

Em segundo lugar, a interpretação conduz ao estabelecimento de *conceitos explicativos*. Se se verificasse que os estudantes estrangeiros que tiveram maior interação com os norte-americanos eram mais favoráveis em suas atitudes, ou nestas mostravam maior mudança, os pesquisadores esperariam ser capazes de deduzir algo a respeito do processo pelo qual a ligação pessoal influi nas atitudes. Se as diferenças de atitude fôsem maiores entre estudantes diferentes quanto à *variedade* de suas experiências com os norte-americanos, os pesquisadores poderiam chegar à conclusão provisória de que a ligação pessoal muda atitudes pelo fato de destruir este-reótipos. Se as diferenças de atitude fôsem maiores entre estudantes que diferiam quanto à *intimidade* das atividades em que participavam com norte-americanos, poder-se-ia fazer a inferência de que a ligação pessoal influi nas atitudes pelo fato de dar ao indivíduo uma oportunidade para observar qualidades pessoais que não são aparentes em contatos mais superficiais. Se a diferença fôsse maior entre estudantes que diziam ter tido um *amigo íntimo norte-americano* e os que não o tinham tido, poder-se-ia deduzir que, através do aumento da afetividade do sentimento pessoal, a ligação com outra pessoa conduz à mudança de atitude. Esta função da interpretação está tão estreitamente ligada à função da teoria da pesquisa social, que será tratada no Capítulo 14, dedicado à relação entre pesquisa e teoria.

Nossa discussão dos processos usados em análise e interpretação abrangerão os seguintes tópicos: *estabelecimento de categorias*, a aplicação, através da *codificação*, de categorias aos dados brutos, a *tabulação de respostas*, a *análise estatística dos dados*, a *inferência de relações causais*, o *uso de dados não-quantificados*.

### O Estabelecimento de Categorias

#### OS PRINCÍPIOS DE CLASSIFICAÇÃO COMO BASE PARA CONJUNTOS DE CATEGORIAS

Num levantamento de bairro, uma amostra da população devia responder à seguinte pergunta: "Que tipos de pessoas vivem neste bairro?" O objetivo da pergunta não era obter informação objetiva a respeito do bairro (o que poderia ter

sido obtido, de maneira mais exata e com menos esforço, através de números de recenseamento), mas verificar até que ponto as pessoas do bairro pensavam em função de distinções de grupos étnicos. Eis algumas das respostas:

Quase que só pessoas pobres como eu.

Pessoas de cor e nós.

Negros, italianos, judeus e muitos outros.

Existe muito o que dizer sobre as pessoas que moram por aqui. Mas não quero complicações. A gente precisa viver onde pode.

Pessoas de cor.

É um bairro "duro." Aqui vivem todos os tipos de pessoas.

Sei que existem alguns irlandeses.

Muitos judeus.

Um grande número de negros..

Realmente não sei; sou novo aqui.

Evidentemente, para que várias centenas de respostas como essas possam ser organizadas de modo a serem usadas para responder às perguntas da pesquisa, precisam ser agrupadas em certo número de categorias. A fim de decidir quais são as categorias significativas, é preciso escolher algum princípio de classificação. A pergunta da pesquisa ou as hipóteses — se formuladas — apresentam a base para a escolha dos princípios de classificação. Suponha-se que esse estudo estivesse voltado para a consciência da existência do grupo negro; o princípio adequado de classificação seria um princípio baseado em referência explícita aos negros. Esse princípio de classificação imediatamente sugere duas categorias:

Menciona negros

Não menciona negros

Essas duas categorias formam um "conjunto de categorias." Um conjunto de categorias deve atender a algumas regras básicas:

1. O conjunto de categorias deve ser derivado de um único princípio de classificação.
2. O conjunto de categorias deve ser exaustivo; vale dizer, deve ser possível colocar qualquer resposta numa das categorias do conjunto.

3. As categorias do conjunto devem ser mutuamente exclusivas; não deve ser possível colocar determinada resposta em mais de uma categoria do conjunto.

