apenas persistem, na verdade: elas crescem.⁶

Inaceitável: Neste trabalho, quero ver algumas das características dos sistemas sociais que criamos em nossas

organizações. Uma coisa que percebi é que uma vez que você cria um sistema, ele parece nunca ir embora. Ocorre exatamente o contrário: eles têm uma tendência a crescer. Podemos dizer que os sistemas são muito parecidos a crianças neste aspecto.⁷

Cada exemplo inaceitável acima é um exemplo de plágio, e são delitos sérios. Temos que admitir que há "áreas cinzentas". Algumas idéias são mais ou menos de domínio público, não "pertencendo" a ninguém. Você pode ter uma idéia que outra pessoa já escreveu. Se tiver dúvidas com relação a alguma situação específica, discuta-a com seu orientador.

Discuti esse tópico detalhadamente porque é importante que você situe sua pesquisa no contexto do que outros fizeram e disseram; o uso impróprio do material de outras pessoas é um delito sério. Dominar a atribuição adequada de idéias é parte do seu "amadurecimento" acadêmico.

Desenho e Execução do Estudo

Um relatório de pesquisa com descobertas e conclusões interessantes pode ser causa de frustração se você não conseguir expor o desenho metodológico e a execução do estudo. A validade das descobertas científicas depende muito da maneira como os dados foram coletados e analisados.

Ao relatar o desenho e a execução de um projeto de *survey*, inclua sempre o seguinte: a população, a moldura e o método de amostragem, o tamanho da amostra, o método de coleta dos dados, a taxa de respostas obtidas e o método de processamento e análise dos dados. Com a experiência, você será capaz de relatar estes detalhes em espaço relativamente pequeno sem omitir nada do que os leitores precisam para a avaliação do estudo.

Análise e Interpretação

Tendo posto o estudo na perspectiva da pesquisa prévia e tendo descrito seus desenho e execução, você deve apresentar os dados. A seção sobre diretrizes para relatar a análise informará mais sobre esta questão. Mas lembremos algumas idéias gerais.

A apresentação e a manipulação dos dados, e sua interpretação, devem ser integradas num todo lógico. É frustrante

para os leitores descobrirem uma coleção de análises e descobertas aparentemente não relacionadas, com a promessa de que todas as pontas soltas serão amarradas mais tarde no relatório. Cada etapa da análise deve fazer sentido na hora que é relatada. Você deve apresentar sua razão para uma análise particular, apresentar os dados relevantes a ela, interpretar os resultados e depois indicar para onde nos levam os resultados.

Resumo e Conclusões

Seguindo o ditado antes citado, você deve garantir o resumo do relatório da pesquisa. Deve evitar rever cada descoberta específica, mas deve rever todas as significativas, mostrando novamente sua significância geral.

O relatório deve ser concluído declarando o que você descobriu sobre o tema em questão e para onde deverão apontar futuras pesquisas. Uma revisão rápida de artigos de periódicos recentes provavelmente mostraria alta frequência de ocorrência desta declaração conclusiva: "Está claro que precisamos de muito mais pesquisa." Esta é provavelmente uma conclusão sempre verdadeira, mas de pouco valor a menos que você possa oferecer sugestões pertinentes sobre a natureza das pesquisas futuras. Reveja as limitações do seu próprio estudo e sugira maneiras de pesquisadores futuros evitarem estas limitações.

Diretrizes para Relatar as Análises

A apresentação das análises dos dados deve dar o máximo de detalhes sem parecer confusa. Isso pode ser alcançado examinando-se continuamente os objetivos descritos abaixo.

Os dados devem ser apresentados de modo a permitir aos leitores fazer as recomputações que quiserem. Por exemplo, no caso das tabelas de porcentagem, os leitores devem ser capazes de fundir categorias e recalcular as porcentagens. Eles devem ter informação bastante para permitir percentuar a tabela na direção oposta à sua.

Todos os aspectos da análise devem ser descritos com detalhes suficientes para permitir a um analista secundário duplicar a análise a partir do mesmo corpo de dados. Isto significa ser capaz de criar os mesmos índices e escalas, produzir as mesmas tabelas, chegar às mesmas equações de regressão, obter os mesmos fatores e cargas de fatores etc. É claro que

essa recriação raramente será feita, mas se o relatório for apresentado de forma a permitir que se faça isso, seus leitores estarão muito mais bem-equipados para avaliá-lo.

Seus leitores devem poder replicar completamente toda a pesquisa de maneira independente. Devem poder identificar a mesma população, selecionar da mesma maneira o mesmo tamanho de amostra, coletar dados da mesma maneira, executar as mesmas análises e ver se chegam às mesmas conclusões. Lembre a discussão anterior na qual dissemos que a replicabilidade é uma norma essencial da ciência em geral. Um único estudo não prova coisa alguma; somente uma série de estudos pode começar a fazê-lo. A menos que os estudos possam ser replicados, não haverá nenhuma série significativa de estudos.

Mencionei antes a importância de se integrar dados, análises e interpretações no relatório. Podemos dar uma diretriz mais específica. Tabelas, gráficos e figuras devem ser integrados ao texto do relatório, aparecendo próximos à porção do texto que os discute. Estudantes, algumas vezes, descrevem suas análises no corpo do relatório e põem as tabelas no apêndice, ao final.⁸ Este procedimento atrapalha muito o leitor. Como regra geral, é melhor (1) descrever o propósito de se apresentar uma tabela, (2) apresentá-la, e (3) revê-la e interpretá-la.

Seja explícito ao tirar conclusões. Apesar de as pesquisas serem geralmente feitas para tirar conclusões gerais, você deve apontar cuidadosamente as bases específicas para elas. Senão, você pode levar o leitor a aceitar conclusões infundadas.

Mostre as qualificações ou condições garantidas na avaliação das conclusões. Tipicamente, quem fez o *survey* está em melhor posição para conhecer os limites e as incertezas das conclusões, e você deve dar ao leitor a vantagem desse conhecimento. Não fazê-lo pode desorientar pesquisas futuras e resultar em perda de verbas para pesquisas.

Concluirei com uma declaração já feita no início do capítulo. Os relatórios de pesquisas devem ser escritos no melhor estilo literário possível. Escrever claramente é mais fácil para uns do que para outros, e é sempre mais difícil do que escrever mal. Portanto, revise sempre o que escreveu.

Resumo

Uma pesquisa perfeitamente projetada, cuidadosamente executada e brilhantemente analisada não terá validade a menos que você possa comunicar suas descobertas aos outros. Este capítulo tentou dar algumas diretrizes gerais e específicas para alcançar este objetivo. Seus melhores guias são lógica, clareza e honestidade. Não há substituto para a prática.

Notas

¹GALL, John. *Systemantics: How Systems Work and Especially How They Fail*. New York: Pocket Books, 1975. p.33-34. (Nota: Gall fez antes uma citação completa de Parkinson.)

²Isto seria acompanhado por uma nota de rodapé: GALL, John. *Systemantics: How Systems Work and Especially How They Fail*. New York: Pocket Books, 1975. p.33.

³Isto seria acompanhado por uma nota de rodapé: GALL, John. *Systemantics: How Systems Work and Especially How They Fail*. New York: Pocket Books, 1975. p.33.

⁴Nota: este formato exige que você forneça a referência bibliográfica completa em sua bibliografia.

⁵É inaceitável citar diretamente o material de alguém sem usar aspas e dar a referência bibliográfica completa.

⁶É inaceitável editar o trabalho de outra pessoa e apresentá-lo como seu.

⁷É inaceitável parafrasear as idéias de alguém e apresentá-las como suas.

⁸Devo dizer que alguns editores preferem receber textos dessa maneira, porque pessoas diferentes vão processar os materiais gráficos e de texto durante o processo da edição. Com os atuais processadores de texto, porém, não é difícil criar duas versões de um trabalho — um para leitores, outro para publicação.

Leituras Adicionais

BART, Pauline, FRANKEL, Linda. *The Student Sociologist's Handbook*. Glenview, IL: Scott, Foresman, 1981.

BECKER, Leonard, GUSTAFSON, Clair. *Encounter with Sociology: The Term Paper*. San Francisco: Boyd and Fraser, 1976.

A Pesquisa de *Survey* no Contexto Social

CAPÍTULO 19

A ÉTICA NA PESQUISA DE *SURVEY*

CAPÍTULO 20

O CONSUMIDOR INFORMADO DE PESQUISA DE *SURVEY*

Na maior parte, este livro abordou as habilidades para fazer pesquisa de *survey*. Foram feitos todos os esforços para situar este método de pesquisa no contexto da ciência em geral. Os últimos capítulos refletem minha opinião de que mesmo esta perspectiva não é suficientemente abrangente.

Os comentários preparatórios da introdução do livro apontaram as maneiras pelas quais a pesquisa de *survey* afeta a vida de todos nós, pesquisadores ou não. Portanto, não basta considerar a pesquisa de *survey* apenas como uma técnica científica neutra. Em vez disso, devemos estar cientes do contexto social no qual os *surveys* são conduzidos.

O Capítulo 19 discute a ética em pesquisa, e como ela se relaciona aos métodos do *survey*. Uma discussão das diretrizes éticas gerais em pesquisas de *survey* é seguida por uma série de exemplos de situações de pesquisa envolvendo questões éticas. Mesmo que poucas situações possam ser consideradas inequivocamente éticas ou antiéticas, acredito que uma consideração atenta delas o tornará mais consciente dos complicados problemas éticos que freqüentemente os pesquisadores de *survey* têm que enfrentar.

GRUBER, James, PRYOR, Judith, BERGE, Patricia. *Materials and Methods for Sociology Research*. New York: Neal-Schuman, 1980.

LI, Tze-Chung. *Social Science Reference Sources: A Practical Guide*. Westport, CT: Greenwood Press, 1980.

O Capítulo 20 conversa com você de maneira um pouco diferente. Você pode não conduzir muitos *surveys* depois de ler este livro, mas certamente lerá sobre muitos e, em alguns casos, terá que tomar decisões baseando-se neles. Meu propósito neste capítulo é habilitá-lo como um eficiente consumidor de pesquisa de *survey*.

Capítulo 19

A Ética na Pesquisa de *Survey*

Um manual teoricamente orientado sobre métodos de pesquisa de *survey* daria ao estudante apenas uma imagem ideal de como os *surveys* devem ser feitos. Em seções anteriores, procurei enfatizar que há várias questões administrativas e práticas envolvidas num processo de pesquisa, de forma que não é possível seguir exatamente um modelo ideal. O pesquisador bem-informado deve estar ciente destas limitações e ser capaz de equilibrar fatores administrativos e científicos para chegar ao melhor compromisso possível.

Este capítulo aborda outra restrição não científica — as questões *éticas* na pesquisa de *survey*. Estas questões não são parte do método científico.¹ Mesmo assim, formam um grupo de normas que os cientistas da maioria das disciplinas são obrigados a seguir. Em muitos casos, essas normas éticas entram em conflito direto com os procedimentos científicos, assim como com questões administrativas. Você deve estar ciente dos possíveis conflitos para poder conduzir uma pesquisa mais científica e mais ética.

A ciência em si é amoral. A lei da gravidade e a correlação entre educação e preconceito não são nem morais, nem imorais. Entretanto, os cientistas não são amorais, nem o são quem pode usar os resultados da pesquisa científica. A pesquisa científica pode ser conduzida e/ou usada para propósitos morais ou imorais. É claro, a moralidade de um homem é a imoralidade de outro. São lições que os físicos nucleares aprenderam antes e os cientistas sociais ainda estão aprendendo.

Não há como assegurar que todos os cientistas serão sempre motivados por questões éticas quando embarcam numa

pesquisa científica, nem há como assegurar que as descobertas científicas serão usadas com propósitos éticos. Entretanto, podemos mostrar um corpo de normas éticas mais ou menos consensual em relação à feitura de pesquisas. Este capítulo apresenta alguns dos problemas éticos mais comuns na pesquisa de *survey* e sugere algumas soluções éticas para eles que não comprometem perigosamente a qualidade “científica” da pesquisa.

Participação Voluntária

A pesquisa de *survey* quase sempre é uma intrusão na vida das pessoas. O entrevistador batendo à porta ou o questionário chegando pelo correio assinala o início de uma atividade que os respondentes não pediram e que pode requerer grande quantidade de tempo e esforço. Além disso, o *survey* freqüentemente pede ao respondente revelar informação pessoal — atitudes e características pessoais — que podem não ser conhecidas por seus amigos ou associados. O *survey* necessita que ele revele tal informação a um completo estranho.

Outros profissionais, como médicos e advogados, também requerem tal informação. Mas elas podem ser justificadas pelo fato de a informação ser necessária para o interesse pessoal do cliente a quem requerem a informação. Pesquisadores de *survey* raramente podem usar esta justificativa. Assim como os pesquisadores médicos, podemos apenas dizer que nossa pesquisa pode, no futuro, ajudar as pessoas em geral.

Um princípio fundamental da ética na pesquisa médica é que a participação experimental deve ser *voluntária*. A mesma norma se aplica à pesquisa de *survey*. Ninguém deve ser forçado a participar. Mas esta norma é bem mais facilmente aceita na teoria do que na prática. Novamente, a pesquisa médica permite um paralelo útil. Muitas drogas experimentais são testadas em prisioneiros. Nos casos mais rigorosamente éticos, os prisioneiros conhecem a natureza — e os possíveis perigos — do experimento, são informados que a participação é absolutamente voluntária e são ainda instruídos de que não devem esperar nenhuma compensação especial (como uma liberdade condicional precoce) pela participação. Mesmo nestas condições, os voluntários são freqüentemente motivados pela crença de que terão benefícios pessoais de sua cooperação.

Quando os professores em aulas introdutórias de sociologia pedem aos alunos que preencham questionários que eles esperam analisar e publicar, devem sempre enfatizar aos estudantes que a participação em *surveys* é completamente voluntária. Mesmo assim, muitos estudantes vão temer que, se não participarem, suas notas serão afetadas durante o curso. Neste caso, os professores devem estar conscientes das sanções implícitas e tomar as providências necessárias para torná-las óbvias. Podem sair da sala enquanto os questionários estiverem sendo preenchidos e postos numa caixa, ou pedir aos alunos que enviem os questionários pelo correio ou os ponham numa caixa próxima à porta quando entrarem na sala, na próxima aula.

Nesta situação, os professores devem ser encorajados, se for apropriado, a envolver os alunos no processamento e na análise dos dados, dando-lhes uma experiência de aprendizado pessoalmente valiosa. É claro que os alunos deverão receber crédito em qualquer relatório publicado sobre o estudo. (Esta técnica não deve ser usada como um método para obter mão-de-obra barata.)

A idéia de que a participação em *surveys* deve ser voluntária vai diretamente de encontro a uma preocupação científica. Se técnicas estatísticas forem usadas legitimamente na análise dos dados, então cada membro da amostra aleatória deve participar. Mesmo com planos estatísticos menos rigorosos, você desejará uma alta taxa de respostas para assegurar uma amostra razoavelmente representativa. Portanto, mesmo não podendo eticamente requerer a participação, você geralmente fará o possível para obtê-la.

Num *survey* por entrevistas, os entrevistadores geralmente são treinados a persuadir pessoas indecisas. “Só levará alguns segundos”, “Os resultados do estudo serão valiosos para as pessoas residentes nesta área” e “Posso voltar em momento mais conveniente?” são frases importantes do vocabulário do entrevistador. O diretor da pesquisa descobrirá alguns entrevistadores mais persuasivos do que outros, e estes provavelmente serão designados para ir atrás de membros da amostra que recusaram o pedido inicial de participação. (A experiência mostra que uma entrevistadora grávida de oito meses é mais bem-sucedida do que a média neste sentido.) O limite entre a persuasão ética e a coerção é impreciso.

Num *survey* pelo correio, você normalmente envia correspondências de acompanhamento aos membros da amostra que não responderam ao pedido inicial de participação. Geralmente,

as correspondências de acompanhamento são seguidas de apelos especiais por cooperação; novamente, o limite entre a persuasão e a coerção não é muito preciso.

Sem Prejuízo para os Entrevistados

A pesquisa de *survey* nunca deve prejudicar os entrevistados que se voluntariaram a cooperar. Talvez o exemplo mais claro desta norma na prática tenha a ver com revelar informação que envergonharia os respondentes ou colocaria em risco suas vidas familiares, suas amizades, empregos, e assim por diante. Esta situação será discutida em maiores detalhes na seção sobre anonimato e confidencialidade.

Também é possível que os entrevistados sejam prejudicados durante a entrevista, e você deve estar ciente desta possibilidade e evitar que isso ocorra. Mesmo com o entrevistador mais profissional (ou seja, neutro), os entrevistados podem se sentir incomodados em relatar comportamento desviado e atitudes que consideram pouco comuns, ou admitir baixos salários, receber pensão do governo, e coisas deste tipo.

Os *surveys* freqüentemente forçam os respondentes a enfrentar aspectos de si mesmos que eles normalmente não consideram. Em retrospectiva, alguns comportamentos do passado podem parecer injustos ou imorais, e eles podem agora se arrepende destes comportamentos. Assim, o *survey* pode ser uma fonte de agonia pessoal para os respondentes. Eles podem começar a achar que não são tão morais e éticos quanto deveriam, e este pensamento pode incomodá-los.

Não há maneira de se proteger contra tais possibilidades. Mas alguns itens do questionário têm maior probabilidade de produzir tais reações do que outros, e você deve ser sensível a isto. Se um item tiver probabilidade de produzir reações desagradáveis nos respondentes, você deve ter fortes razões científicas para fazer tal pergunta. A menos que seja vital para os objetivos de pesquisa, remova este item. Se ele for essencial e sensível, você se verá em terreno eticamente perigoso e sofrerá sua própria agonia.

Apesar de o fato quase sempre passar despercebido, respondentes podem ser prejudicados pela análise e relatório dos dados. De vez em quando, eles lêem os livros publicados a partir do *survey* de que participaram. Respondentes razoavelmente

sofisticados serão capazes de se localizar em índices e tabelas e se verem caracterizados — apesar de não estarem identificados pelos nomes — como fanáticos, impatrióticos, irreligiosos etc. Na melhor das hipóteses, esta descoberta pode incomodá-los e ameaçar as imagens que têm deles mesmos. Ainda que o propósito da pesquisa seja explicar por que algumas pessoas têm preconceitos e outras não.

