

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

<sup>1</sup> FOX, Richard J., CRASK, Melvin R., KIM, Jonghoon. Mail *Survey* Response Rates. *Public Opinion Quarterly*, v.52, p.467-491, inverno 1988.

<sup>2</sup> 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.

<sup>3</sup> 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.<sup>1</sup> 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 Abramson<sup>2</sup> 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 colegas<sup>3</sup> 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 engenho 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:

Ói, escuta só essas perguntas e responda. 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.

## Segundo 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

## 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 das 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.

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.<sup>4</sup>

## 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.

## 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.<sup>5</sup> Todos os *surveys* e questionários diferem uns dos outros, e os entrevistadores devem ser treinados 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 sentir-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. Frequentemente, 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 subsequentes tiverem que se superpor ao início das entrevistas verdadeiras. É melhor alguns entrevistadores começarem 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 discar 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 trapaceiam. 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.<sup>6</sup>

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

<sup>1</sup> 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 decrescem 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.

<sup>2</sup> 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.

<sup>3</sup> 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.

<sup>4</sup> Anos atrás, me pediram para falar de um *survey* sobre qualidade de vida que eu desenhei para os idosos do Havaí. Perguntado como lidávamos 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 afundava 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?'".

<sup>5</sup> O impacto do treinamento do entrevistador na qualidade dos dados coletados foi demonstrado por BILLET, 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.

<sup>6</sup> 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 traçado 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 traçado. 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 traçado) podia ser *codificada* e *armazenada* sob a forma de buracos num cartão e subsequentemente *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 *noze* 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

```

00000000000000000000000000000000000000000000000000000000000000000000000000
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
111111111111111111111111111111111111111111111111111111111111111111111111111111
222222222222222222222222222222222222222222222222222222222222222222222222222222
3333333333333333333333333333333333333333333333333333333333333333333333333333
444444444444444444444444444444444444444444444444444444444444444444444444444
55555555555555555555555555555555555555555555555555555555555555555555555555
66666666666666666666666666666666666666666666666666666666666666666666666666
77777777777777777777777777777777777777777777777777777777777777777777777777
88888888888888888888888888888888888888888888888888888888888888888888888888
99999999999999999999999999999999999999999999999999999999999999999999999999
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6
7890123456789012345678901234567890123456789012345678901234567890123456789012
34567890123456789012345678901234567890123456789012345678901234567890123456789012
    
```

### 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 impróprio 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 operam 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 ensanduiçar 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

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.

A 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:

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 disar 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

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.<sup>1</sup>

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 referiram 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

<p><b>VAR. 1: Fun Atu</b> Qual é sua função atual na Universidade Chapman?</p> <ol style="list-style-type: none"> <li>1. Administração</li> <li>2. Professor permanente</li> <li>3. Professor contratado, em vias de tornar-se permanente</li> <li>4. Professor contratado temporariamente</li> <li>5. Diretoria</li> <li>6. Função mista</li> <li>7. Outra</li> </ol>	<p><b>VAR. 4: Grau</b> Qual o mais alto grau acadêmico que você possui?</p> <ol style="list-style-type: none"> <li>1. Bacharel</li> <li>2. Mestre</li> <li>3. Doutor</li> <li>4. Outro</li> </ol>
<p><b>VAR. 2: Pos Acad</b> Qual a sua posição acadêmica?</p> <ol style="list-style-type: none"> <li>1. Professor Assistente</li> <li>2. Professor Associado</li> <li>3. Professor</li> <li>4. Professor Adjunto</li> <li>5. Outro</li> <li>6. Não aplicável</li> </ol>	<p><b>VAR. 5: Sexo</b> Qual o seu sexo?</p> <ol style="list-style-type: none"> <li>1. Feminino</li> <li>2. Masculino</li> </ol>
<p><b>VAR. 3: Camp Acad</b> Qual é o seu campo acadêmico?</p> <ol style="list-style-type: none"> <li>1. Negócios</li> <li>2. Educação</li> <li>3. Humanidades</li> <li>4. Ciências do Movimento e Exercícios</li> <li>5. Ciências Naturais</li> <li>6. Ciências Sociais</li> <li>7. Mistó ou outros</li> <li>8. Não aplicável</li> </ol>	<p><b>VAR. 6: Não/Rumo</b> A universidade não tem rumo.</p> <ol style="list-style-type: none"> <li>1. Concordo fortemente</li> <li>2. Concordo</li> <li>3. Discordo</li> <li>4. Discordo fortemente</li> <li>5. Não sei</li> </ol>
	<p><b>VAR. 7: Para/Pesq</b> A universidade está se afastando do ensino e se movendo na direção da pesquisa.</p> <ol style="list-style-type: none"> <li>1. Concordo fortemente</li> <li>2. Concordo</li> <li>3. Discordo</li> <li>4. Discordo fortemente</li> <li>5. Não sei</li> </ol>

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 correspondentes 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 *revisados* 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

<sup>1</sup> 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



## 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.

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.

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.

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

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, estaremos 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.

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

Podem 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

excluídas, 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 frequentemente, 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".* Frequentemente, 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 sobrepor-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 propositalmente 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.* Frequentemente, 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.* Frequentemente, 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 produzem 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 frequentemente 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 sinais. 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 factual 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 considerá-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.

## Avaliação da Análise e do Relatório

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 menção 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 moldar 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.)

## Resumo

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 erros 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*

## 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.