Um conjunto pode ser composto por mais de duas categorias, desde que essas regras não sejam violadas. Por exemplo, um exame superficial das respostas enumeradas acima logo mostra que "não menciona negros" inclui muitas respostas diferentes. A pessoa que não deseja falar sobre os vizinhos por temor de perturbações ou aquela que disse que era nova no bairro são muito diferentes daquelas que mencionam outros grupos, e não os negros. Isso indica a necessidade de subcategorias para distinguir entre os diferentes tipos de respostas que não mencionam os negros. Por exemplo, pode parecer psicologicamente significativo distinguir entre pessoas que não mencionam negros, mas mencionam outros grupos, e pessoas que não mencionam nem os negros nem outros grupos. Nesse caso, o conjunto estaria composto por três categorias:

Menciona negros (com ou sem menção de outros grupos)  
 Não menciona negros, mas menciona outros grupos  
 Não menciona negros, nem outros grupos

É preciso decidir quanto a essa extensão. A observação das respostas reais sugere que esse conjunto de categorias ainda não faz justiça ao seu conteúdo. Os grupos não-negros mencionados apresentam várias diferenças significativas: algumas pessoas descrevem os grupos através de termos econômicos (pessoas pobres); outras fazem referência a atributos pessoais ("duros"); outros mencionam grupos religiosos ou nacionais (judeus, irlandeses). Cada um desses atributos, escolhidos pelas pessoas para descrever os grupos de que estão conscientes, apresenta um possível princípio de classificação que poderia conduzir a outro conjunto de categorias, mais ou menos como no caso dos negros. Evidentemente, em princípio é possível usar todos esses atributos para o estabelecimento de conjuntos de categoria. Na prática, isso é frequentemente pouco econômico e pouco compensador, pois nem todos esses princípios de classificação se ligam aos objetivos da pesquisa.

Se, como o supusemos, o estudo se interessa pela consciência da presença dos negros, poderia parecer que todos

esses outros princípios de classificação não são significativos. Todavia, essa opinião se baseia em noção muito restrita da tarefa de análise. O fato de certa porcentagem mencionar os negros ainda não é uma indicação suficiente de sua consciência dos negros. Pode ser que qualquer outro grupo étnico do bairro tivesse sido mencionado da mesma forma. O fato de haver ou não uma consciência específica da presença de negros só pode ser decidido se se fizer uma codificação da menção de outros grupos raciais, nacionais ou religiosos. No entanto, a classificação das respostas através de menção de status econômico ou atributos pessoais de pessoas do bairro provavelmente não teria qualquer função.

A fim de apresentar essas classificações adicionais de referências a outros grupos étnicos, é preciso estabelecer outros conjuntos de categorias. No entanto, cada um desses conjuntos deve obedecer às regras apresentadas na página 441. A lista seguinte, que constitui o "código" para a pergunta, poderia ser o resultado final:

#### *Grupos raciais*

1. Negros: mencionados ....; não mencionados ....
2. Outros grupos raciais: mencionados ....; não mencionados ....

#### *Grupos de nacionalidade*

3. Irlandeses: mencionados ....; não mencionados ....
4. Italianos: mencionados ....; não mencionados ....
5. Outros grupos nacionais: mencionados ....; não mencionados ....

#### *Grupos religiosos*

6. Judeus: mencionados ...; não mencionados ...
7. Católicos: mencionados ....; não mencionados ....
8. Protestantes: mencionados ....; não mencionados ....
9. Outros grupos religiosos: mencionados ....; não mencionados ....

10. *Menção dos grupos acima enumerados*

Menciona um ou mais .....; não menciona qualquer deles, mas menciona outros grupos humanos .....; não menciona qualquer deles, afirmando "eu não conheço" .....; não responde à pergunta .....

Observe-se que, embora estejamos lidando com as respostas a apenas uma pergunta, nossos interesses específicos de resposta nos levaram a *dez* conjuntos de categorias para classificar as pessoas. Cada uma das pessoas que respondem pode ser colocada em uma das duas categorias de cada um dos nove primeiros conjuntos, e em uma das quatro categorias do conjunto final.