Uma vez, fiz um *survey* com mulheres da Igreja Episcopal do Norte da Califórnia. Pedi aos pastores de uma amostra das igrejas que distribuíssem questionários para uma amostra específica dos membros, os coletassem e os devolvessem ao escritório da pesquisa. Um pastor leu os questionários da amostra antes de devolvê-los e fez um sermão enfático e rigoroso para a congregação, dizendo que muitos deles eram ateus e iriam para o inferno. Mesmo que o pastor não conhecesse a identidade dos respondentes que deram as respostas, parece certo que muitos destes foram pessoalmente feridos por isto.

Como a participação voluntária, não prejudicar os entrevistados é uma norma fácil de aceitar em teoria, mas difícil de assegurar na prática. Sensibilidade com a questão e experiência com suas aplicações podem melhorar seu desempenho.

Anonimato e Sigilo

A situação mais clara envolvendo a proteção do interesse e bem-estar dos respondentes se relaciona à proteção de suas identidades. Se revelar suas respostas pode prejudicá-los de alguma maneira, é ainda mais importante seguir esta norma. Duas técnicas — *anonimato* e *sigilo* — ajudam a proteger a identidade dos respondentes, embora as duas sejam freqüentemente confundidas.

Anonimato

Um respondente pode ser considerado *anônimo* quando você mesmo não consegue relacionar uma dada resposta a algum respondente. Isto significa que um respondente de *survey* por entrevista nunca poderá ser considerado anônimo, porque o entrevistador colhe a informação de um entrevistado identificável. (Isto supondo que sejam seguidos métodos padrão de amostragem.) Um exemplo de anonimato é

um *survey* pelo correio no qual nenhum número de identificação é posto no questionário antes de seu retorno ao escritório da pesquisa.

É claro que o anonimato complica os planos de acompanhamento para aumentar as taxas de resposta. Se você não sabe quem de sua amostra não respondeu o questionário, não poderá entrar em contato apenas com não-respondentes. Mas, como alternativa, você pode enviar o questionário novamente para todos os membros da amostra original, pedindo aos que já responderam que ignorem o segundo questionário, ou você pode usar a técnica do cartão postal discutida no Capítulo 9.

Apesar das dificuldades de garantir o anonimato, em algumas situações é melhor você pagar o preço necessário. Por exemplo, em estudos sobre uso de drogas entre estudantes universitários, você pode preferir não saber a identidade dos respondentes, por duas razões. Primeira, garantir o anonimato pode aumentar a probabilidade e a honestidade das respostas. Segunda, você pode não querer encarar a possibilidade de ser questionado pela polícia quanto aos nomes dos usuários de drogas. No segundo caso, um *survey* anônimo irá lhe poupar ter que escolher entre (1) violar sua promessa aos respondentes, (2) arriscar ir para a prisão por ocultar evidência criminosa, ou (3) rezar para que os questionários acidentalmente peguem fogo e sejam destruídos.

Sigilo

Num *survey confidencial*, você é capaz de identificar as respostas de uma certa pessoa, mas prometeu que não o fará. Por exemplo, num *survey* por entrevistas, você poderia tornar pública a renda dos respondentes, mas garantiu que não o faria.

Várias técnicas podem ajudá-lo a manter a promessa. Para começar, os entrevistadores e outros com acesso à identificação dos respondentes devem ser treinados sobre suas responsabilidades éticas. Assim que possível, todos os nomes e/ou endereços devem ser removidos dos questionários e substituídos por números de identificação. Um arquivo-mestre de identificação deve ser criado, ligando números a nomes para permitir a correção posterior de informação faltante ou contraditória, mas este arquivo não deve ficar disponível, exceto por razões legítimas.

Sempre que um *survey* for confidencial, em vez de anônimo, você é responsável por torná-lo claro ao respondente. Usar o termo "anônimo" significando "confidencial" não deve ser tolerado.

Identidade Inferida

Mesmo num *survey* realmente anônimo, às vezes é possível identificar um respondente, principalmente em questões abertas. Se ele listar sua ocupação como "presidente da Companhia ABC", é óbvia sua identidade. Algumas vezes, a análise multivariada de questões fechadas permite a identificação.

Essa possibilidade sempre existe em *survey*, anônimo ou não; não é possível eliminá-la completamente. Mas você nunca deve tentar fazer tais identificações, e certifique-se de que sua equipe de pesquisa também não o faça. Além disso, não relate dados agregados de tal modo que os leitores possam fazê-lo. Por isso, o Departamento do Censo Norte-Americano não relata dados agregados contendo menos do que quinze casos por célula numa tabela.

Identificação Escondida

Ocasionalmente, se garante anonimato aos respondentes, mas na verdade eles são identificáveis. Algumas vezes, o endereço de retorno do escritório da pesquisa tem um "número de caixa postal" diferente para cada respondente. Outras vezes, os pesquisadores põem números de identificação embaixo dos selos colados nos envelopes de retorno. Provavelmente, alguns pesquisadores já escreveram números com suco de limão.

Suspeito que virtualmente em todos estes casos — e eles são provavelmente poucos — os pesquisadores tentaram manter o sigilo dos dados e não fizeram nada que prejudicasse os respondentes. (Tenho a impressão que os pesquisadores mais convencidos de sua "moralidade" pessoal são os mais dispostos a fazer isto.) Qualquer que seja a motivação ou o valor científico, tais práticas são inaceitáveis. Se um *survey* for confidencial, diga. Não minta. Finalmente, todos os *surveys* devem ser pelo menos confidenciais.

Identificação da Finalidade e do Patrocinador

Muitas vezes, enfrentamos o dilema de, se o respondente souber a finalidade e/ou o patrocinador da pesquisa, isso poder afetar as respostas que dará. Em casos mais extremos, pode afetar a probabilidade de cooperação. Se o entrevistador se apresentar como fazendo um estudo sobre preconceito, podemos inferir que o entrevistado tomará cuidado para não parecer preconceituoso. Se o patrocinador do estudo for identificado com a previdência social, as respostas de quem recebe benefícios da previdência podem ser modificadas para se prevenir contra a perda dos benefícios. Se a universidade local tiver má reputação entre a população — digamos, como resultado de demonstrações estudantis —, um *survey* patrocinado por ela pode sofrer oposição e produzir baixa taxa de respostas.

Quase qualquer especificação de finalidade e/ou patrocinador terá algum efeito sobre as taxas de respostas e sobre as respostas dos participantes, afetando a qualidade científica dos dados e as conclusões a serem tiradas. Ainda assim, enganar o respondente levanta uma importante questão ética.

Minha opinião é que você deve ser honesto com os respondentes sobre o patrocinador do estudo. Se ele estiver sendo feito por um departamento de pesquisa da universidade para o governo estadual, isto deve ser bem enfatizado na apresentação do entrevistador (ou na carta de introdução que acompanha um questionário auto-administrado). Além disso, me preocupam especialmente as organizações fictícias — nomes de escritórios de pesquisa criados somente para um estudo em particular.

Honestidade a este respeito não é apenas mais ético, mas também prático. Sempre que você tentar enganar os respondentes com relação aos patrocinadores da pesquisa, corre o risco de a verdade ser conhecida durante o estudo. Nesse caso, isto levará o estudo a sofrer mais do que se fosse feito abertamente por uma agência ou patrocinador de má reputação.

Honestidade sobre o propósito do estudo é mais difícil, porque a pesquisa freqüentemente servirá a vários propósitos analíticos. Portanto, você não pode ser completamente honesto com os respondentes, já que não pode antecipar completamente as finalidades do *survey*. Uma declaração precisa sobre os propósitos básicos do *survey* provavelmente afeta as respostas muito mais do que a descoberta do patrocinador.

Em vista disto, tendo a ser um pouco mais liberal em relação ao propósito do *survey* do que em relação ao patrocínio. As seguintes diretrizes parecem boas: (1) Você não deve revelar nada sobre o propósito do estudo que possa afetar a confiabilidade das respostas. (2) Ao mesmo tempo, deve dizer aos respondentes o que puder sobre os propósitos, quando tal informação não tiver muita chance de afetar as respostas. (3) As explicações sobre o propósito devem ser mais genéricas do que específicas. (4) Não dê razões fictícias para o estudo.

Análise e Relatório

Assim como você tem obrigações éticas para com os respondentes, tem também obrigações éticas para com seus leitores da comunidade científica. Estas não são sempre consideradas no domínio da ética, portanto, faço alguns comentários.

Ao fazer um *survey* rigoroso, você conhece melhor as deficiências técnicas dele do que qualquer outra pessoa e deve informá-las. Se, próximo à conclusão do estudo, você descobrir que algum subgrupo da população foi omitido da amostra, torne este fato conhecido do leitor. Se você acreditar que vários respondentes não compreenderam bem um item do questionário, informe isto aos leitores também. Quaisquer defeitos no desenho do estudo ou da análise que possa ter efeito sobre as conclusões devem ser informados abertamente.

As descobertas negativas devem ser relatadas, se estiverem relacionadas à análise. Infelizmente, persiste um mito de que, em relatórios científicos, somente as descobertas positivas devem ser relatadas (os editores de periódicos algumas vezes também são culpados por isso). Entretanto, do ponto de vista da comunidade científica, saber que duas variáveis não estão relacionadas uma com a outra freqüentemente é tão importante quanto saber que estão. Algumas vezes, como na discussão do *The American Soldier* do Capítulo 15, a falta das correlações esperadas pode ser ainda mais útil. Não há razão para ficar envergonhado pela descoberta de que duas variáveis não estão relacionadas entre si.

Igualmente, tome cuidado para não cair vítima da tentação de “livrar a cara” descrevendo descobertas empíricas como produtos de uma estratégia analítica pré-planejada, quando isso não for o caso. Muitas descobertas acontecem por acaso, mesmo que pareçam óbvias em retrospecto. Decorar tais eventos

com descrições de hipóteses fictícias não é apenas desonesto, mas também tende a enganar os pesquisadores inexperientes e levá-los a achar que toda pesquisa científica é pré-planejada.

Se aparecer uma associação inesperada entre variáveis, ela deve ser apresentada como inesperada. Se toda a estratégia analítica for radicalmente reestruturada durante o estudo, o leitor deve ficar sabendo. Essa revelação é benéfica, porque, assim, outros pesquisadores saberão que outra estratégia, aparentemente funcional, não é apropriada para o estudo em questão. Em geral, a ciência progride através da honestidade e é atrasada pelos enganos motivados pelo ego.

Um Código de Ética Profissional

Uma vez que as questões éticas na pesquisa social são importantes e ambíguas, a maioria das associações profissionais criou e publicou códigos formais de conduta que descrevem o que é considerado um comportamento profissional aceitável e o que é inaceitável. Como ilustração, apresento na Figura 19-1 o código de conduta da Associação Americana de Pesquisa de Opinião Pública (AAPOP), uma associação interdisciplinar de pesquisadores de *survey* acadêmicos e comerciais.

FIGURA 19-1

Código de ética e práticas profissionais

Nós, membros da Associação Americana de Pesquisa de Opinião Pública, subscrevemos os princípios descritos no seguinte código. Nossos objetivos são apoiar uma prática ética bem fundada na feitura da pesquisa de opinião pública e no uso desta pesquisa para fazer política e se tomar decisões nos setores públicos e privados, assim como melhorar a compreensão pública dos métodos de pesquisa de opinião pública e o uso adequado dos resultados de tais pesquisas.

Nos comprometemos a manter altos padrões de competência científica e integridade na condução, análise e relatório do nosso trabalho, em nossas relações com os respondentes do *survey*, com nossos clientes, com os usuários da pesquisa para propósitos decisórios e com o público em geral. Também nos comprometemos a rejeitar todas as tarefas e designações que requeiram atividades inconsistentes com os princípios deste código.

O CÓDIGO

I. Princípios da Prática Profissional na Condução do Nosso Trabalho

A. Devemos ter o cuidado adequado no desenvolvimento de desenhos de pesquisa e instrumentos de *survey*, na coleta, no processamento e na análise de dados, dando todos os passos necessários para garantir a confiabilidade e a validade dos resultados.

1. Devemos recomendar e empregar somente os instrumentos e métodos de análises que, em nosso juízo profissional, sejam adequados para a pesquisa.
2. Não devemos escolher instrumentos de pesquisa e métodos de análises por sua capacidade de levar a conclusões enganosas.
3. Não devemos, conscientemente, interpretar resultados de pesquisas, nem tacitamente permitir interpretações inconsistentes com os dados disponíveis.
4. Não devemos, conscientemente, atribuir às interpretações mais confiança do que a que os dados verdadeiramente merecem.

B. Devemos descrever nossos dados e descobertas de forma precisa e apropriadamente detalhados em todos os relatórios de pesquisa, seguindo os padrões de exposição mínima especificados na Seção III.

C. Se algum dos nossos trabalhos for objeto de investigação formal devido a denúncia de violação deste Código, executada com a aprovação do Conselho Executivo da AAPOP, devemos fornecer informação adicional sobre a pesquisa com o detalhamento necessário, para que um colega pesquisador possa fazer uma avaliação profissional da pesquisa.

II. Princípios de Responsabilidade Profissional em Nossas Relações com as Pessoas.

A. O Público:

1. Se tivermos conhecimento do aparecimento em público de sérias distorções de nossa pesquisa, devemos publicamente revelar o que for necessário para corrigir tais distorções, incluindo uma declaração na imprensa, no corpo legislativo, na agência reguladora, ou outro grupo apropriado, ante o qual foram apresentadas as descobertas distorcidas.

B. Clientes ou Patrocinadores:

1. Ao trabalhar para um cliente particular, manteremos confidencial toda informação de sua propriedade obtida sobre o cliente e sobre a condução e descobertas da pesquisa feita para o cliente, exceto quando a divulgação da informação for expressamente autorizada pelo cliente, ou quando a divulgação se tornar necessária nos termos da Seção I-C ou II-A deste Código.
2. Conheceremos as limitações de nossas técnicas e capacidades e aceitaremos somente os trabalhos que pudermos razoavelmente esperar realizar dentro dessas limitações.

C. A Profissão:

1. Reconhecemos nossa responsabilidade em contribuir para a ciência da pesquisa de opinião pública e em disseminar o mais livremente possível as idéias e as descobertas que surgirem de nossa pesquisa.
2. Não mencionaremos nossa filiação à AAPOP como evidência de competência profissional, uma vez que a Associação não certifica pessoas ou organizações.

D. O Respondente:

1. Nos esforçaremos para evitar o uso de práticas ou métodos que possam prejudicar, humilhar ou seriamente enganar os respondentes de um *survey*.
2. A menos que o respondente abra mão da confidencialidade para usos especificados, manteremos sigilo sobre toda informação que possa levar à identificação

do respondente por suas respostas. Também não divulgaremos ou usaremos os nomes dos respondentes para fins que não sejam os da pesquisa, a menos que os respondentes nos dêem permissão para fazê-lo.

III. Padrão para Divulgação Mínima

Boas práticas profissionais obrigam todos os pesquisadores de opinião pública a incluir, em qualquer relatório de resultados de pesquisa, algumas informações essenciais sobre como a pesquisa foi feita. Pelo menos os seguintes itens devem ser divulgados:

1. Quem patrocinou a pesquisa e quem a conduziu.
2. A redação exata das perguntas feitas, incluindo o texto de qualquer instrução ou explicação precedente para o entrevistador ou para os respondentes que possam afetar a resposta.
3. Uma definição da população estudada e uma descrição da moldura da amostragem usada para identificar esta população.
4. Uma descrição do procedimento de seleção da amostra, fornecendo uma indicação clara do método pelo qual os respondentes foram selecionados ou se estes foram inteiramente auto-selecionados.
5. Tamanho das amostras e, se aplicável, taxas de resposta e informação sobre os critérios de aceitação e procedimentos de triagem.
6. Uma discussão da precisão das descobertas, incluindo, se apropriado, estimativas de erros de amostragem e uma descrição de qualquer procedimento de ponderação ou de estimativa usado.
7. Quais resultados se baseiam em partes da amostra, ao em vez de na amostra total.
8. O método, o local e as datas da coleta de dados.

FONTE: American Association for Public Opinion Research, 1989. Usado com permissão.

Ética - Ilustrações Relevantes

A ética da pesquisa de *survey*, ou de qualquer pesquisa científica, não é muito bem definida. Neste capítulo, indiquei relativamente poucas diretrizes claras, e a maior parte delas está sujeita a debates, porque representam minha própria orientação pessoal.

Minha maior preocupação é *sensibilizá-lo* para as questões éticas na pesquisa de *survey*. Ser capaz de reconhecer as questões éticas em situações reais de pesquisa é muito mais importante do que decorar algumas normas éticas. Pensando nisso, concluímos o capítulo descrevendo várias situações de pesquisa, a maioria delas real e algumas hipotéticas. A maioria delas pode ser debatida honestamente. Você deve identificar as questões éticas dessas situações e considerar os passos que podem ser dados para que a pesquisa seja a mais ética possível.

1. Como instrutor em pesquisa de mercado, você pede a seus alunos que completem questionários que usará na análise de um problema de pesquisa que lhe interessa.
2. Num estudo sobre as atitudes de recém-formados em Direito, você faz um acordo com a Ordem dos Advogados do Estado para incluir um questionário nas provas da Ordem. Completar o questionário seria uma exigência para o credenciamento.
3. Você concorda em fazer um *survey* para uma prefeitura. Como o governo municipal não está sendo bem aceito pela população, você instrui os entrevistadores a dizerem que estão fazendo a pesquisa para a Universidade.²
4. Sua análise dos dados produziu tantas surpresas que suas hipóteses iniciais foram totalmente alteradas pelas descobertas que apareceram em sua análise preliminar. As conclusões finais são tais que você se envergonha de não ter começado com hipóteses apropriadas. Para se salvar, você escreve seu relatório como se o tivesse.
5. Após um *survey* por entrevistas sobre comportamento desviante, policiais lhe ordenam que identifique os

entrevistados que relataram haver participado do saque a lojas durante um tumulto recente. Para não agir como cúmplice, você cumpre a ordem.

