

# 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:<sup>1</sup>

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.<sup>2</sup>

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.

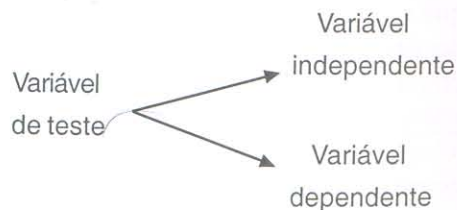
## O Paradigma de Elaboração

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.<sup>3</sup> 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.<sup>4</sup> 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	Leste	84
	Sul	49

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 delinquentes 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 delinquência 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.<sup>5</sup> 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.<sup>6</sup> 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.<sup>7</sup> 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.<sup>8</sup> 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

<sup>1</sup> STOUFFER, Samuel A. et al. *The American Soldier*. Princeton, N.J.: Princeton University Press, 1949. v.1. p.122 et seq., esp. p.127.

<sup>2</sup> 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.

<sup>3</sup> 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.

<sup>4</sup> 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.

<sup>5</sup> GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. p.92.

<sup>6</sup> ROSENBERG, Morris. *The Logic of Survey Analysis*. New York: Basic Books, 1968.

<sup>7</sup> Ibidem. p.88-89.

<sup>8</sup> 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.