Como o fato de não assinalar um "mencionado" específico pode ser considerado como suposição do correspondente "não mencionado", a aparência do código pode ser simplificada eliminando-se todas as categorias de "não mencionado"; no entanto, cada um dos nove primeiros conjuntos continua a ser um conjunto de duas categorias, uma das quais indicada por um sinal e a outra por ausência de sinal. Todavia, essa simplificação nem sempre é boa. O fato de não assinalar uma determinada categoria "mencionado" pode representar um esquecimento na codificação. Esses esquecimentos podem ser desestimulados (e certamente identificáveis) através da exigência de que a colocação de uma pessoa numa categoria "não mencionado" deva exigir um ato tão positivo quanto a sua colocação numa categoria "mencionado". Além disso, para fazer uma pequena antecipação, quando se faz a codificação de tais dados para a tabulação por máquina, é muito desejável que todas as categorias tenham uma identificação explícita. As máquinas mais usadas contam, separadamente, o número de casos em cada categoria. A soma de contagens em todas as categorias de um conjunto deve ser igual ao número total de casos; se a categoria "não mencionado" e a "mencionado" são explicitamente codificadas, podemos verificar se todos os casos foram contados.

Várias qualidades desse código devem ser indicadas. Em primeiro lugar, para criar os dez conjuntos de categorias, foram usados quatro critérios fundamentais de classificação: menção de grupos raciais *versus* não menção de grupos raciais; menção de grupos nacionais *versus* não menção de grupos nacionais; menção de grupos religiosos *versus* não

menção de grupos religiosos; menção de qualquer desses tipos de grupos *versus* não menção de qualquer desses grupos. Deve-se notar que cada um desses quatro critérios de classificação permite certo número de conjuntos de categorias. Assim, "menção de grupos nacionais *versus* não menção de grupos nacionais" inclui os seguintes conjuntos de categorias: menção de irlandeses *versus* não menção de irlandeses; menção de italianos *versus* não menção de italianos; menção de outros grupos nacionais *versus* não menção de outros grupos nacionais.

Em segundo lugar, as categorias em cada conjunto são mutuamente exclusivas, e também incluem todas as possibilidades significativas para esse conjunto de categorias, de forma que cada resposta possa ser colocada, sem ambigüidade, numa categoria ou outra; por exemplo, determinada pessoa menciona ou não menciona os negros. (Pode ter também mencionado outros grupos raciais, bem como judeus e irlandeses. Cada um desses elementos de sua resposta é considerado separadamente nesse código, mas cada um deles pode ser claramente colocado no seu conjunto significativo de categorias.)

Em terceiro lugar, o código total é exaustivo na medida em que existe um lugar para cada resposta possível (sobretudo porque existe a categoria residual que "aceita tudo", que é a de número 10). Todavia, em quarto lugar, não é exaustiva em função de todos os possíveis princípios de classificação que poderiam ser aplicados às respostas. Além dos já eliminados como não significativos para o problema da pesquisa, não inclui, por exemplo, as distinções entre a maneira, depreciativa ou não, de fazer referência aos grupos.

Finalmente, deve-se discutir a inclusão na lista das categorias "católicos" e "protestantes." A tabulação posterior mostrou que 3 por cento das pessoas mencionaram os católicos e 0 por cento os protestantes. Na realidade, isso tinha sido previsto no estabelecimento das categorias. A idéia de inclusão desses dois grupos decorreu, em parte, do conhecimento da composição real do bairro e, em parte, da expectativa e interesse na demonstração de que esses dois grupos, embora tivessem muitos membros no bairro, não eram considerados como tão separados ou distintos quanto os outros grupos.