6. Ao completar sua análise, você descobre que vinte e cinco das duas mil entrevistas foram falsificadas pelos entrevistadores. Você decide ignorar este fato em seu relatório.
7. Uma pessoa que não foi selecionada para a amostra do *survey* entra em contato com você e insiste em ser entrevistada. Você faz a entrevista e depois joga o questionário fora.
8. Você obtém uma lista dos assinantes da *National Review* para um estudo sobre comportamento do consumidor. Os questionários são enviados para esse grupo com a explicação de que eles foram "selecionados aleatoriamente".
9. A compatibilidade racial é considerada essencial para um estudo sobre o preconceito racial. Entretanto, isso significa que, na grande cidade sendo estudada, os entrevistadores da minoria estarão trabalhando em condições geralmente piores do que os entrevistadores brancos.
10. Num *survey* sobre educação superior, você quer examinar os efeitos de vários fatores de *background* sobre o desempenho acadêmico. Para medir o desempenho, você obtém a média dos alunos na diretoria da universidade.
11. Garante-se aos respondentes que o questionário que estão respondendo é anônimo. Na verdade, foi escondido um número de série no questionário, para permitir a análise de outras informações coletadas sobre os respondentes a partir de outras fontes.
12. Num estudo sobre comportamento sexual, você deseja superar a relutância dos respondentes em relatar o que eles podem considerar comportamento aberrante. Então,

você usa o seguinte item: "Todos se masturbam de vez em quando; com qual frequência você se masturba?"

13. É dito aos respondentes de um *survey* que o estudo está sendo feito para determinar como as pessoas se sentem com relação a uma série de questões públicas. Na verdade, você está interessado em determinar as fontes de oposição a um determinado assunto.

14. Você descobre que 85% do corpo discente de uma universidade fumam maconha regularmente. A publicação dessa descoberta com certeza provocará muito debate na comunidade. Já que você não planeja analisar detalhadamente o uso de drogas, decide ignorar o fato.

15. Você assina um contrato para fazer um estudo e preparar um relatório para um patrocinador sobre certo tema. Você descobre que os dados dão uma oportunidade para examinar um tema relacionado, mas que não é de interesse do patrocinador. Você usa a verba do projeto para cobrir os custos da análise e a preparação do trabalho, que você apresenta na reunião de uma associação profissional.

16. Para testar até que ponto os respondentes tentariam salvar a cara expressando atitudes sobre questões das quais estão mal informados, você pede atitudes sobre uma questão fictícia.

17. Um questionário de pesquisa circula entre os estudantes como parte da documentação para a matrícula na universidade. Apesar de não se dizer aos alunos que eles devem preencher o questionário, espera-se que eles acreditem que devem, garantindo uma maior taxa de resposta.

Resumo

Os requisitos da pesquisa *científica* exigem muita imaginação e criatividade do pesquisador. As restrições administrativas e práticas freqüentemente complicam ainda mais as coisas, impedindo procedimentos ideais de pesquisa. As questões éticas podem significar um peso extra em suas costas. Ao descobrir que não há tempo nem dinheiro suficiente para fazer o melhor estudo possível, você pode chegar a um compromisso brilhante, administrativamente viável e cientificamente adequado, para depois descobrir que ele viola princípios éticos.

Como repeti em todo o livro, a boa pesquisa científica é freqüentemente difícil de ser executada. A pesquisa científica ética pode ser ainda mais difícil, mas você não pode desconsiderar essas questões. Você deve fazer pesquisa cientificamente adequada, administrativamente viável e eticamente defensável. Você não deve violar os direitos e o bem-estar dos cidadãos tentando ajudá-los.

Notas

¹ Quando o livro foi primeiramente publicado em 1973, vários revisores se opuseram à inclusão de um capítulo sobre ética em pesquisa. Em resposta aos meus comentários sobre "os direitos dos respondentes", lembro-me de um revisor que disse: "E sobre os direitos da ciência?" Entretanto, desde então penso que nenhum manual sobre pesquisa social foi publicado sem alguma abordagem substancial das questões éticas. Em minha opinião, isso é um sinal saudável.

² Lembre-se do Capítulo 9 em que a pesquisa metodológica mostrou que o patrocínio de pesquisas pela universidade aumenta a taxa de respostas.

Leituras Adicionais

BEAUCHAMP, Tom L., FADEN, Ruth R., WALLACE JR., Jay, WALTERS, LeRoy. *Ethical Issues in Social Science Research*. Baltimore: The Johns Hopkins Press, 1982.

DIENER Edward, CRANDALL, Rick. *Ethics in Social and Behavioral Research*. Chicago: University of Chicago Press, 1978.

- GOLLIN, Albert E. In Search of a Useable Past. *Public Opinion Quarterly*, v.49, p.414-424, outono 1985.
- MACRAE JR., Duncan. *The Social Function of Social Science*. New Haven, CT: Yale University Press, 1976.
- MCGUIRE, Steven. A Feminist Ethic for Science. *Humanity and Society*, p.461-467, novembro 1984.
- RICHTER JR., Maurice. *Science as a Cultural Process*. Cambridge, MA: Schenkman, 1972.
- ROPER, Burns W. Some Things That Concern Me. *Public Opinion Quarterly*, v.48, p.303-309, outono 1983.
- SJOBERG, Gideon. *Ethics, Politics, and Social Research*. Cambridge, MA: Schenkman, 1967.

Capítulo 20

O Consumidor Informado de Pesquisa de *Survey*

Pode ter havido momentos durante a leitura deste livro em que você achou que o material não era apropriado para você e pode até achar difícil acreditar agora que um dia será um pesquisador de *surveys* profissional. Provavelmente você está certo. Em bases per capita, não há tantos pesquisadores de *survey* assim no mundo.

São pequenas as chances ($p < 0,001$) de, algum dia, você ser responsável pelo desenho e execução de um *survey* probabilístico de área, de múltiplas etapas, sobre a força de trabalho da nação, para determinar se há uma relação estatisticamente significativa entre gênero e renda entre subgrupos emparelhados de trabalhadores assalariados. Mas você sabe, basicamente, o que isso significa e poderia dizer algo sobre os vários passos de um estudo como esse: desenho do questionário, amostragem, coleta de dados, cálculo das rendas médias para os empregados em horário integral ou meio-horário etc. Mesmo se precisar rever este e outros livros para refrescar a memória em alguns detalhes, profissionais fazem isso o tempo todo.

Talvez você nunca dirija sozinho um grande *survey*, mas será capaz de avaliar as pesquisas de outros. Você se tornou o que é chamado, algumas vezes, de "consumidor informado". Não é uma posição insignificante, porque você estará exposto à pesquisa de *survey* quase todos os dias para o resto de sua vida. Você lerá relatórios de *surveys* em jornais e revistas, escutará sobre pesquisas de *survey* no rádio e na televisão, e os políticos e industriais lhe dirão o que a "pesquisa

mostra". Portanto, é importante que você seja capaz de se proteger e de separar o joio do trigo se, algum dia, for plantar seu próprio trigo.

Neste capítulo, dou algumas dicas de consumidor para você usar na revisão de relatórios de pesquisas de *survey*. Em parte, essas diretrizes servem como uma revisão de quase tudo o que foi dito antes. Sinta-se à vontade para usar essa seção como uma revisão do livro, mas veja que minha intenção principal é reforçar sua capacidade de avaliar pesquisas de *survey*, e não sua capacidade de desenhá-la e executá-la. Por isso, pulei vários tópicos do livro e me concentrei apenas nos tópicos para os quais são pertinentes as habilidades do consumidor. Contudo, organizarei o guia do consumidor de modo a seguir mais ou menos as mesmas linhas do resto do livro.

Sempre que você se deparar com o que o "survey mostra", poderá fazer as perguntas listadas a seguir. Nem todas as perguntas se aplicam a todos os *surveys*, mas a lista lhe proporciona o suficiente para que você possa escolher.

Desenho da Pesquisa

- Qual foi a finalidade do estudo: exploração, descrição, explicação ou uma combinação? O desenho da pesquisa foi adequado à sua finalidade?
- Quem conduziu a pesquisa? Quem pagou por ela? O que motivou o estudo? Se as conclusões corresponderem aos interesses do patrocinador ou do pesquisador, isso não as desqualifica, mas ainda assim você deve ficar bem atento.
- Qual foi a unidade de análise? Ela foi apropriada à finalidade do estudo? As conclusões tiradas da pesquisa são apropriadas para a unidade de análise? Por exemplo, os pesquisadores pesquisaram domicílios e terminaram fazendo afirmações sobre indivíduos, ou vice-versa?
- Este estudo é interseccional ou longitudinal? Fique atento às afirmações longitudinais feitas a partir de observações interseccionais.

- Se foram coletados dados longitudinais, certifique-se de que tenham sido feitas medidas comparativas em cada ponto do tempo. As mesmas perguntas foram feitas em cada vez?

- Se foi feito um estudo de painel, verifique o número de pessoas que saíram ("esgotamento de painel") durante o estudo.

Medição

- Quais são os nomes dos conceitos sob estudo?

- Os pesquisadores delinearão diferentes dimensões das variáveis, quando preciso?

- Estas distinções foram mantidas durante a análise e o relatório?

- Quais indicadores foram escolhidos como medidas daquelas dimensões e conceitos? Cada indicador é uma medida válida do que pretende medir? De que mais poderia o indicador ser uma medida? É uma medida confiável? A confiabilidade foi testada?

- As medidas das variáveis correspondem à maneira como outros pesquisadores as mediram? Se não, os pesquisadores atuais têm boas razões para inovar?

- Exatamente como são redigidas as perguntas do questionário? Tome cuidado com pesquisadores que parafraseiam as perguntas.

- Todas as perguntas são feitas de forma clara e não ambígua? Elas poderiam ter sido mal interpretadas pelos respondentes? Se sim, o que poderiam significar as respostas dadas, além do que os pesquisadores presumiram?

- Os respondentes são capazes de responder as perguntas feitas? Se não, eles podem responder assim mesmo, mas as respostas podem não significar coisa alguma.

- Será que alguma das questões tem dupla resposta? Procure conjunções como *e* e *ou*. Está sendo perguntado aos respondentes se eles concordam ou discordam de duas idéias, quando eles podem preferir concordar com uma e discordar da outra?
- As perguntas têm termos negativos? Se sim, os respondentes podem ter interpretado mal as perguntas e respondido inadequadamente.
- Há perigo de desejabilidade social em alguma das perguntas? Será alguma resposta tão certa ou tão errada que os respondentes possam ter respondido com base no que as pessoas pensariam deles?
- Como regra geral, é bom testar todos os itens do questionário, fazendo as perguntas a si mesmo, para ver como você as responderia. As dificuldades que você tiver em responder podem também se aplicar a outros. A seguir, tente supor diferentes pontos de vista (por exemplo, liberal versus conservador, religioso versus irreligioso) e pergunte como as questões soariam para alguém com esse ponto de vista.
- Quais são os níveis de medição — nominal, ordinal, intervalos ou razão — das variáveis? Eles são os níveis apropriados?
- Se foram feitas perguntas fechadas, as categorias de respostas fornecidas foram apropriadas, em número suficiente e mutuamente excludentes?
- Se foram feitas perguntas abertas, como foram categorizadas as respostas? Os pesquisadores se protegeram contra seus próprios vieses na codificação das respostas?
- Foram usadas medições compostas, por exemplo, índices, escalas e tipologias? Se sim, elas são apropriadas à finalidade do estudo? Elas foram construídas corretamente?
- Fazer amostragem era adequado ou deveriam ter sido estudados todos os elementos?
- Foi feita uma amostra probabilística? Se não, tome cuidado com as conclusões tiradas do estudo, principalmente qualquer afirmação descritiva.
- Sobre qual população os pesquisadores desejam tirar conclusões?
- Qual moldura de amostragem foi usada? Ela é uma representação adequada da população que interessa ao pesquisador? Quais elementos da população foram omitidos da moldura da amostragem, e quais elementos estranhos a ela foram incluídos?
- Quais técnicas específicas de amostragem foram empregadas: amostragem aleatória simples, amostragem sistemática ou amostragem por conglomerado? Os pesquisadores estratificaram a moldura da amostragem antes de amostrar? As variáveis de estratificação foram escolhidas sabiamente? Ou seja, elas são relevantes para as variáveis em estudo?
- Qual o tamanho da amostra selecionada? Qual foi a taxa de respostas? Isto é, qual porcentagem da amostra que respondeu? Há qualquer diferença provável entre os que responderam e os que não responderam?
- Mesmo supondo que os respondentes são representativos dos selecionados para a amostra, qual é o erro de amostragem a ser esperado de uma amostra desse tamanho? Os pesquisadores informaram o erro de amostragem?
- Os pesquisadores testaram a representatividade, comparando, por exemplo, a distribuição de sexo entre a população e os respondentes, ou suas idades, etnia, escolaridade ou renda?

- E, finalmente, os indivíduos estudados (ou outra unidade de análise) representam a maioria da população da qual foram selecionados? Ou seja, as conclusões tiradas da amostra nos dizem algo sobre uma população significativa, incluindo os seres humanos em geral?

Análise dos Dados

- Quais técnicas estatísticas foram usadas na análise dos dados? Elas são apropriadas aos níveis de medição das variáveis envolvidas?
- Os pesquisadores fizeram todas as análises relevantes? Todas as variáveis apropriadas foram examinadas? Teria sido importante, por exemplo, replicar os resultados separadamente para homens e mulheres?
- É possível que a correlação observada entre duas variáveis tenha sido causada por uma terceira variável antecedente, tornando espúria a relação observada?
- Foram usados testes de significância estatística? Se sim, foram interpretados corretamente? A significância estatística foi confundida com significância substantiva?
- Alguma descoberta da pesquisa realmente faz diferença? É importante? Alguma diferença observada entre subgrupos, por exemplo, é uma diferença grande e significativa? Há qualquer implicação para ação? Nem todas as descobertas precisam ter implicações para a ação, mas esta implicação é uma medida para avaliar pesquisas.
- Os pesquisadores foram muito além das descobertas feitas ao tirarem conclusões e implicações?
- Existem falhas lógicas na análise e interpretação dos dados?

Relatório dos Dados

- Os pesquisadores situaram esse projeto no contexto de pesquisa anterior sobre o tema? Esta pesquisa acrescenta, modifica, replica ou contradiz estudos anteriores?
- Em geral, os pesquisadores relataram exaustivamente os detalhes do desenho e da execução do estudo? Há partes do relatório particularmente vagos ou incompletos?
- Os pesquisadores relataram falhas ou limitações no desenho e na execução do estudo? Há sugestões para melhores pesquisas sobre esse tópico no futuro?

Confio que esta lista de checagem possa ajudá-lo na revisão e avaliação dos *surveys* que você encontrar em suas atividades diárias. Espero que você tenha aproveitado esse livro e que ele o tenha ajudado na sua aprendizagem. Tenho um gosto especial pela pesquisa de *survey*, e foi um grande prazer compartilhar com você essa poderosa técnica de pesquisa.

Leituras Adicionais

- BABBIE, Earl. *Social Research for Consumers*. Belmont, CA: Wadsworth, 1982.
- FELLIN, Phillip, TRIPODI, Tony, MEYER, Henry J. (Ed.). *Exemplars of Social Research*. Itasca, IL: F. E. Peacock Publishers, 1969.
- GOLLIN, Albert E. Polling and the News Media. *Public Opinion Quarterly*, v.51, p.S86-S94, inverno 1987.
- GROVES, Robert M. Research on Survey Data Quality. *Public Opinion Quarterly*, v.51, p.S156-S172, inverno 1987.
- HUFF, Darrell. *How to Lie with Statistics*. New York: W. W. Norton, 1954.
- JAFFE, A. J., SPIRER, Herbert F. *Misused Statistics: Straight Talk for Twisted Numbers*. New York: Marcel Dekker, 1987.
- LAZARFELD, Paul F., SEWELL, William H., WILENSKY, Harold L. (Ed.). *The Uses of Sociology*. New York: Basic Books, 1967.
- SUDMAN, Seymour, BRADBURN, Norman. The Organizational Growth of Public Opinion Research in the United States. *Public Opinion Quarterly*, v.51, p.S67-S78, inverno 1987.

TRIPODI, Tony, FELLIN, Phillip, MEYER, Henry J. *The Assessment of Social Research*. Itasca, IL: F. E. Peacock Publishers, 1969.

B I B L I O G R A F I A

- ALMOND, Gabriel, VERBA, Sidney. *The Civic Culture*. Princeton, NJ: Princeton University Press, 1963.
- ANDERSON, Barbara A., SILVER, Brian D., ABRAMSON, Paul R. The Effects of the Race of the Interviewer on Race-Related Attitudes of Black Respondents in SRC/CPS National Election Studies. *Public Opinion Quarterly*, v.52, p.289-324, outono 1988.
- BABBIE, Earl R. *Science and Morality in Medicine*. Berkeley, CA: University of California Press, 1970.
- BABBIE, Earl. *Observing Ourselves: Essays in Social Research*. Belmont, CA: Wadsworth, 1986.
- BABBIE, Earl. *Social Research for Consumers*. Belmont, CA: Wadsworth, 1982.
- BABBIE, Earl. *The Practice of Social Research*. Belmont, CA: Wadsworth, 1989.
- BAINBRIDGE, William Sims. *Survey Research: a Computer-Assisted Introduction*. Belmont, CA: Wadsworth, 1989.
- BART, Pauline, FRANKEL, Linda. *The Student Sociologist's Handbook*. Glenview, IL: Scott, Foresman, 1981.
- BEAUCHAMP, Tom L. et al. *Ethical Issues in Social Science Research*. Baltimore: The Johns Hopkins Press, 1982.
- BECKER, Leonard, GUSTAFSON, Clair. *Encounter with Sociology: The Term Paper*. San Francisco: Boyd and Fraser, 1976.
- BEVERIDGE, W. I. B. *The Art of Scientific Investigation*. New York: Vintage Books, 1950.
- BILLIET, Jacques, LOOSVELDT, Geert. Improvement in the Quality of Responses to Factual Survey Questions by Interviewers. *Public Opinion Quarterly*, v.52, p.190-211, verão 1988.