Neste caso, a "categorização de dados ausentes" — isto é, a apresentação de uma categoria para uma resposta cuja ocorrência não se espera — pode conduzir a um resultado óbvio, isto é, a demonstração de que os membros de grupos dominantes ou socialmente aceitos freqüentemente não são percebidos através de sua participação nesses grupos. No entanto, o princípio aqui existente está longe de ser óbvio, e é muito freqüentemente esquecido. Baseia-se na noção de que as idéias para categorização sempre devem vir de duas fontes: um conhecimento íntimo das provas existentes e conhecimento geral e análise prévia dos possíveis tipos de respostas, baseada em considerações teóricas, lógicas ou práticas. A aplicação dessa regra ao estabelecimento de categorias conduz à possível descoberta de ausência significativa de alguma resposta, o que poderia, de outra forma, passar despercebido.

Existem tipos de estudo e tipos de dados para os quais o estabelecimento de categorias é ainda mais simples que neste exemplo. Quando se usam perguntas para assinalar, ou qualquer forma de escala de avaliação, aplicada numa situação de observação, o analista dos dados praticamente não precisa dar atenção ao problema do estabelecimento de categorias. Seus dados foram pré-categorizados nos estágios anteriores da pesquisa. Isso também é verdade em todos os estudos em que as observações são registradas em categorias prescritas. Essa pré-categorização, quando realizada como parte do desenvolvimento de instruções para pesquisa, é, naturalmente, muito semelhante ao processo empregado no estabelecimento de categorias depois da coleta de dados.

#### DEFINIÇÃO DE CATEGORIAS COMPLEXAS

Em algumas situações, o estabelecimento de categorias é consideravelmente mais difícil e exige muito mais tempo que o exemplo do bairro, acima apresentado. O que, neste caso, tornou o estabelecimento de categorias um processo relativamente fácil foi o fato de as respostas serem relativamente simples e nítidas; as categorias poderiam ser facilmente definidas de maneira não-ambígua. Embora as categorias devam ser sempre definidas dessa maneira, a tarefa é muito mais difícil com alguns tipos de conteúdo que com outros.

Por exemplo, Merton e seus colaboradores (estudo inédito) perguntaram a senhoras brancas: "O que acha que os moradores de cor sentem por viverem na mesma comunidade com brancos?" E perguntaram a senhoras negras: "O que acha que os moradores brancos sentem por viverem na mesma comunidade com negros?" As respostas iam desde atribuição de atitudes muito favoráveis até atribuição de atitudes muito desfavoráveis. Eis alguns exemplos disso:

Desejam viver aqui, de forma que possam dizer que são iguais.

Parecem muito amistosos conosco.

Gostam da idéia.

Penso que não se incomodam; nós nos entendemos.

Alguns gostam disso, outros não.

Pensam que isso os rebaixa.

Não estou em contato com eles, e por isso não sei.

Detestam isso.

De acordo com a discussão anterior, não seria difícil criar um conjunto simples de categorias, baseado no princípio de classificação de atitude favorável *versus* atitude desfavorável atribuída à outra raça. Neste caso, seria necessário uma categoria para atitude favorável, uma para desfavorável, outra para neutra, além de uma categoria residual. No entanto, entre as atitudes favoráveis e desfavoráveis que são atribuídas, aparecem certas sutilezas de sentido. Uma pessoa que diz: "Desejam viver aqui, de forma que possam dizer que são iguais", revela algo diferente daquela que diz: "Gostam da idéia." De forma semelhante, aquela que diz: "Pensam que isso os rebaixa" pode querer dizer coisa diversa de quem diz: "Detestam isso." Qual é o critério distintivo? Aparentemente, em ambos os casos algumas das pessoas que respondem atribuem o suposto sentimento da outra raça a um valor que, segundo acreditam, é aceito por esse grupo. Por isso, a atribuição de tais valores foi usada como outro princípio de classificação. Criou-se um conjunto de categorias em que uma categoria representava a atribuição de atitude favorável da outra raça a um valor material, social ou espiritual, ou a um benefício que obtinham com a situação (ou a atribuição de sua atitude desfavorável a perda ou desvantagem correspondentes), e outra categoria representava uma

ausência de afirmações sobre valor. Para simplificação da operação subsequente de codificação, os dois conjuntos de categorias foram combinados da seguinte maneira:

Atitude favorável atribuída aos membros da outra raça, explicada através de valores ou benefícios que obtêm com o fato de viverem no mesmo conjunto residencial com a raça da pessoa que responde.

Atitude favorável atribuída à outra raça, sem explicação explícita em termos de valor.

Atitude neutra ou acomodatória atribuída à outra raça.

Atitude desfavorável atribuída aos membros da outra raça, e explicada em função de valores ou benefícios que obtêm com a segregação.

Atitude desfavorável atribuída à outra raça, sem explicação explícita em termos de valor.

Outras respostas, ausência de resposta, não sabe.

Antes de usar uma lista de tais categorias, é necessário especificar, tão exatamente quanto possível, o conteúdo de cada categoria. Isso é feito através da adição de uma ou duas sentenças explicativas a cada categoria e da exemplificação de seu sentido.

Por exemplo, no caso da primeira categoria da lista acima apresentada, a sentença explicativa era a seguinte: Tais respostas supõem que a outra raça obtêm algo com o fato de viver com a raça da pessoa que responde. Por ex.: "Desejam viver aqui, de forma que possam dizer que são iguais."

O trabalho com essas categorias complexas exige muito esforço e cuidado na classificação. Mesmo quando as categorias são cuidadosamente estabelecidas, seu emprego apresentará mais problemas que o uso de categorias definidas de forma restrita e exata. Se uma pessoa responde, por exemplo: "Gostam de tudo que encontram aqui, e sabem por que", não se sabe se essa afirmação supõe ou não um benefício. É preciso estabelecer outras regras, a fim de lidar com tais respostas. Neste caso, uma regra adequada para a primeira categoria poderia ser a seguinte: "Esta categoria aplica-se apenas às afirmações que especificam a natureza do valor ou benefício."

#### SELEÇÃO DE PRINCÍPIOS DE CLASSIFICAÇÃO PARA A CATEGORIZAÇÃO DE MATERIAL ASSISTEMÁTICO

Surgem problemas especiais na categorização de material assistemático — por exemplo, em protocolos de observação, histórias de caso, discursos de agitadores, entrevistas assistemáticas, etc. Já aludimos a essas dificuldades nos capítulos anteriores, ao discutir a observação participante e o emprego, na pesquisa, de documentos pessoais e de conteúdo de comunicação.

Num estudo que empregue instrumentos assistemáticos para a obtenção de dados significativos para hipóteses ou problemas de pesquisa claramente formulados, os princípios adequados para a classificação de respostas são claramente prescritos pela natureza do estímulo ou pelas perguntas e respostas. No entanto, ao trabalhar com provas assistemáticas, o primeiro problema é decidir quanto aos aspectos do material que devem ser categorizados — isto é, quais os princípios de classificação que devem ser usados para o estabelecimento de conjuntos de categorias.

O estabelecimento de princípios de classificação é difícil, sobretudo, em estudos exploratórios, pois tais estudos, por definição, não partem de hipóteses explícitas. No momento da coleta de dados, o pesquisador não sabe quais são os aspectos que se revelarão mais importantes. Por isso, geralmente precisa coligir um grande número de dados; assim, na análise, tem o problema de lidar não apenas com material assistemático, mas também com grande quantidade de tal material, grande parte do qual pode não ter significação para seu objetivo.

O primeiro passo na análise de dados de um estudo exploratório é criar hipóteses de trabalho que permitam princípios de classificação. Geralmente, o pesquisador lê cuidadosamente todo o seu material e mantém-se alerta para indicações nos dados. Existem vários processos que podem ajudá-lo em sua tarefa. Um deles é estudar, quando possível, o material de um grupo que apresente contraste com aquele que está estudando, a fim de obter idéias a respeito de diferenças importantes entre os dois. Por exemplo, num estudo sobre delinquência, é adequado ler registros de casos

em instituições sociais, não apenas a respeito de delinquentes, mas também de outros jovens que estejam sob os cuidados de tal instituição. Vale dizer, contrasta casos que diferem, nitidamente, na característica que está sendo pesquisada, a fim de ver quais as outras diferenças concomitantes que podem ser percebidas nos casos contrastantes.