- BISHOP, George F., TUCHFARBER, Alfred J., OLDENDICK, Robert W. Opinions on Fictitious Issues: The Pressure to Answer Survey Questions. *Public Opinion Quarterly*, v.50, p.240-250, verão 1986.
- BLALOCK, Hubert. *Social Statistics*. New York: McGraw-Hill, 1979.
- BRADBURN, Norman M., SUDMAN, Seymour e Sócios. *Improving Interview Method and Questionnaire Designs*. San Francisco: Jossey-Bass, 1979.
- BROWNLEE, K. A. A Note on the Effects of Nonresponse on Surveys. *Journal of the American Statistical Association*, v.52, n.177, p.29-32, 1957.
- BUTTERFIELD, Herbert. *The Origins of Modern Science*. New York: Macmillan, 1960.
- CARMINES, Edward G., ZELLER, Richard A. *Reliability and Validity Assessment*. Beverly Hills, CA: Sage, 1979.
- COCHRAN, William G. *Sampling Techniques*. New York: John Wiley & Sons, 1963.
- COLE, Stephen. *The Sociological Method*. Chicago: Markham, 1972.
- COZBY, Paul C. *Using Computers in the Behavioral Sciences*. Palo Alto, CA: Mayfield, 1984.
- DAVIS, James A. *Elementary Survey Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- DIENER, Edward, CRANDALL, Rick. *Ethics in Social and Behavioral Research*. Chicago: University of Chicago Press, 1978.
- DILLMAN, Don A. *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons, 1978.
- DONALD, Marjorie N. Implications of Nonresponse for the Interpretation of Mail Questionnaire Data. *Public Opinion Quarterly*, v.24, n.1, p.99-114, 1960.
- DURKHEIM, Emile. *Suicide: A Study in Sociology*. New York: Free Press, 1951.
- DURKHEIM, Emile. *The Rules of Sociological Method*. New York: Free Press, 1962.
- EDWARDS, Perry, BROADWELL, Bruce. *Data Processing*. 2nd.ed. Belmont, CA: Wadsworth, 1982.
- FELLIN, Phillip, TRIPODI, Tony, MEYER, Henry J. (Ed.). *Examples of Social Research*. Itasca, IL: F. E. Peacock Publishers, 1969.
- FORSLUND, Morris A. Patterns of Delinquency Involvement: An Empirical Typology. Trabalho apresentado na Reunião Anual da Western Association of Sociologists and Anthropologists, Lethbridge, Alberta, 8 fev. 1980.
- FOWLER, H. W. *A Dictionary of Modern English Usage*. New York: Oxford University Press, 1965.
- FOX, Richard J., CRASK, Melvin R., KIM, Jonghoon. Mail Survey Response Rates. *Public Opinion Quarterly*, v.52, p.467-491, inverno 1988.
- FRANKEL, Martin R., FRANKEL, Lester R. Fifty Years of Survey Sampling in the United States. *Public Opinion Quarterly*, v.51, p.S127-S138, inverno 1987.
- FREEMAN, Linton. *Elementary Applied Statistics*. New York: John Wiley & Sons, 1968.
- GALL, John. *Systemantics: How Systems Work and Especially How They Fail*. New York: Pocket Books, 1975.
- GLOCK, Charles Y. (Ed.). *Survey Research in the Social Sciences*. New York: Russell Sage Foundation, 1967.
- GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967.
- GLOCK, Charles Y., STARK, Rodney. *Christian Beliefs and Anti-Semitism*. New York: Harper & Row, 1966.
- GLOCK, Charles Y., STARK, Rodney. *Religion and Society in Tension*. Chicago: Rand McNally, 1965.
- GOLD, Julius, KOLB, William. *A Dictionary of the Social Sciences*. New York: Free Press, 1964.
- GOLDEN, Patricia (Ed.). *The Research Experience*. Itasca, IL: F. E. Peacock, 1976.
- GOLLIN, Albert E. In Search of a Usable Past. *Public Opinion Quarterly*, v.49, p.414-424, outono 1985.
- GOLLIN, Albert E. Polling and the News Media. *Public Opinion Quarterly*, v.51, p.S86-S94, inverno 1987.
- GORDEN, Raymond L. *Interviewing: Strategy, Techniques, and Tactics*. Homewood, IL: Dorsey Press, 1969.

- GOYDER, John. Face-to-Face Interviews and Mailed Questionnaires: The Net Difference in Response Rate. *Public Opinion Quarterly*, v.49, p.234-252, verão 1985.
- GROVES, Robert M., MATHIOWETZ, Nancy A. Computer Assisted Telephone Interviewing: Effects on Interviewers and Respondents. *Public Opinion Quarterly*, v.48, p.356-369, primavera 1984.
- GROVES, Robert M. Research on Survey Data Quality. *Public Opinion Quarterly*, v.51, p.S156-S172, inverno 1987.
- GRUBER, James, PRYOR, Judith, BERGE, Patricia. *Materials and Methods for Sociology Research*. New York: Neal-Schuman, 1980.
- HAMMOND, Phillip (Ed.). *Sociologists at Work*. New York: Basic Books, 1964.
- HANSEN, M. H., HURWITZ W. N., MADOW, W. G. *Sample Survey Methods and Theory*. New York: John Wiley & Sons, 1953. 2 V.
- HEISE, David (Ed.). *Microcomputers in Social Research*. Beverly Hills, CA: Sage, 1981.
- HENKEL, Ramon E. *Tests of Significance*. Beverly Hills, CA: Sage, 1976.
- HILDEBRAND, David K., LAING, James D., ROSENTHAL, Howard. *Analysis of Ordinal Data*. Beverly Hills, CA: Sage, 1977.
- HIRSCHI, Travis, SELVIN, Hanan. *Principles of Survey Analysis*. New York: Free Press, 1973.
- HOLSTI, Ole. *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley, 1969.
- HORNIK, Jacob, ELLIS, Shmuel. Strategies to Secure Compliance for a Mall Intercept Interview. *Public Opinion Quarterly*, v.52, p.539-551, inverno 1988.
- HUFF, Darrell. *How to Lie with Statistics*. New York: W. W. Norton, 1954.
- HUNT, Morton. *Profiles of Social Research: The Scientific Study of Human Interactions*. New York: Basic Books, 1985.
- HYMAN, Herbert. *Survey Design and Analysis*. New York: Free Press, 1955.
- IRVINE, William. *Apes, Angels, and Victorians*. New York: Meridian Books, 1959.
- IVERSEN, Gudmund R., NORPOTH, Helmut. *Analysis of Variance*. Beverly Hills: Sage, 1976.
- JAFFE, A. J., SPIRER, Herbert F. *Misused Statistics: Straight Talk for Twisted Numbers*. New York: Marcel Dekker, 1987.
- JENDREK, Margaret Platt. *Through the Maze: Statistics with Computer Applications*. Belmont, CA: Wadsworth, 1985.
- KAHN, Robert L., CANNELL, Charles F. *The Dynamics of Interviewing*. New York: John Wiley & Sons, 1967.
- KAPLAN, Abraham. *The Conduct of Inquiry*. San Francisco: Chandler Publishing Co., 1964.
- KIM, Jae-On, MUELLER, Charles W. *Introduction to Factor Analysis*. Beverly Hills, CA: Sage, 1978.
- KISH, Leslie. Chance, Statistics and Statisticians. *Journal of the American Statistical Association*, v.73, n.361, p.1-6, mar. 1978.
- KISH, Leslie. *Survey Sampling*. New York: John Wiley & Sons, 1965.
- KNOKE, David, BURKE, Peter J. *Log-Linear Models*. Beverly Hills, CA: Sage, 1980.
- KROSNIC, Jon A., ALWIN, Duane F. A Test of the Form-resistant Correlation Hypothesis: Ratings, Rankings and the Measurement of Values. *Public Opinion Quarterly*, v.52, p.526-538, inverno 1988.
- KUHN, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1970.
- LABOVITZ, Sanford, HAGEDORN, Robert. *Introduction to Social Research*. New York: McGraw-Hill, 1971.
- LAVRAKAS, Paul J. *Telephone Survey Methods: Sampling, Selection and Supervision*. Beverly Hills, CA: Sage, 1987.
- LAZARSELD, Paul F., PASANELLA, Ann K., ROSENBERG, Morris (Ed.). *Continuities in "The Language of Social Research"*. New York: Free Press, 1972. seções III-IV.
- LAZARSELD, Paul F., ROSENBERG, Morris (Ed.). *The Language of Social Research*. New York: Free Press, 1955. p.15-108.

- LAZARFELD, Paul F., SEWELL, William H., WILENSKY, Harold L. (Ed.). *The Uses of Sociology*. New York: Basic Books, 1967.
- LEWIS-BECK, Michael S. *Applied Regression: An Introduction*. Beverly Hills, CA: Sage, 1980.
- LOFLAND, John. *Analyzing Social Settings*. Belmont, CA: Wadsworth, 1984.
- LOPATA, Helena Znaniecki. Widowhood and Husband Sanctification. *Journal of Marriage and the Family*, p.439-450, maio 1981.
- MACRAE JR., Duncan. *The Social Function of Social Science*. New Haven, CT: Yale University Press, 1976.
- MCCALL, George, SIMMONS, J. L. (Ed.). *Issues in Participant Observation: A Text and Reader*. Reading, MA: Addison-Wesley, 1969.
- MCGUIRE, Steven. A Feminist Ethic for Science. *Humanity and Society*, v.8, p.461-467, nov. 1984.
- MERTON, Robert K. (Ed.). *Sociology Today*. New York: Basic Books, 1959.
- MERTON, Robert K., LAZARFELD, Paul F. (Ed.). *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"*. New York: Free Press, 1950.
- MILLER, Delbert. *Handbook of Research Design and Social Measurement*. New York: Longman, 1983.
- MITCHELL, Arnold. *The Nine American Lifestyles*. New York: Warner Books, 1983.
- MORRISON, Denton, HENKEL, Ramon (Ed.). *The Significance Test Controversy: A Reader*. Chicago: Aldine-Atherton, 1970.
- OKSENBERG, Lois, COLEMAN, Lerita, CANNELL, Charles F. Interviewers Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, v.50, p.97-111, primavera 1986.
- OPPENHEIM, A. N. *Questionnaire Design and Attitude Measurement*. New York: Basic Books, 1966.
- PAYNE, Stanley. *The Art of Asking Questions*. Princeton, NJ: Princeton University Press, 1965.
- PUBLIC OPINION QUARTERLY, v.51, n.4 (Parte 2). Esta edição comemorativa dos 50 anos da publicação (inverno 1987) contém diversos artigos revendo a história das pesquisas de survey.
- RAY, William, RAVIZZA, Richard. *Methods Toward a Science of Behavior and Experience*. Belmont, CA: Wadsworth, 1985.
- REESE, Stephen D. et al. Ethnicity-of-Interviewer Effects Among Mexican-Americans and Anglos. *Public Opinion Quarterly*, v.50, p.563-572, outono 1986.
- REYNOLDS, H. T. *Analysis of Nominal Data*. Beverly Hills, CA: Sage, 1977.
- RICHARDSON, Stephen A. et al. *Interviewing: Its Forms and Functions*. New York: Basic Books, 1965.
- RICHTER JR., Maurice. *Science as a Cultural Process*. Cambridge, MA: Schenkman, 1972.
- ROPER, Burns W. Some Things That Concern Me. *Public Opinion Quarterly*, v.48, p.303-309, outono 1983.
- ROSENBERG, Morris. *The Logic of Survey Analysis*. New York: Basic Books, 1968.
- SCHEAFFER, Richard L., MENDENHALL, William, OTT, Lyman. *Elementary Survey Sampling*. North Scituate, MA: Duxbury Press, 1979.
- SCHUMAN, Howard, PRESSER, Stanley. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context*. New York: Academic Press, 1981.
- SELLTIZ, Claire et al. *Research Methods in Social Relations*. New York: Holt, Rinehart & Winston, 1959.
- SHANKS, J. Merrill, TORTORA, Robert D. Beyond CATI: Generalized and Distributed Systems for Computer-Assisted Surveys, preparado para o Bureau of Census, Primeira Conferência Anual de Pesquisas, Reston, VA, 20-23 de mar. 1985.
- SHARP, Vicki. *Statistics for the Social Sciences*. Boston: Little, Brown, 1979.
- SJOBERT, Gideon. *Ethics, Politics and Social Research*. Cambridge, MA: Schenkman, 1967.
- SMITH, Tom W. The Art of Asking Questions. *Public Opinion Quarterly*, v.51, p.S95-S108, inverno 1987.
- SMITH, Tom. That Which We Call Welfare by Any Other Name Would Smell Sweeter. *Public Opinion Quarterly*, v.51, p.75-83, primavera 1987.
- STEVENS, S. S. On the Theory of Scales of Measurement. *Science*, v.103, p.677-680, 1946.

- STOUFFER, Samuel A. *Communism, Conformity and Civil Liberties*. New York: John Wiley & Sons, 1955.
- STOUFFER, Samuel A. et al. *The American Soldier*. Princeton, NJ: Princeton University Press, 1949.
- STOUFFER, Samuel A. *Social Research to Test Ideas*. New York: Free Press, 1962.
- STRUNK, Jr., William, WHITE, E. B. *The Elements of Style*. New York.: Macmillan, 1959.
- SUDMAN, Seymour, BRADBURN, Norman M. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass, 1982.
- SUDMAN, Seymour, BRADBURN, Norman M.. The Organizational Growth of Public Opinion Research in The United States. *Public Opinion Quarterly*, v.51, p.S67-S78, inverno 1987.
- TOBEN, Bob. *Space-Time and Beyond*. New York: E. P. Dutton, 1975.
- TRIPODI, Tony et al. *The Assessment of Social Research*. Itasca, IL: F. E. Peacock Publishers, 1969.
- TZE-CHUNG Li. *Social Science Reference Sources: A Practical Guide*. Westport, CT: Greenwood Press, 1980.
- WALLACE, Walter. *The Logic of Science in Sociology*. Chicago: Aldine-Atherton, 1971.
- WARNER, Lloyd. *Democracy in Jonesville*. New York: Harper & Row, 1949.
- WATSON, James D. *The Double Helix*. New York: The New American Library, 1968.
- WEEKS, Michael F., KULKA, Richard A., PIERSON, Stephanie A. Optimal Call Scheduling for a Telephone Survey. *Public Opinion Quarterly*, v.51, p.540-549, inverno 1987.
- WHITEHEAD, Alfred North. *Science and the Modern World*. New York: Macmillan, 1925.
- YERG, Beverly J. Reflections on the Use of the RTE Model in Physical Education. *Research Quarterly for Exercise and Sport*, p.42, março de 1981.
- ZEISEL, Hans. *Say It with Figures*. New York: Harper & Row, 1957.

A P Ê N D I C E

- A Tabela de Números Aleatórios
- B Erro Estimado de Amostragem para uma Binomial (Nível de Confiança de 95%)
- C Distribuição do Qui-Quadrado
- D Áreas Normais de Curva



TABELA DE NÚMEROS ALEATÓRIOS

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	16367	39440	53537	71341	57004	00849	74917	97758	16379
37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
77921	06907	06907	06907	27756	53498	18602	70659	90655	15053	21946	81825	44394	42880
99562	72905	72905	72905	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
96301	91977	91977	91977	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
89579	14342	14342	14342	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
85475	36857	36857	36857	53060	59533	38867	63300	08158	17983	16439	11458	18593	64952
28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01554	85590	91610	78188
63553	40961	48232	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	81259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64839	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	76989	13602	51851	46104	88916	19509	25625	58104

FONTE: Abreviado do *Handbook of Tables for Probability and Statistics*. Second Edition. Cleveland: The Chemical Rubber Company, 1968. Edited by William H. Beyer. Reproduzido com permissão dos editores, The Chemical Rubber Company.

(continua)

48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
54164	58492	22421	74103	47070	25306	74468	26384	58151	06646	21524	15227	96909	44592
32639	32363	05597	24200	13263	38005	94342	28728	35806	06912	17012	64161	18296	22851
29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	35188	18510
02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
81525	72295	04839	96423	24878	82651	66566	14788	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	37937	63904	45776	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	62775	35605	81263	39667	47358	56873	56307	61607	49518	89656	20103	77490	18062
98427	07523	33362	64270	01638	92477	66969	89420	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	46151	22923	32261	85653

(continua)

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	19127	63318	29686	03387	59846
31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325
78919	19474	23632	27889	47917	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	93197	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

(continua)

82486	84846	99254	67632	43218	50076	21368	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
06299	01121	05418	38982	55758	92237	26759	86367	21216	98444	08303	56613	91511	75928
79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076	72811	22717	50585
85613	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
52608	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	14196	00921	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	74239
46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	53203	18098	47625	88684
03237	45430	55417	63282	90816	17345	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

B

ERRO ESTIMADO DE AMOSTRAGEM PARA UMA BINOMIAL (NÍVEL DE CONFIANÇA DE 95%)

Como usar esta tabela: encontre a interseção entre o tamanho da amostra e a distribuição percentual aproximada da binomial na amostra. O número que aparece na interseção representa o erro amostral estimado, ao nível de 95% de confiança, expresso em pontos percentuais (mais ou menos).

Exemplo: Numa amostra de 400 respondentes, 60% respondem sim e 40% respondem não. O erro amostral é estimado a mais ou menos 4,9 pontos percentuais. O intervalo de confiança, portanto, está entre 55,1% e 64,9%. Estimariamos (95% de confiança) que a proporção da população total que diria sim está em algum lugar dentro do intervalo.

Distribuição Percentual Binomial

Tamanho da Amostra	Distribuição Percentual Binomial				
	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
1.000	3.2	3.1	2.9	2.5	1.9
1.100	3	3	2.8	2.4	1.8
1.200	2.9	2.8	2.6	2.3	1.7
1.300	2.8	2.7	2.5	2.2	1.7
1.400	2.7	2.6	2.4	2.1	1.6
1.500	2.6	2.5	2.4	2.1	1.5
1.600	2.5	2.4	2.3	2	1.5
1.700	2.4	2.4	2.2	1.9	1.5
1.800	2.4	2.3	2.2	1.9	1.4
1.900	2.3	2.2	2.1	1.8	1.4
2.000	2.2	2.2	2	1.8	1.3

DISTRIBUIÇÃO DO QUI-QUADRADO

gl	Probabilidade						
	.99	.98	.95	.90	.80	.70	.50
1	.03157	.03628	.00393	.0158	.0642	.148	.455
2	.0201	.0404	.103	.211	.446	.713	1.386
3	.115	.185	.352	.584	1.005	1.424	2.366
4	.297	.429	.711	1.064	1.649	2.195	3.357
5	.554	.752	1.145	1.610	2.343	3.000	4.351
6	.872	1.134	1.635	2.204	3.070	3.828	5.348
7	1.238	1.564	2.167	2.833	3.822	4.671	6.346
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343
10	2.558	3.059	3.940	4.869	6.178	7.267	9.342
11	3.053	3.609	4.575	5.578	6.989	8.178	10.341
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338
18	7.015	7.906	9.390	10.865	12857	14.440	17.338
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336
28	13.565	14.847	16.928	18.935	8	23.647	27.336
29	14.256	15.574	17.708	19.768	22.475	22.577	28.336
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336

Para valores mais altos de gl, a expressão $\sqrt{2\chi^2} - \sqrt{2gl-1}$ pode ser usada como um desvio normal com variância unitária, lembrando que a probabilidade do χ^2 corresponde à única cauda da curva normal.

FONTE: Agradeço ao Executor Literário do falecido Sir Ronald A. Fisher, F.R.S., a Frank Yates, F.R.S., e a Longman Group Ltd., London, pela permissão para publicar a Tabela IV do livro *Statistical Tables for Biological, Agricultural, and Medical Research*. (6th.ed., 1974.)

Probabilidade

Df	.30	.20	.10	.05	.02	.01	.001
1	1.074	1.642	2.706	3.841	5.412	6.635	10.827
2	2.408	3.219	4.605	5.991	7.824	9.210	13.815
3	3.665	4.642	6.251	7.815	9.837	11.341	16.268
4	4.878	5.989	7.779	9.488	11.668	13.277	18.465
5	6.054	7.289	9.236	11.070	13.388	15.086	20.517
6	7.231	8.558	10.645	12.592	15.033	16.812	22.457
7	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	17.332	19.311	22.307	24.996	28.259	30.578	37.697
16	18.841	20.465	23.542	26.296	29.633	32.000	39.252
17	15.511	21.615	24.769	27.587	30.995	33.409	40.790
18	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	31.391	34.027	37.916	41.337	45.419	48.278	56.893
29	32.461	35.139	39.087	42.557	46.693	49.588	58.302
30	35.530	36.250	40.256	43.773	47.692	50.892	59.703

D

ÁREAS NORMAIS DE CURVA

ÁREAS NORMAIS DE CURVA										
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	.1945	.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	.4972	.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

FONTE: Abreviado da Tabela I do *Statistical Tables and Formulas* por A. Hald (New York: John Wiley & Sons, 1952). Usado com a permissão de John Wiley & Sons, Inc.

- A**
- amostra** Conjunto de *respondentes* selecionados para estudo de forma a garantir que o que for aprendido sobre os integrantes da amostra também é verdade para a *população* da qual foram selecionados. Ver Capítulos 3 e 5.
- amostra aleatória simples** Tipo de *amostra probabilística* na qual as unidades que formam uma *população* recebem números; gera-se então um conjunto de números aleatórios, e as unidades com estes números são incluídas na amostra. Apesar de a teoria da probabilidade e os cálculos que ela provê suporem este método básico de amostragem, por motivos práticos ele é raramente usado. Uma alternativa equivalente é a *amostra sistemática*, com início aleatório. Ver Capítulo 5.
- amostra bola de neve** Metodologia de *amostragem não-probabilística* freqüentemente usada em pesquisa de campo. Pede-se a cada participante sugerir outras pessoas para serem entrevistadas. Ver Capítulo 5.
- amostra de probabilidade de área** Forma de *amostra por conglomerado em múltiplas etapas*, na qual áreas geográficas, como bairros, servem como *unidade de amostragem* na primeira etapa. Unidades selecionadas na primeira etapa da amostragem são então listadas — todas as residências em cada bairro selecionadas são anotadas após visita ao bairro —, e estas listas são subamostradas. Ver Capítulo 5.
- amostra intencional** Ver *amostra por julgamento* e Capítulo 5.
- amostra MIPs** Método de igual probabilidade de seleção. Desenho de *amostra* em que cada membro da *população* tem a mesma chance de ser selecionado na amostra. Ver Capítulo 5.

- amostra não-probabilística** *Amostra* selecionada de forma não sugerida pela teoria da probabilidade. Exemplos incluem amostras *por julgamento (intencionais)*, *por quota* e *bola de neve*. Ver Capítulo 5.
- amostra por conglomerados em múltiplas etapas** Amostra em múltiplas etapas em que grupos naturais [*clusters*] são inicialmente amostrados, com os membros de cada grupo selecionado sendo em seguida sub-amostrados. Exemplo: selecionar num guia escolar uma amostra de faculdades e universidades, conseguir listas dos alunos em todas as escolas selecionadas e, em seguida, amostrar alunos de cada uma. Este procedimento é discutido no Capítulo 5. Ver ainda *amostra probabilística por área*.
- amostra por cota** Tipo de *amostra não-probabilística* em que unidades são selecionadas para a *amostra* com base em características pré-especificadas, de modo que a amostra total tenha a mesma distribuição de características que se supõe existir na *população* estudada. Ver Capítulo 5.
- amostra por julgamento** Tipo de *amostra não-probabilística* em que você seleciona as unidades a serem observadas com base no seu próprio julgamento sobre quais delas serão mais úteis ou representativas. Também chamada *amostra intencional*. Ver Capítulo 5.
- amostra probabilística** Termo genérico para *amostra* selecionada conforme a teoria da probabilidade, tipicamente com algum mecanismo de *seleção aleatória*. Tipos específicos de amostras probabilística são *amostra probabilística por área*, *amostra MIPS*, *PPT* e *amostra sistemática*. Ver Capítulo 5.
- amostra sistemática** Tipo de *amostra probabilística* no qual cada unidade k^a numa lista é selecionada para inclusão em amostra, por exemplo, cada vigésimo quinto aluno na listada escola. Computa-se o k dividindo-se o tamanho da *população* pelo tamanho desejado da amostra, e k é o *intervalo de amostragem*. Dentro de certas limitações, a amostragem sistemática é o equivalente funcional da *amostragem aleatória simples* e normalmente mais fácil de fazer. Tipicamente, a primeira unidade é selecionada aleatoriamente. Ver Capítulo 5 e a *amostra bola de neve*.
- amostragem de zonas eleitorais selecionadas** Técnica de *amostragem não-probabilística* às vezes usada em pesquisas de opinião política. Baseada em padrões eleitorais passados, a pesquisa estabelece um conjunto de zonas eleitorais que, no todo, parecem representativas dos eleitores em toda a jurisdição estudada. O risco é que as zonas amostradas e/ou toda a jurisdição podem ter mudado o bastante para que a *amostra* já não seja representativa do todo. Ver Capítulo 5.
- amostragem por conglomerado** Ver *amostragem por conglomerado em múltiplas etapas*.

- amostras paralelas** Técnica de desenho de pesquisa em que *amostras* de populações relacionadas são selecionadas para análise contextual. Por exemplo, poderíamos selecionar para estudo uma amostra de estudantes e também seus professores. Ver Capítulo 4.
- amplitude** Medida de *dispersão* composta dos valores mais alto e mais baixo de uma *variável* num conjunto de observações. Na sua turma, por exemplo, a amplitude etária pode ser de 17 a 37. Ver Capítulo 16.
- análise bivariada** Análise simultânea de duas variáveis para determinar a relação empírica entre elas. A montagem de uma tabela percentual simples e a computação de um coeficiente de correlação simples são exemplos de análises bivariadas. Ver Capítulo 14.
- análise de conteúdo** Método de pesquisa baseado na codificação de comunicações registradas, como livros, jornais, discursos, comerciais de TV etc. Ver Capítulos 2 e 5.
- análise de regressão** Método de análise de dados em que as relações entre *variáveis* são representadas na forma de uma equação, chamada *equação de regressão*. Ver Capítulo 17.
- análise de trajetória** Forma de *análise multivariada* na qual as *relações causais* entre *variáveis* são apresentadas em formato gráfico. Ver Capítulo 17.
- análise de variância** Forma de análise de dados na qual a variância de uma *variável dependente* é examinada para toda a *amostra* e para subgrupos separados criados com base em alguma(s) *variável(is) independente(s)*. Ver Capítulo 17.
- análise discriminante** Técnica analítica que adapta a lógica da *análise de regressão* para *variáveis* de nível *nominal*. Ver Capítulo 17.
- análise fatorial** Método algébrico complexo para determinar as dimensões ou fatores gerais existentes num conjunto de observações concretas. Ver Capítulo 17.
- análise multivariada** Análise de relações simultâneas entre diversas *variáveis*. Examinar simultaneamente os efeitos de idade, sexo e classe social sobre religiosidade seria um exemplo de análise multivariada. Ver Capítulos 4, 15 e 17.
- análise quantitativa** Representação e manipulação numérica de observações para descrever e explicar os fenômenos que tais observações refletem. Ver Capítulo 3.
- análise secundária** Forma de pesquisa na qual os dados coletados e processados por um pesquisador são reanalisados por outro — muitas vezes para uma finalidade diferente. Isto é especialmente apropriado nos casos de dados de *survey*. Arquivos de dados são repositórios ou bibliotecas para armazenamento e distribuição de dados para análise secundária. Ver Capítulo 3.

análise univariada Análise de uma só *variável* para fins descritivos. *Distribuições de freqüência, médias, e medidas de dispersão* são exemplos de análise univariada, diferente de análises *bivariada* e *multivariada*. Ver Capítulo 14 e *análise multivariada*.

anonimato Descreve um *survey* no qual a identidade do *respondente* não pode ser determinada por ninguém, nem mesmo pelo diretor do *survey*. (*Nota: surveys* por entrevista raramente são anônimos, já que o entrevistador normalmente conhece a identidade do entrevistado.) Ver Capítulo 19 e *confidencial*.

atributos Características de pessoas ou coisas. Ver *variáveis* e Capítulo 7.

C

carga fatorial Em *análise fatorial*, medidas da *correlação* entre um item e um dos fatores calculados. Ver Capítulo 17.

censo Enumeração das características de uma *população*. Frequentemente, censo se assemelha a *survey*, a diferença sendo que o censo coleta dados de todos os integrantes da população, enquanto o *survey* se limita a uma *amostra*. Ver Capítulos 3 e 5.

codificação Processo através do qual dados brutos são transformados em forma padronizada adequada para processamento e análise em computador. Ver Capítulo 11.

coeficiente de reproducibilidade Medida da extensão pela qual um score de *escala* permite a reconstrução exata dos dados específicos que entraram na montagem da escala. Ver Capítulo 8.

coeficiente de trajetória Medida da associação entre duas *variáveis* numa análise de *trajetória*. Ver Capítulo 17.

conceito Idéia ou imagem mental usada para resumir e representar um conjunto de objetos, experiências ou pensamentos. Por exemplo, o conceito "cadeira" representa um conjunto variado de objetos. Conceitos típicos das ciências sociais são "classe alta" e "alienado". Ver Capítulo 1.

conceitualização Processo mental através do qual noções vagas e imprecisas (*conceitos*) se tornam mais específicas e precisas. Imagine que você quer estudar preconceito. O que quer dizer preconceito? Há tipos diferentes de preconceito? Quais são? Ver Capítulo 7.

confiabilidade Qualidade de métodos de medida que sugere que os mesmos dados teriam sido coletados cada vez em repetidas observações do mesmo fenômeno. No contexto de um *survey*, esperamos que a pergunta "Você foi à igreja na semana passada?" teria mais confiabilidade que a pergunta "Quantas vezes na vida você já foi à igreja?" Não confundir com *validade*. Ver Capítulo 7.

conglomerado [*cluster*] Agrupamento natural de unidades, usado em amostragem por conglomerado em múltiplas etapas. Exemplos: nas Faculdades há conglomerados de alunos; em quarteirões há conglomerados de residências. Ver Capítulo 5.

correlação Termo ambíguo usado para indicar o co-relacionamento ou correspondência entre *variáveis*; quando os valores de uma variável mudam, os valores da outra mudam padronizadamente. Fala-se de correlação positiva quando os valores das duas variáveis aumentam ou diminuem juntos e de correlação negativa quando os valores de uma variável aumentam enquanto os da outra diminuem. O termo é também usado com significado mais específico para se referir à correlação produto-momento *r* de Pearson. Ver Capítulos 1 e 16.

correlação produto-momento Ver *correlação* e Capítulo 16.

D

dedução Modelo lógico no qual expectativas específicas ou *hipóteses* são desenvolvidas com base em princípios gerais. Partindo do princípio de que todos os reitores são durões, você pode antecipar que este reitor não o deixará mudar de curso. Esta antecipação resulta de dedução. Ver *indução* e Capítulo 1.

definição operacional Definição concreta e específica de algo em termos das operações através das quais observações devem ser categorizadas. A definição operacional de "tirar nota A neste curso" poderia ser "responder corretamente pelo menos 90% das questões no exame final". Ver Capítulo 7.

desconfirmabilidade Possibilidade de especificar condições sob as quais uma *hipótese* é considerada falsa. A menos que uma hipótese seja desconfirmável, não pode ser genuinamente testada. Ver Capítulo 1.

dimensão Aspecto de uma *variável*; por exemplo, podemos falar da "dimensão da crença" da variável religiosidade ou da "dimensão de política externa" da variável orientação política. Ver Capítulo 7.

discagem digital aleatória Um aspecto do *ETAC*. Computadores podem ser usados para gerar aleatoriamente conjuntos de números com sete algarismos, constituindo a *amostra* do *survey*, em vez de selecionar nomes e números de assinantes telefônicos. Ver Capítulos 5 e 10.

dispersão Distribuição de valores em torno de algum valor central, como uma *média*. A *amplitude* é um exemplo simples de medida de dispersão. Poderíamos afirmar que a *média* de idade num grupo é de 37,9 anos e que a amplitude varia de 12 a 89 anos. Ver Capítulo 14.

distribuição amostral Já que as características de uma *amostra* se aproximam (mas não perfeitamente) das características da *população* da qual ela foi selecionada, isto se refere às aproximações que se produziram se fosse selecionado um grande número de amostras, cada uma dando uma estimativa desta característica para toda a população. Ver Capítulo 5.

distribuição de frequência Descrição do número de vezes que os vários *atributos* de uma *variável* são observados numa *amostra*. Relatar que 53% de uma amostra são homens e 47% mulheres é um exemplo de distribuição de frequência. Outro exemplo é informar que 15 das cidades estudadas tinham populações abaixo de 10.000, 23, populações entre 10.000 e 25.000, e assim por diante. Ver Capítulo 14.

E

elemento Em amostragem, os elementos são as unidades que compõem a população, como indivíduos ou eleitores. Ver Capítulo 5.

entrevista Encontro para coleta de dados, no qual uma pessoa (um entrevistador) faz perguntas a outra (um *respondente*). Entrevistas podem ser feitas pessoalmente ou por telefone. Ver Capítulo 10.

entrevistas por telefone assistidas por computador (ETAC) Adaptação de computadores para estruturar e facilitar entrevistas por telefone, incluindo seleção de amostras, apresentação de *questionários*, registros e análise de dados. Ver Capítulo 10.

equação de regressão Equação gerada para explicar e prever os valores de uma variável independente com base nos valores de uma ou mais variáveis dependentes. Se Y for sempre o dobro de X , a equação de regressão seria $Y = 2X$. A maioria das equações de regressão é um pouco mais complexa. Ver Capítulo 17.

erro não-amostral Imperfeições de qualidade dos dados, resultantes de outros fatores que não erros de amostragem. Exemplos: má compreensão das questões pelos *respondentes*, registros errôneos por entrevistadores e codificadores, erros de digitação etc. Ver Capítulo 16.

erro padrão Desvio padrão de uma *distribuição de amostragem*. Ver Capítulo 5.

escala Tipo de medida composta compreendendo diversos itens tendo uma estrutura lógica ou empírica entre si. Exemplos de escalas: *distância social de Bogardus*, *Guttman*, *Likert* e *Thurstone*. Diferente de índice. Ver Capítulo 8.

escala de distância social de Bogardus Técnica de medição para determinar a disposição das pessoas em participarem de relações sociais — em vários graus de intimidade — com outros tipos de pessoas. É uma técnica particularmente eficiente na medida em que diversas respostas discretas podem ser resumidas sem perda de nenhum detalhe original dos dados. Esta técnica é descrita no Capítulo 8.

escala de Guttman Tipo de medida composta usada para resumir várias observações discretas e representar uma *variável* mais geral. Ver Capítulo 8.

escala de Thurstone Tipo de medida composta construída de acordo com pesos atribuídos por “juizes” a diversos indicadores de algumas *variáveis*. Ver Capítulo 8.

escala Likert Tipo de medida composta desenvolvida por Rensis Likert para melhorar os níveis de medição em pesquisa social pelo uso de categorias padronizadas de respostas em *questionários de survey*. Itens Likert são os que usam categorias de resposta como “concordo fortemente”, “concordo”, “discordo” e “discordo fortemente”. Tais itens podem ser usados na montagem de genuínas escalas Likert ou na construção de outros tipos de medidas compostas. Ver Capítulos 7 e 8.

esgotamento de painel Quando, em *survey* de painel, alguns *respondentes* não participam de todas as etapas do estudo. Ver Capítulo 4.

especificação (modelo de elaboração) Em geral, processo através do qual *conceitos* se tornam mais específicos. Termo técnico usado em conexão com o *modelo de elaboração*, representando o resultado da elaboração no qual uma relação inicialmente observada entre duas *variáveis* é replicada entre alguns subgrupos criados pela *variável de controle* e não entre outros. Nesta situação, terão sido especificadas as condições sob as quais a relação original existe, por exemplo, entre homens mas não entre mulheres. Ver Capítulo 15.

especificações (construção do questionário) Conjunto de instruções detalhadas sobre o tratamento adequado de situações ambíguas ou complexas em *surveys*. Sempre que o entrevistador não sabe como lidar com determinada situação ou resposta, a solução deve ser encontrada nas especificações. Ver Capítulo 10.

estatística Apesar de este termo ser usado como referência geral a dados numéricos, ele tem um significado mais específico em relação à amostragem de *survey*: uma descrição numérica de uma *amostra*, como “45% de homens”. Corresponde a um *parâmetro*, que é uma descrição numérica de toda uma *população*. Em amostragens, estatísticas de amostras são calculadas como estimativas de parâmetros da população. Ver Capítulo 5.

estatística descritiva Computações estatísticas descrevendo as características de uma *amostra* ou a relação entre *variáveis* numa amostra. Estatística descritiva apenas resume um conjunto de observações amostrais, enquanto *estatística inferencial* vai além da descrição de observações específicas para inferir sobre a *população* maior da qual foram retiradas as observações amostrais. Ver Capítulo 16.

estatística inferencial Corpo de computações estatísticas relevantes à feita de inferências baseadas em observações amostrais de uma *população* maior. Ver também *estatística descritiva* e Capítulo 16.

- estímulo** Em experimento, evento cujo impacto está sendo testado. Em pesquisa médica, pode ser uma nova droga. Em criminologia, um programa de aconselhamento. O estímulo experimental corresponde à *variável independente* da análise de *survey*. Ver Capítulo 2.
- estratificação** Agrupamento de unidades que formam uma população em grupos homogêneos (ou estratos) antes da amostragem. Este procedimento, que pode ser usado juntamente com amostragens *aleatória simples, sistemática* ou *por conglomerado*, melhora a representatividade da *amostra*, pelo menos em termos das *variáveis* de estratificação. Ver Capítulo 5.
- estrato** Agrupamento relativamente homogêneo de *unidades de amostragem* da qual *amostras* são selecionadas. Exemplo: amostrando estudantes, podemos amostrar separadamente calouros, 2º ano etc. Cada um dos anos seria um estrato. Todos são relativamente homogêneos no sentido de compartilharem a mesma condição escolar e serem mais ou menos semelhantes em termos das características relacionadas a ano, como idade. Amostragem estratificada exige amostragem dentro dos estratos para obter maior *representatividade* na amostra como um todo. Ver Capítulo 5.
- estudo contextual** *Survey* no qual se coletam dados para descrever as condições nas quais os *respondentes* funcionam. Pode-se fazer um *survey* com estudantes e também colher dados sobre suas escolas. Ver Capítulo 4.
- estudo de coorte** Estudo no qual um grupo específico é estudado ao longo do tempo, embora se possa coletar dados de membros diferentes do grupo em cada conjunto de observações. Exemplo: um estudo da história ocupacional da classe de 1970, com *questionários* enviados a cada cinco anos. Ver Capítulo 4.
- estudo de painel** Tipo de *estudo longitudinal* no qual dados são coletados da mesma *amostra* (o painel) em diversos pontos do tempo. Ver Capítulos 3 e 4.
- estudo de tendências** Tipo de *estudo longitudinal* no qual uma característica de uma *população* é monitorada por algum tempo. Exemplo: as séries de pesquisas da Gallup mostrando as preferências do eleitorado político durante uma campanha, embora amostras diferentes sejam entrevistadas em cada momento. Ver Capítulo 4.
- estudo explicativo** Estudo projetado para revelar *relações causais* entre *variáveis*. Ver Capítulo 4.
- estudo inter-seccional** Estudo baseado em observações que representam um só ponto temporal, ao contrário de um *survey longitudinal*. Ver Capítulo 4.
- estudo piloto** Estudo completo miniaturizado, para testar todos os aspectos do desenho do estudo. Ver Capítulo 12.

estudo sociométrico Exame de redes de relações, tais como amizades, tomada de decisões etc. Ver Capítulo 4.

EIAC Vide abaixo *entrevistas por telefone assistidas por computador*.

experimento Método de pesquisa com manipulação controlada da *variável independente* (o estímulo) para determinar seu impacto na *variável dependente*. Normalmente compreende o uso de um *grupo de controle* e *pré e pós-teste*. Ver Capítulo 1.

explicação Resultado de um modelo de elaboração em que a relação entre duas variáveis é "explicada" pelo controle de uma variável antecedente. Neste caso, dizemos que a relação original era "espúria" e não genuína. Ver Capítulo 15.

F

falácia ecológica Conclusões errôneas sobre indivíduos baseadas somente em observação de grupos. Ver Capítulo 2.

G

gamma Medida de associação apropriada a variáveis *ordinais*. Ver Capítulo 16.

graus de liberdade Uma quantidade requerida para avaliar a significação de uma relação estatística. Num conjunto de valores, é o número de valores que precisamos conhecer para calcular o restante. Sabendo que duas pessoas juntas têm dez reais, ao saber quanto uma tem, sabemos automaticamente quanto a outra tem. Neste caso, dizemos que há um grau de liberdade, porque só um valor pode variar. Se neste exemplo houvesse três pessoas, teríamos dois graus de liberdade. Ver Capítulo 16.

grupo de controle Em experimentação, um grupo de participantes a que nenhum estímulo experimental é dado e que, exceto por isto, se assemelha ao grupo experimental em todos os demais aspectos. A comparação do grupo de controle com o grupo experimental no final do *experimento* mostra o efeito do *estímulo* experimental. Ver Capítulo 2.

grupo experimental Num *experimento*, os sujeitos que não recebem tratamento, servindo de comparação para os sujeitos experimentais. Ver Capítulo 2.

grupos focais Pequenos grupos reunidos para discussões orientadas de um determinado assunto, como um novo produto de consumo, um tema político etc. Podem fornecer um exame profundo de atitudes e preferências, mas em geral não têm uma representatividade rigorosa. Ver Capítulo 5.

H

- hipótese** Expectativa sobre a natureza das coisas derivada de um teoria. Afirmação de que algo deve ser observado no mundo real se a teoria for correta. Ver *dedução* e Capítulo 1.
- hipótese nula** Em conexão com *teste de hipótese* e com *testes de significação estatística*, é a hipótese que sugere não haver relação entre as *variáveis* estudadas. Pode-se concluir que duas variáveis estão relacionadas depois da rejeição estatística da hipótese nula. Ver Capítulo 16.
- homogeneidade** Qualidade de semelhança. Um grupo de pessoas é homogêneo na medida em que elas são semelhantes entre si. Ver Capítulo 5.

I

- índice** Tipo de medida composta que resume várias observações específicas e representa alguma dimensão mais geral. Contrastada com *escala*. Ver Capítulo 8.
- indução** Modelo lógico em que princípios gerais são desenvolvidos a partir de observações específicas. Constatando que judeus e católicos votam mais provavelmente no Partido Democrata do que protestantes, pode-se concluir que minorias religiosas nos EUA se filiam mais ao Partido Democrata e explicar por quê. Isto seria um exemplo de indução. Ver também *dedução* e Capítulo 1.
- intercambiabilidade de índices** Termo cunhado por Paul Lazarsfeld referido à proposição lógica de que se alguma *variável* geral se relaciona com outra variável, todos os indicadores da variável devem ter aquela relação. Ver Capítulo 13.
- interpretação** Termo técnico usado em conexão com o *modelo de elaboração*. Representa o resultado de pesquisa em que se descobre que uma *variável de controle* é o fator mediador através do qual uma *variável independente* produz seu efeito sobre a *variável dependente*. Ver Capítulo 15.
- intersubjetivo** Caracterizado por acordo entre observadores. Quando afirmamos que algo é "objetivamente" verdadeiro, usualmente queremos dizer que observadores diferentes concordam que é verdade. Ver Capítulo 1.
- intervalo de amostragem** Distância padrão entre *elementos* selecionados de uma *população* numa *amostra sistemática*. Ver Capítulo 5.
- intervalo de confiança** Faixa de valores dentro da qual se estima que um *parâmetro* de população vá cair. Um *survey*, por exemplo, pode mostrar que 40% de uma amostra são a favor do candidato A. Embora a melhor estimativa de apoio entre todos os eleitores também seja de 40%, não se deve esperar que ele seja exatamente este. Podemos, pois, computar um intervalo de confiança (por exemplo de 35% a 45%), dentro do qual está provavelmente o

verdadeiro percentual da população. É necessário especificar um *nível de confiança* juntamente com cada intervalo de confiança. Ver Capítulos 5 e 16.

L

- lambda** Medida de associação apropriada a variáveis nominais. Ver Capítulo 16.
- lei** Em teoria, uma generalização universal sobre um conjunto de fatos, por vezes também chamado de princípio. Um exemplo é a lei da gravidade. Ver Capítulo 1.
- linha de regressão** Representação gráfica de uma *equação de regressão*. Linha traçada através de um conjunto de pontos que chega mais perto deles (medida como a distância entre os pontos e a linha), fornecendo uma descrição gráfica da relação entre as *variáveis*. Ver Capítulo 17.
- Livro-código** Documento usado em processamento e análise de dados que indica a localização de *variáveis* diferentes num arquivo de dados e o significado dos códigos usados para representar *atributos* diferentes destas variáveis. Ver Capítulo 11.
- lógica** Sistema de feitura de inferência. As regras através das quais podem-se derivar *hipóteses* ou conclusões de um conjunto de observações ou suposições iniciais. Embora alguns sistemas de lógica sejam rigorosamente explicitados na filosofia, todos nós operamos dentro de sistemas lógicos implícitos (ver, por exemplo, *racional*). Ver Capítulo 1.

M

- matriz de correlação** Formato tabular para apresentação de *correlações* bivariadas dentre diversas *variáveis*. A lista de variáveis é apresentada horizontalmente no topo da tabela e para baixo do lado. Qualquer célula da tabela contém a correlação entre as duas variáveis cuja interseção define aquela célula. Ver Capítulo 16.
- média** Medida computada somando os valores de diversas observações e dividindo pelo número de observações. Ver Capítulo 14.
- média** Termo ambíguo, normalmente sugerindo típico ou normal. Exemplos específicos de médias são a *média aritmética*, a *moda* e a *mediana*. Ver Capítulo 14.
- mediana** Medida representando o valor do caso "no meio" de um conjunto ordenado de observações. Se as idades de cinco homens são 16, 17, 20, 54 e 88 a mediana seria 20. (A *média* seria 39.) Ver Capítulo 14.
- medição** Processo fundamental de pesquisa, no qual *unidades de análise* são identificadas com *atributos* específicos nas *variáveis* estudadas. Por exemplo, o processo de decidir que alguns eleitores são liberais, outros conservadores etc. Ver Capítulo 2.

medida de intervalo Nível de medição que descreve uma *variável* cujos *atributos* são ordenados, com distância igual entre atributos adjacentes. Exemplo: a escala Fahrenheit de temperatura, já que a distância entre 17° e 18° é a mesma que entre 89° e 90°. Ver *medida nominal*, *medida ordinal* e *medida de razão* e Capítulo 7.

medida de razão Nível de medição descrevendo uma *variável* cujos *atributos* possuem todas as qualidades das medidas *nominais*, *ordinais* e *intervalo*, além de se basearem num "genuíno ponto zero". Idade seria um exemplo. Ver Capítulo 7.

medida nominal Nível de medição descrevendo uma *variável* cujos *atributos* são apenas diferentes, em contraste com medidas *ordinais*, *intervalo* ou *de razão*. Sexo é um exemplo de medida nominal. Ver Capítulo 7.

medida ordinal Nível de medida que descreve uma *variável* cujos *atributos* podem ser ordenados ao longo de alguma dimensão. Um exemplo seria status socioeconômico composto dos atributos alto, médio e baixo. Ver *medida nominal*, *medida de intervalo* e *medida de razão*, e o Capítulo 7.

moda Medida representando o valor ou *atributo* mais freqüentemente observado. Se uma *amostra* tem 1.000 protestantes, 275 católicos e 33 judeus, protestante é a categoria modal. Ver Capítulo 14.

modelo de elaboração Método lógico para entender relações causais entre *variáveis*. A relação entre duas variáveis é examinada entre subconjuntos criados por uma *variável de controle*. Ver Capítulos 3 e 15.

modelos log-lineares Forma de análise de dados que usa cálculos logarítmicos para simplificar a análise de complexas tabulações cruzadas multivariadas. Ver Capítulo 17.

moldura de amostragem Lista, ou quase lista, de unidades que compõem uma *população* da qual a *amostra* é selecionada. Para que a amostra seja representativa da população é essencial que a moldura da amostragem inclua todos ou quase todos os membros da população. Ver Capítulo 5.

mutuamente excludente Qualidade desejada num conjunto de respostas que acompanha uma questão fechada — que apenas uma resposta se aplique a cada *respondente*. Ver Capítulo 1.

N

nível de confiança A probabilidade estimada de que um *parâmetro* de população caia dentro de um determinado *intervalo de confiança*. Podemos ter, assim, 95% de confiança de que entre 35% e 45% dos eleitores são a favor do candidato A. Ver Capítulos 5 e 16.

nível de significação No contexto de *testes de significação estatística*, é o grau de probabilidade de que uma relação empírica, observada, possa ser atribuída a erro de amostragem. Uma relação é significativa no nível 0,05 se sua probabilidade de ser apenas função de erro de amostragem não for maior que 5 em 100. Ver Capítulo 16.

O

objetividade Não existe. Ver *intersubjetividade* e o Capítulo 1.

observação participante Técnica de pesquisa na qual pesquisadores participam diretamente dos eventos estudados, seja revelando, seja escondendo sua identidade de pesquisadores. Ver Capítulo 2.

operacionalização Um passo além da conceituação. Operacionalização é o processo de desenvolver *definições operacionais*. Ver Capítulos 1 e 7.

P

paradigma Modelo ou quadro de referência para observação e entendimento. Nas ciências sociais, três paradigmas importantes são funcionalismo, interacionismo e teoria do conflito. Ver Capítulo 1.

parâmetro Característica de uma *população*. Ver Capítulo 5 e *estatística*.

pergunta contingente Questão feita apenas a alguns *respondentes*, determinada pelas respostas deles a outra pergunta. Por exemplo, todos os respondentes podem ser indagados se pertencem à Cosa Nostra e só os que responderem sim seriam indagados com que freqüência vão a reuniões e piqueniques da companhia. Esta última seria uma pergunta contingente. Ver Capítulo 7.

perguntas abertas Itens de *questionário* que pedem aos respondentes dar as próprias respostas em suas próprias palavras. Ver Capítulos 3 e 7, bem como *perguntas fechadas*.

perguntas fechadas Itens de *questionário* onde o *respondente* escolhe entre respostas padronizadas. Exemplo: Qual é o seu sexo? [] Masculino [] Feminino. Ver Capítulo 7.

periodicidade Padrão cíclico numa lista de *unidades de amostragem*, tais como o mesmo número de apartamentos em cada andar, o mesmo número de casas em cada quarteirão etc. Se o período cíclico corresponder ao *intervalo de amostragem* numa *amostra sistemática*, pode-se selecionar uma amostra não-representativa, como selecionar apenas apartamentos perto do túnel de lixo ou somente casas na esquina. Ver Capítulo 5.

ponderação Procedimento usado em conexão com amostragem, através do qual as unidades selecionadas com probabilidades desiguais recebem pesos, de forma a tornar representativa a *amostra* selecionada. Ver Capítulo 5.

população Especificação do *universo* a ser amostrado. Por exemplo, podemos especificar o universo “americanos” como “todos os adultos americanos morando em residências a partir de 1º de julho de 1990”. Ver Capítulo 5.

população do survey Lista ou quase lista específica de *elementos* dos quais se seleciona uma amostra de *survey*. Exemplo: nossa população pode ser “eleitores em Chicago em 1990”, mas nossa população do *survey* (dos últimos se seleciona de fato a *amostra*) pode ser “eleitores em Chicago que constam da lista de eleitores registrados em 15 de julho de 1990”. Em última análise, as amostras que selecionamos são representativas das populações do *survey* das quais as amostras são tiradas. Ver Capítulo 5.

pós-teste Em *experimentos*, medidas da *variável dependente* depois da administração do *estímulo* ao *grupo experimental*. Ver Capítulo 2.

pré-teste (experimentos) Em *experimentos*, medidas da *variável dependente* antes da administração do *estímulo* ao *grupo experimental*. Ver Capítulo 2.

pré-teste (pesquisa de survey) Teste de elementos de um desenho de pesquisa, como itens de *questionário*, técnicas de amostragem etc. Ver Capítulo 12.

probabilidade proporcional ao tamanho (PPT) Tipo de amostra por conglomerados em múltiplas etapas nas quais conglomerados não são selecionados com probabilidades iguais (ver *amostras MIPS*), mas com probabilidades proporcionais a seus tamanhos, medidas pelo número de unidades a serem sub-amostradas. Ver Capítulo 5.

probabilístico Seguir as leis da probabilidade ao formatar um padrão descritivo ou associação entre *variáveis*. Por exemplo, amostragem probabilística gera *amostras* que dão estimativas próximas — mas não perfeitas — das características da *população* da qual foram selecionadas. Podemos também observar que crianças de famílias desfeitas “mais provavelmente” serão delinquentes do que outras, apesar de esta não ser uma relação perfeita. Ver Capítulo 1.

proposição Conclusão tirada logicamente de um conjunto de suposições iniciais (axiomas) num sistema teórico. Ver Capítulo 1.

Q

quantificação Conversão de dados para formato quantitativo ou numérico através de *codificação*. Ver Capítulos 3 e 11.

questionário Documento com perguntas e outros tipos de itens que visam obter informações para análise. Questionários

são principalmente usados em *pesquisas de survey* e também em *experimentos*, pesquisas de campo e outros modos de observação. Ver Capítulos 3 e 9.

questões matriciais Formato de *questionário* no qual uma tabela de itens é mostrada aos *respondentes*, cada um sendo respondido segundo um conjunto padronizado de respostas, tais como “concordo fortemente”, “concordo” etc. Ver Capítulo 7.

qui-quadrado Teste de *significação estatística* apropriado para variáveis *nominais* e *ordinais*. Ver Capítulo 16.

R

racional Razoável; que faz sentido. Racionalidade é um *paradigma* dentro do qual a ciência opera e que é implícito em grande parte da vida cotidiana. É importante compreender que a vida não é racional, mas que racionalidade é uma moldura que impomos às nossas experiências e observações num esforço para lidar eficazmente com a vida. Ver Capítulo 1.

razão amostral Proporção de *elementos* na *população* selecionados para a *amostra*. Ver Capítulo 5.

redução proporcional de erro (RPE) Método lógico para conceituar associações entre *variáveis*. Por exemplo, se tivermos de adivinhar a filiação político-partidária de uma *amostra* de pessoas, cometeremos muitos erros. Mas cometeremos menos erros se soubermos a filiação dos seus pais e apostarmos que é a mesma. A redução de erros, enquanto percentual, é o RPE. Ver Capítulo 16.

regressão curvilínea Forma de *análise de regressão* que estuda relações não-lineares entre variáveis. Ver Capítulo 17.

regressão linear Modelo de regressão que testa se a relação entre duas *variáveis* pode ser representada por uma linha reta. Ver Capítulo 17.

regressão múltipla Modelo de regressão no qual uma equação é desenvolvida para estimar valores de uma *variável dependente* através do conhecimento de valores de diversas *variáveis independentes*. Ver Capítulo 17.

regressão parcial Modelo de *regressão* no qual uma *variável dependente* é estimada com base numa *variável independente*, enquanto uma terceira *variável* se mantém constante. Ver Capítulo 17.

relação causal Relação entre duas *variáveis* na qual uma causa a outra. Por exemplo, pode-se afirmar que aumento na educação causa redução no preconceito. Ver Capítulo 1.

relacionamento [rapport] Relacionamento fácil, natural e confiante entre entrevistador e respondente, tornando mais provável que este responda perguntas mais exata e completamente. Ver Capítulo 3.

replicação Geralmente, duplicação de um experimento para expor ou reduzir erro. Também um termo técnico usado em relação ao *modelo de elaboração*, referindo-se ao *resultado da elaboração* no qual a relação inicialmente observada entre duas *variáveis* persiste quando uma *variável de controle* é mantida constante. Ver Capítulo 15.

representatividade Qualidade pela qual uma *amostra* tem a mesma distribuição de características que a *população* da qual ela foi selecionada. Por implicação, descrições e *explicações* derivadas de uma análise da amostra podem representar o mesmo na população. Representatividade é aprimorada por *amostragem probabilística* e permite generalizabilidade e o uso de estatística inferencial. Ver Capítulo 5.

respondente Pessoa que fornece dados para análise respondendo a um *questionário* de *survey*. Ver Capítulo 10.

seleção aleatória Seleção de *amostras* pelo uso de métodos de probabilidade. Cada *elemento* na *população* amostrada deve ter uma chance conhecida não-zero de seleção. Ver Capítulos 2 e 5.

S

sigilo Descreve um *survey* no qual pesquisadores conhecem ou podem conhecer as respostas de *respondentes* específicos do *survey*, mas garantem que elas permanecerão em segredo. Ver Capítulo 19 e *anônimo*.

silogismo Modelo, às vezes usado para raciocínio na lógica clássica, em que um conjunto de afirmações leva logicamente a uma conclusão. Exemplo: (1) Manuais contêm palavras. (2) Este livro é um manual. (3) Portanto, este livro deve conter palavras. Ver Capítulo 1.

sondar Técnica usada em entrevistas para conseguir resposta mais completa a uma questão. É uma frase ou pergunta não-diretiva, usada para encorajar o respondente a aprofundar uma resposta. Exemplos: "mais alguma coisa?" e "como assim?" Ver Capítulo 10.

survey longitudinal Desenho de estudo compreendendo coleta de dados em diferentes ocasiões no tempo, em contraste com *estudos transversais*. Ver Capítulo 4 e *estudos de tendências, de coorte e painel*.

T

tabela de contingência Formato para mostrar relações entre *variáveis* sob a forma de distribuições percentuais. Ver Capítulo 14.

taxa de respostas Número de participantes num *survey* dividido pelo número selecionado na *amostra*, em forma de porcentagem. Também chamada de taxa de questionários completados ou, em *surveys*

auto-administrados, de taxa de retorno — o percentual de *questionários* enviados e devolvidos. Ver Capítulo 9.

teoria Explicação abrangente de algum setor da existência, incluindo (1) definições de *elementos* constituindo o que vai ser explicado, (2) conjunto de suposições e axiomas que servirão como ponto de partida da teoria e (3) conjunto de afirmações inter-relacionadas sobre as relações entre os elementos. Ver Capítulo 1.

teste de hipótese Determinação de se as expectativas representadas por uma *hipótese* de fato existem no mundo real. Ver Capítulo 1.

teste de significação estatística Classe de computações estatísticas que indicam a probabilidade de que a relação observada entre *variáveis* numa amostra se dever apenas a erro de amostragem. Ver *estatística inferencial* e Capítulo 16.

tipologia Classificação (tipicamente *nominal*) de observações em termos de seus *atributos* em duas ou mais *variáveis*. Um exemplo de tipologia seria a classificação de jornais em urbanos liberais, rurais liberais, urbanos conservadores ou rurais conservadores. Ver Capítulo 8.

U

unidade de observação Unidade da qual dados são coletados. Tipicamente, é a *unidade de análise*, mas as duas podem diferir, como quando entrevistas são feitas em residências (unidade de observação) e dados coletados sobre os membros individuais (unidades de análise). Ver Capítulo 5.

unidades de amostragem Unidades (por exemplo, indivíduos, residências etc.) amostradas. Em geral, a unidade de amostragem é a mesma que a *unidade de observação*, mas não necessariamente. Por exemplo, residências podem ser amostradas mas, em seguida, cada indivíduo nas residências selecionadas é entrevistado (observado). Ver Capítulo 5.

unidades de análise O "que" ou "quem" está sendo estudado. Em pesquisa de ciências sociais as unidades mais típicas de análise são indivíduos. Ver Capítulo 4.

universo Massa geral e abstrata de pessoas sobre as quais se quer tirar conclusões. Exemplo: "americanos". Mas na seleção de *amostras* é necessário ser mais específico na identificação de quem temos em mente. Ver *população, população de survey* e Capítulo 5.

V

validação externa Processo de testar a *validade* de uma medida, como um *índice* ou *escala*, examinando suas relações com outros supostos indicadores da mesma *variável*. Por exemplo, se o índice

realmente mede preconceito, deve correlacionar com outros indicadores de preconceito. Ver Capítulo 8.

validação interna Processo pelo qual os itens individuais de uma medida composta se correlacionam com a própria medida. Assim se testa a decisão de incluir todos os itens na medida composta. Ver ainda *validação externa* e Capítulo 8.

validade Termo usado para descrever uma medida que reflete com exatidão o *conceito* que se pretende medir. Por exemplo, seu QI parece ser uma medida mais válida de inteligência que o número de horas que você passa numa biblioteca. É importante entender que nunca se pode provar a validade definitiva de uma medida. Podemos, porém, concordar quanto à sua validade relativa, baseados nas validades *aparente, por critério, de conteúdo, por construção, interna e externa*. Não deve ser confundido com *confiabilidade*. Ver Capítulo 7.

validade aparente Qualidade de um indicador que o faz parecer uma medida razoável de alguma *variável*. Parece fazer sentido, sem muita explicação, que ir freqüentemente à igreja é indicador de religiosidade. Tem validade aparente. Ver Capítulos 7 e 8.

validade de construção O grau em que uma medição se relaciona da forma esperada com outras *variáveis* num sistema de relações teóricas. Ver Capítulo 7.

validade de conteúdo O grau com que uma medida cobre a faixa de significados incluídos no *conceito*. Ver Capítulo 7.

validade relacionada a critério Em que grau uma medida se relaciona a algum critério externo. Por exemplo, a validade da junta escolar é demonstrada por sua capacidade de prever o sucesso dos alunos na faculdade. Ver Capítulo 7.

variáveis Agrupamentos lógicos de *atributos*. A variável “sexo” se compõe dos atributos “masculino” e “feminino”. Ver Capítulos 5 e 7.

variável binomial Uma *variável* com apenas dois *atributos* é binomial. Um exemplo é sexo, com os atributos masculino e feminino. Ver Capítulo 5.

variável de controle *Variável* mantida constante, para esclarecer a relação entre duas outras variáveis. Exemplo: tendo descoberto uma relação entre educação e preconceito, a variável sexo pode ser mantida constante, examinando a relação entre educação e preconceito somente entre homens e depois somente entre mulheres. Neste exemplo, sexo é a variável de controle. Ver Capítulo 15.

variável de distorção No modelo de elaboração, uma variável de controle que interage com as variáveis independente e dependente, fazendo a relação das duas parecer o contrário do que de fato é. Ver Capítulo 15.

variável dependente A *variável* que se supõe depender de ou ser causada por outra (chamada *variável independente*). Se se descobre que renda é em parte função da quantidade de educação formal, a renda está sendo tratada como variável dependente. Ver Capítulo 15.

variável independente *Variável* cujos valores não são problemáticos numa análise, mas considerados como dados. Supõe-se que uma variável independente cause ou determine o valor de uma *variável dependente*. Se descobrirmos que religiosidade é em parte função do sexo — mulheres são mais religiosas que homens —, sexo é a variável independente e religiosidade a dependente. Qualquer variável pode ser tratada como independente numa parte da análise e como dependente em outra parte. Religiosidade pode se tornar uma variável independente na explicação de crime. Ver Capítulo 15.

variável ordinal Variável cujos atributos diferem entre si em alguma medida de grandeza, como por exemplo variar de alto a baixo ou de fraco a forte etc. Ver Capítulo 7.

variável supressora No modelo de elaboração, uma variável de controle que interage com as variáveis independente e dependente, camuflando a relação entre elas. Ver Capítulo 15.

viés Qualidade de um instrumento de medida que tende a resultar em má interpretação do que está sendo medido numa determinada direção. Por exemplo, o item de *questionário* “Você não concorda que o presidente está fazendo um bom trabalho?” seria tendencioso pois, em geral, incentivaria repostas mais favoráveis. Ver Capítulo 7.

ÍNDICE REMISSIVO

A

- Abramson, Paul R., 261, 275
Ace, Goodman, 115
Ação, 436
Afirmações explicativas, 96
Aleatorização, 205
Alwin, Duane F., 208, 211
American Association for Public Opinion Research (AAPOR), 81, 456-459
American Soldier, The, 455
Americanos negros, 60, 79, 261
Amostra de área complexa, exemplo de, 167
Amostragem aleatória simples, 135
Amostragem. *Ver também* Amostragem probabilística
conceitos e terminologia, 120-125
de mulheres da Igreja Episcopal, 162-165
de professores de escolas
de medicina, 161-162
de domicílios em Oakland, 165-176
desproporcional, 148-151
e Bureau do Censo, 78-79
e o consumidor, 469-470
estratificada, 137-139
intencional, 153-154
não-probabilística, 117, 152-153
por conglomerado em
múltiplas etapas, 140-144, 165-176
por cotas, 115, 154-155
por julgamento, 153-154
precisão de dados da, 114-117
Amostragem com probabilidade proporcional ao
tamanho (PPT), 145-147
exemplo de, 162-165
Amostragem sistemática, 135-137, 139-140
Estratificação implícita, 139-140
Amostras aleatórias simples, 117
Amostras estratificadas, 117
Amostras estratificadas sistemáticas, 117
Amostras MIPSE, 120
Amostras paralelas, 106
Amostras por conglomerado, 117, 161-166
Amostras sistemáticas, 117
Amplitude, 243
Amplitude semi-interquartilica, 343
Análise.
Ver também Análise multivariada; Análise
de regressão; Análise univariada
avaliativa, 320
bivariada, 349, 353-356
de conteúdo, 70-72, 87, 156
de dados, 72-73, 470-471
de estudo-piloto, 312-313
de item, 229-230, 233
de pesquisa de *survey*, 223-224
de pré-teste, 309-310
de trajetória, 415-417, 431
de variância, 421-424, 431
discriminante, 424-428, 432
ética na, 455-456
fatorial, 417-421, 431
relatório de, 440, 441-442
secundária, 79
unidades de, 98-101
Análise bivariada, 341
construindo tabelas na, 353-356
formatos de tabela na, 356
Análise multivariada, 97, 337, 357-360
técnicas avançadas para, 409-432
Análise de regressão, 409-410, 431
cuidados, 415
curvilínea, 414-415, 431

linear, 410-421, 431
múltipla, 413, 431
parcial, 414, 431

Análise univariada, 337

dispersão em, 343
distribuição em, 337-339
tendência central em, 339-343
variáveis em, 343-344

Anderson, Barbara A., 261, 275

Anonimato, 451-452

Anti-semitismo, 416-417

Associação

e medições, 331, 333-343
medidas de, 385-391, 416

Associação Americana de Ciência Política, 81

Associação Americana de Sociologia, 81

Associações, profissionais, 81

Atributos, 182, 281

B

Bishop, George F., 192, 211

Brahe, Tycho, 48

Bureau do Censo. *Ver* U.S. Bureau of the Census

Bush, George, 116-117

Butterfield, Herbert, 48, 55

C

Califórnia, Universidade da

em Berkeley, 81, 97, 272

em Los Angeles, 81

Campo, 281

Cantril, Hadley, 181

Cargas fatoriais, 418

Carmines, Edward, 197, 211

Cartões IBM, 281

Censo, 78

Chicago, Universidade de, 81

Christian Beliefs and Anti-Semitism (Glock e Stark), 233

Ciência, 37, 43-45

características de, 45-54

e ciência social, 57-75

e pesquisa de *survey*, 82-86, 87-90

imagem tradicional da, 327-331

na prática, 41-43

visão crítica, 39-41

visão tradicional da, 37-39

Ciência social, 57-58

características de, 62-66

como científica, 87-90

e regularidades sociais, 58-62

métodos de pesquisa em, 67-75

programas de computador para, 284

Cientistas, 43-45, 180

Classe social, 181

e envolvimento com a igreja, 83, 375-376

Codificação, 87, 288-289, 297-299. *Ver também*

Codificação nas margens; Pré-codificação

para folhas de leitura óptica, 299

para folhas de transferência, 293

Codificação nas margens, 294

Coefficientes de trajetória, 416

Coleta de dados, 245, 328

estudos-piloto, 311-312

pré-teste, 307

Companhia de Máquinas de Tabulação, 281

Computadores, 280

em pesquisa de *survey*, 280-288

Conceitos, 42

Conceituação, 179-182

Confiabilidade, 195-196

e validade, 198

Conglomerados [*clusters*], 140

Consumidor, de pesquisa de *survey*, 465-471

Copérnico, Nicolau, 47

Correlação, 43

Correlação produto-momento (*r*) de Pearson, 390-391,

406

Crask, Melvin R., 252-253, 258

D

Dados

análise de, 72-73, 470

e cientistas, 180

em cartões perfurados, 281-282

faltantes, em índices, 227-229

precisão de, 114-117

tipos de, 185-186
redução de, 214, 235-236, 243, 384-385

Darwin, Charles, 48, 61

Dedução, 41

Delinqüência, 419-420

Depressão, a, 72, 79

Desconfiabilidade, 52

Descrição, 96, 179, 436

Descrições de subgrupos, 344-349

Desenhos de amostra
exemplos de, 159-177
e o Bureau do Censo, 79
pré-teste de, 303-304
tipos de, 134-147

Desenhos de pesquisa de *survey*, 93-94 *Ver também*
Desenho de pesquisa; Desenho de amostra; Desenhos de *survey*

Desenho de pesquisa, 466-467

Desenhos de *survey*,
básicos, 101-105
escolha de, 108-109
relatório de, 440
variações do básico, 105-107

Desvio padrão, 343

Desvio quartílico, 343

Determinismo
probabilístico, 48-49
em ciência, 48
em ciência social, 63
em pesquisa de *survey*, 83

Dewey, Thomas, 115

Dimensões, 182
medição de, 184

Discagem digital aleatória, 133

Dispersão, 343

Distribuição amostral binomial, 126-131

Distribuições de frequência, 338

Drogas, 104-105

Dukakis, Michael, 116

Durkheim, Émile, 72-73, 75

E

Einstein, Albert, 44, 50

Elaboração, 381

Elemento, 121

Ellis, Shmuel, 155, 158

"Emparelhamento racial", 354-356

Entrada de dados, 293
direta, 294-295
pelos entrevistadores, 296
pré-codificado para, 297-299

Entrevista por telefone assistida
por computador (ETAC), 271, 272

Entrevistador(es)
apresentação e aparência, 261-263
despedindo, 274-275
entrada de dados por, 296
importância do, 259-260
instruções para, 208-209
treinamento do, 266-267

Entrevistando
operação de, 270-275
regras para, 261-265
survey, 87

Entrevistas
prática, 268
redistribuindo, 273
verificando, 274

Entrevistas de *survey*
entrevistador para, 259-260, 266-269
entrevistando para, 261-265, 270-275

"Entrevista por interceptação em *shoppings*", 155

Envolvimento
e classe social, 375-376
tipos de, em religião, 183

Equação de regressão, 410

Erro(s),
amostragem, 124, 125, 137, 486
não-amostrais, 393
padrão, 128
redução proporcional do, 385, 406, 412

Escalas, 243
construção de, 233-243
de Distância Social de Bogardus, 234-235
de diferenciação semântica, 389
Guttman, 236-240
Likert, 189, 232-233
Thurstone, 235-236
versus índices, 214-217

Escola de Columbia, 363

Escolas de medicina
amostrando professores de, 161-162
atitude sobre o Medicare em, 363

Escores, de índices, 225-227, 238-239

Especificação, 374-376, 381

Especificidade
Da ciência, 51
Da ciência social, 65
Da pesquisa de *survey*, 84-86

Estatística(s), 124
social, 385, 405-407

Estatística descritiva, 383, 405
medidas de associação em, 385-391, 405
redução de dados em, 383-385

Estatística inferencial, 383, 391, 405
inferências univariadas em, 392-393
testes de significância estatística em,
393-405, 405-407

Estímulo, 67

Estratificação, 138-139
e amostragem por conglomerado
em múltiplas etapas, 144
em amostragem sistemática, 139-140
em exemplo de amostra complexa
de área, 168-171
em exemplo de amostra sistemática
estratificada, 160

Etrato, 144
seleção de quarteirões dentro de, 171-172

Estudos contextuais, 106-107

Estudos de caso, 73-74

Estudos de coorte, 102-103

Estudos de painel, 80, 102, 103-104, 108

Estudos-piloto, 303, 309
avaliando, 313-320
fazendo, 310-313

Estudos sobre as Eleições Nacionais, 261

Estudos sociométricos, 107

Estudos de tendências, 102

ETAC, 271-272

Ética,
código de, 456-459
em pesquisa de *survey*, 445-456, 462
ilustrações, 460-462

Ética protestante, 78

Experimento, 39
controlado, 67-69, 86-87
de laboratório, 156

Explicação, 96-97, 179
como objetivo de pesquisa, 437
em paradigma de elaboração, 372-373, 381

Exploração, 97, 436

Extrapolação, 415

F

Falácia ecológica, 72, 99

Filiação política, 421-424

Firmas de pesquisa de opinião, 79

Fisher, R. A., 80, 118

Folhas de código, 296

Folhas óptico-sensíveis, 296

Folhas de transferência, 293

Forslund, Morris A., 418-419

Fox, Richard J., 252-253

G

Gamma, 387-391, 406

Gallup, George, 79, 115

Generalizabilidade,
na ciência, 49-50
na ciência social, 62-63
na pesquisa de *survey*, 83

Glock, Charles Y., 82, 90, 183, 211, 333, 334,
380, 382

Graus de liberdade, 402

Grupo de controle, 68

Grupo experimental, 68

Grupos focais, 155

Guttman, Louis, 236

H

Harris, Louis, 79

Heterogeneidade, 118-119

Hipóteses, 39, 329
nula, 401

Hipóteses *ex post facto*, 379-380

Hollerith, Herman, 280

Homogeneidade, 118-119

I

Igreja, 82-83, 375-376. *Ver também* Religião; Religiosidade

Igualdade sexual, 349-360
exemplo hipotético de, 394-399

Índices,

- dados faltantes em, 227-229
- e escalonamento Likert, 232-233
- escores de, 229-232
- interpermutabilidade de, 327, 331-332
- relações bivariadas em, 218-222
- relações multivariadas em, 222-225
- seleção de itens para, 217-218
- validação de, 229-232
- versus escalas, 214-217

Indução, 41

Inferências univariadas, na estatística inferencial, 391-393, 406

Instituto Americano de Opinião Pública, 115

Instituto para pesquisa social, 81

Instruções, para questionários, 206-209

Instrumento de pesquisa

- Estudo-piloto, 311
- Pré-testando, 304-306

International Business Machines Corporation (IBM), 281

Interpermutabilidade, de índices, 327, 331-332

Interpolação, 415

Interpretação, 374, 381

Intersubjetividade, 86, 290

- da ciência, 52-53
- da ciência social, 65-66

Intervalo amostral, 136

Intervalos de confiança, 124, 130

Item(ns)

- análise de, 229-230
- relações bivariadas entre, 218-222
- relações multivariadas entre, 222-225
- seleção de, 217-218
- validação de, 319
- viés em, 194

J

Jacquard, Joseph-Marie, 280

K

Kendall, Patricia, 367-368, 382

Kepler, Johannes, 48

Kim, Jonghoon, 252-253, 258

Kish, Leslie, 121, 158

Krosnick, Jan A., 208, 211

L

Lambda, 386-387, 406

Landon, Alfred M., 115

Lazarsfeld, Paul F., 80, 90, 327, 331, 334, 363, 367-368, 382

Leis, 44

Likert, Rensis, 189, 232

Limpeza

- de contingências, 300-301
- de dados, 299-300
- de possíveis-códigos, 299

Linha de regressão, 410, 411-412, 431

Literary Digest, 115-116

Livro de código, 291-292

- construção de, 291-293

Lógica

- da amostragem probabilística, 118-120
- da ciência, 45-48
- da ciência social, 62-63
- da conceituação, 179-182
- da construção de índices e escalas, 216-217
- da elaboração, 381
- da pesquisa de *survey*, 82-83
- da significância estatística, 399-401
- dedutiva, 46-48
- indutiva, 47-48

Lopata, Helena Znaniecki, 389, 407

M

Marginais, 338

Marx, Karl, 78, 180

Matriz de correlações, 389

Mazur, Allan, 87, 90

Mediana, 339-343

Médias, 339-341
Medicare, 376
Medição(ões) 58
 de dimensões, 184
 e associação, 331, 333-334
 e consumidor, 467-468
 níveis de, 187-189
Medições nominais, 187
Medições de intervalo, 187
Medições ordinais, 187
Medições de razão, 188
Medidas de associação, 385-386
 variáveis intervalo ou de razão em, 390-391
 variáveis nominais em, 386-387
 variáveis ordinais em, 387-390
Medidas de tendência central, 339-343
Megabytes, 287
Memória, 287
Método interpretativo, 363
Método de Lazarsfeld, 363
Métodos de amostragem
 tipos de, 117-118
 usos não-*survey* de, 156-157
Métodos de *survey* assistidos por computador, 272
Microcomputadores, 286-289
Mitchell, Arnold, 242, 244
Moda, 339, 341
Modelo de elaboração, 80
 e paradigma de elaboração, 370-379
 e hipóteses ex post facto, 379-380
 história de, 363-370
Modelos log-lineares, 428-430, 432
Modems, 286
Modificabilidade
 da ciência, 53-54
 da ciência social, 66-67
 da pesquisa de *survey*, 86
Modificação, de amostra, 160-161
Moldura de amostragem, 115, 117, 123, 131-133
 exemplo de amostra complexa de área, 167
 exemplo de amostra estratificada sistemática, 159-160

Mulheres episcopais, 83, 84, 375, 376, 451
 amostragem de, 162
Multidimensionalidade, de tipologias, 240
Mutuamente excludentes, 46

N

National Opinion Research Center, 81
Newton, Sir Isaac, 44
Níveis de confiança, 124-125, 130
 e inferências univariadas, 392, 406
Níveis de significância, 400
Números aleatórios, tabela de, 482-485

O

Oakland, California, amostrando domicílios em, 165-176
Observação participante, 74-75, 156-157
Oldendick, Robert W., 192, 211
Operacionalização, 38, 128
 de conceitos, 328
 quadro de referência para, 182-185
Ortodoxia, 333

P

Pagels, Heinz, 89, 90
Paradigma de elaboração, 370-371
 especificação em, 374-376, 381
 explicação em, 372-373, 381
 interpretação em, 374, 381
 modelo de elaboração e, 370-379
 refinamentos em, 377-379
 replicação em, 371-372, 381
Paradigmas, 55
Parâmetro, 124
Parcimônia
 da ciência, 50
 da ciência social, 64
 da pesquisa de *survey*, 84
Participação voluntária, 448-450
Pearson, Karl, 80
Percentuando, 349
 tabelas, 354
Periodicidade, 136, 140

Pesos beta, 416
Pesquisa. *Ver também* Pesquisa de survey
 métodos de pesquisa científica, 67-75
Pesquisa de survey, 77
 análise de, 323-325
 características científicas do, 82-86
 como científica, 87-90
 computadores na, 280-288
 consumidor de, 465-471
 e outros métodos, 86-87
 história do, 78-81
 no contexto social, 447-463
 objetivos do, 95-98
 relatório do, 433-444
Pesquisa Gallup, 90-101
Pesquisas políticas, 114-117
Plágio, 438-440
Política, *Ver também* Voto 78
Ponderação
 e amostragem desproporcional, 148-151
 e inferência estatística, 152
 em exemplo de amostra complexa
 de área, 174-176
 graus de precisão, 150-151
 métodos para, 151
População(ões), 117, 121, 131-134
 de survey, 122, 160
Porcentagens, 338
Pós-teste, 68
Pravda, 71
Precisão, 194
Pré-codificando, 297-299
Preconceito, 59
 e ortodoxia, 333
 experimento em, 67-69
Predição, 351
Pré-testes, 68
 avaliando, 313-320
 fazendo, 303-310
Probabilidade, de seleção, 119-120
Procedimentos objetivos, 39
Procedimentos racionais, 39
Processamento de dados, 279-301
 estudo-piloto, 311-312
 pré-testando, 307-309

Proposições, 38
Ptolomeu, Claudius, 47-48
Public Opinion Quarterly, 81

Q

Quantificação, 86
Quarteirão de recenseamento, 166
Questionário, 77
 estudando o, 267-268
 familiaridade do entrevistador com, 263
 formato para, 198-205
 instruções para, 206-209
 ordenando as questões para, 205-206
 pré-codificação do, 297-299
 reprodução do, 209-210
 validação de itens, 318
Questionários auto-administrados
 correspondências de acompanhamento,
 251-253
 estudo de caso de, 254-257
 distribuição postal e retorno de, 248-249
 monitorando retorno de, 250-251
 pré-codificação de, 297-299
 taxas de resposta para, 253-254
Questões
 abertas, 87, 189-190
 avaliando a clareza de, 313-315
 contingentes, 200-203
 fechadas, 189-190
 guias para a construção de, 189-194
 matriz, 204-205
 no formato geral do questionário, 198-205
 ordem de, 205-206
 seguindo a redação exata de, 263-264
Qui-quadrado, 401-403, 488-489

R

Raça, 378
Radicalismo estudantil, 97
Rapport, 87
Razão amostral, 136
Redes de computadores, 286
Redução proporcional do erro (RPE), 385, 406
 em regressão linear, 412
Reese, Stephen D., 261, 275

Regressão curvilínea, 414-415, 431
Regressão linear, 410-412, 431
Regressão múltipla, 413, 431
Regressão parcial, 414, 431
Regularidades sociais, 58-62
Relações bivariadas, entre itens de índices, 218-222
Relações causais, 38
Relações espúrias, 372
Relações parciais, 381
Relatando, avaliando, 319-320

Relatório

análise de dados no, 441-442
considerações básicas para, 433-436
ética no, 455-456
organização do, 436-441
Relatório de dados, *Ver também* Relatório; *Relatando*

Religião. *Ver também* Igreja
e anti-semitismo, 416-417
e suicídio, 72

Religiosidade, 59, 108. *Ver também* Igreja e
envolvimento, 182-184

Replicação, 371-372, 381

Representatividade, 119-120

Respondentes, 450-451

Respostas

formatos para, 199-200
fundindo categorias de, 345-347
"não sei", 347-349
registrando, 264
sondando, 265
variância em, 315-319

Reynolds, H. I., 429, 432

Ringer, Benjamin, 82, 90

Roosevelt, Franklin D., 115

Roper, Elmo, 79

Rosenberg, Morris, 377, 382

S

Sakumoto, Tom, 192

Scanner óptico, 296

Seleção

aleatória, 68, 125-126
de itens, 217-218

probabilidade de, 119-120

Selvin, Hanan, 405

Shanks, J. Merrill, 272

Sigilo, 451, 452-453

Significância. *Ver também* Testes de significância
estatística

nível de, 400

substantiva, 403, 404, 406

Silogismo, 47

Silver, Brian D., 261, 275

Smith Tom, 194, 211

Sociologia, 87-88

Stark, Rodney, 333, 334

Status social, 180

Stouffer, Samuel A., 60, 72, 75, 79, 80, 90, 380
e o modelo de elaboração, 363-370, 371

Strunk, William, Jr., 433, 442

Suicídio, 72-73

Suits, Daniel, 88

Survey Research Center, 81-271

Survey Research Laboratory, 81

Survey Research Office, 252

Surveys interseccionais, 101, 108

aproximação de *surveys* longitudinais,
104-105

Surveys longitudinais, 101-104

Surveys pelo correio, 247-248

correspondência de acompanhamento,
251-253

distribuição e retorno de, 248-249

estudo de caso de, 254-257

monitorando retornos de, 250-251

taxas de resposta, 253-254

Surveys por entrevista, 259-277

T

Tabela(s)

de mudanças, 104

de números aleatórios, 482-485

explicativas, 360

na análise bivariada, 353-357

na análise multivariada, 357-360

Taxas de resposta, 253-254

Coleção Aprender

1. FOLCLORE EM MINAS GERAIS (2ª edição revista e ampliada)
Saul Martins
2. ASTRONOMIA FUNDAMENTAL
Rodrigo Dias Târsia
3. EDUCAÇÃO ARTÍSTICA - Introdução à História da Arte
(2ª edição revista e ampliada)
Sandra Loureiro de Freitas Reis
4. ESCREVER SEM DOER - Oficina de Redação (3ª reimpressão)
Ronald Claver
5. FONTES DE INFORMAÇÃO ESPECIALIZADA - Características e Utilização (esgotado)
Bernadete Santos Campello e Carlita Maria Campos
6. 1, 2, 3 DA SEMIÓTICA
Júlio Pinto
7. UM TOQUE DE CLÁSSICOS - Durkheim, Marx e Weber
(2ª reimpressão)
Tania Quintaneiro, Maria Lígia de O. Barbosa e Márcia Gardênia de Oliveira
8. LATINA ESSENTIA - *Preparação ao Latim* (3ª edição revista e ampliada)
Antônio Martinez de Rezende
9. DESENVOLVIMENTO HUMANO E PSICOLOGIA - Generalidades, Conceitos, Teorias
Vânia Brina Corrêa Lima de Carvalho
10. O BELO AUTÔNOMO - Textos Clássicos de Estética
Rodrigo Duarte (Org.)
11. OS HETERÓPTEROS AQUÁTICOS DE MINAS GERAIS - Guia Introdutório com Chave de Identificação para as Espécies de Nepomorpha e Gerromorpha
Nico Nieser e Alan Lane de Melo
12. PRINCÍPIOS DE MORFOMETRIA DIGITAL - KS300 para Iniciantes
Marcelo Vidigal Caliarí

13. Sistema de Esgotos
Patricio Gallegos Crespo
14. Técnicas Alternativas de Conservação -
Recuperação de Livros, Revistas, Folhetos e Mapas
(2ª edição revista)
Sônia de Conti Gomes e Rosemary Tofani Motta
15. Manual para Normalização de Publicações
Técnico-Científicas (5ª edição revista e aumentada)
*Júnia Lessa França, Ana Cristina de Vasconcellos,
Maria Helena de Andrade Magalhães e Stella Maris Borges*
16. Iniciação Esportiva Universal - Da Aprendizagem
Motora ao Treinamento Técnico (vol. 1)
Pablo Juan Greco e Rodolfo Novellino Benda (Org.)
17. Iniciação Esportiva Universal - Metodologia
da Iniciação Esportiva na Escola e no Clube (vol. 2)
Pablo Juan Greco (Org.)
18. Estruturas Morfológicas do Português
(1ª reimpressão)
Luiz Carlos de Assis Rocha
19. Métodos de Pesquisas de Survey (1ª reimpressão)
Earl Babbie
20. Asma - A Resposta Inflamatória Pulmonar e seu Controle
Farmacológico
J. N. Francischi e D. M. Conroy
21. aprender ciências - Um Mundo de Materiais
(livro do aluno)
*Maria Emília Caixeta de Castro Lima, Orlando Gomes de Aguiar
Júnior e Selma Ambrosina de Moura Braga*
22. aprender ciências - um mundo de materiais
(livro do professor)
*Maria Emília Caixeta de Castro Lima, Orlando Gomes de Aguiar
Júnior e Selma Ambrosina de Moura Braga*
23. Fontes de Informação para Pesquisadores
e Profissionais
*Bernadete Santos Campello, Beatriz Valadares Cendón
e Jeannette Marguerite Kremer (Org.)*

24. LINGUAGEM E FORMAÇÃO DE CONCEITOS NO ENSINO
DE CIÊNCIAS
Eduardo Fleury Mortimer
25. Lazer, Trabalho e educação - relações históricas,
questões contemporâneas
Christianne Werneck
19. introdução à lógica simbólica
Paulo Roberto Margutti Pinto
20. elevatórias de sistemas de esgotos
Patricio Gallegos Crespo

A presente edição foi composta pela Editora UFMG, em caracteres Gatineau, corpo 11/13, e impressa pela Rona Editora, em sistema offset, papel offset 90g (miolo) e supremo 250g (capa) em abril de 2003.

Sobre o autor

Earl Babbie graduou-se na Universidade de Harvard, obtendo o doutorado na Universidade da Califórnia em Berkeley. Atualmente leciona Sociologia e é chefe do Departamento de Sociologia no Chapman College. É membro da American Sociological Association e da American Association for Public Opinion Research. Além deste livro, é autor de três outros textos sobre métodos de pesquisa: *Social Research of Consumers*, *Observing Ourselves: Essays in Social Research* e *The Practice of Social Research*.

ISBN 857041175-8



9 788570 411754 >

O livro de Earl Babbie, *Métodos de Pesquisas de Survey*, tornou-se um clássico nos cursos de graduação em ciências sociais nas universidades americanas, logo que foi publicado. Ao mesmo tempo que levanta detalhadamente e ilustra os diferentes instrumentos da pesquisa de *survey* e das ocasiões de sua aplicação apropriada, o livro também expõe as razões de ser dos procedimentos usados. É um manual útil para estudantes de metodologia das ciências sociais e marketing, para pesquisadores em geral, além de usuários de pesquisas de *survey*, tais como clientes de levantamento de mercado ou de relatórios sobre indicadores sociais, administradores públicos ou privados que usam dados de *surveys*, leitores de jornais etc.

Biblioteca Central - UnB



B0071533

2



aprender