O pesquisador pode separar seus casos em grupos que parecem comuns, e depois perguntar a si mesmo o que o levou a pensar que os colocados num único grupo sejam semelhantes. Por exemplo, Chein *et al.* (1952), num estudo sobre opiniões de eminentes educadores judeus e assistentes sociais a respeito de alguns problemas de educação de judeus, consideraram adequado selecionar as pessoas, inicialmente, em função de sua visão total do sentido de ser judeu. O pesquisador pode verificar que agrupou os seus casos a partir de características comuns; pode, depois, examiná-los, a fim de verificar se os que têm características semelhantes passaram por experiências semelhantes. Pode descobrir que seu agrupamento se baseia em experiências semelhantes; depois, pode rever os casos, a fim de verificar se tais experiências semelhantes parecem ter levado a consequências semelhantes.

Outro processo que pode estimular a formulação de hipóteses de trabalho é observar questões que parecem surpreendentes diante de expectativas do senso comum ou da teoria, e depois procurar explicações para o fenômeno surpreendente. Por exemplo, Lambert e Bressler (1957), num estudo sobre estudantes hindus nos Estados Unidos, observaram que esses estudantes freqüentemente pareciam mais perturbados por perguntas, a respeito de problemas hindus, propostas por norte-americanos relativamente bem informados, que por estereótipos ou pela ignorância dos que não tinham informação a respeito da Índia; mesmo quando as perguntas não eram feitas com intenção inamistosa, os estudantes pareciam interpretá-las como crítica hostil. Essa observação conduziu à hipótese de que, para pessoas de países coloniais ou que foram coloniais, criam-se algumas "áreas sensíveis." Essas "áreas sensíveis" são aspectos da cultura nacional que, historicamente, constituíram objeto de crítica hostil por parte do país ou países dominantes e foram usadas como desculpa para a manutenção do país na situação

de subordinação. Qualquer referência a essas áreas, qualquer que seja a intenção da pessoa que as apresenta, traz consigo as conotações históricas de hostilidade. No entanto, o estranho (nesse caso, o norte-americano) precisa ter algum conhecimento do país de origem do estudante para ter ciência dessas áreas sensíveis; por isso, as perguntas de pessoas relativamente bem informadas apresentam maior tendência para atingir tais áreas. As perguntas de pessoas sem qualquer informação, que geralmente não se referem a essas áreas sensíveis, são atribuídas à ignorância; as perguntas referentes às áreas sensíveis, que só podem ser propostas por pessoas com alguma informação a respeito do país, são interpretadas como provas de hostilidade.

No entanto, mesmo com hipóteses claras, a análise de material assistemático apresenta problemas especiais. Como, por definição, o material não é o do tipo em que a mesma pergunta pode ser respondida por todos, ou em que as mesmas observações são feitas a respeito de todas as pessoas, sempre existe a possibilidade de que a informação sobre determinado aspecto não exista em alguns documentos. De outro lado, tende a haver muito material que não é diretamente significativo para as hipóteses. Além disso, existe o problema de decidir quanto ao tamanho de unidades de material a que devem ser aplicadas as categorias. Se os dados foram obtidos com um instrumento sistemático — por exemplo, um questionário —, geralmente cada pergunta apresenta uma unidade natural para a categorização, embora alguns conjuntos de categorias possam aplicar-se a unidades maiores, — por exemplo, o questionário como um todo. No entanto, quando lidamos com material assistemático, não existem essas unidades "naturais" e adequadas. Por exemplo, se o pesquisador emprega registros de casos mantidos por agências de serviço social, pode categorizar cada afirmação feita pelo cliente, ou cada sessão do cliente com o assistente social, ou todo o registro do caso. Precisa decidir qual dessas unidades é mais adequada para as respostas às suas perguntas específicas de pesquisa.

#### Codificação: A Categorização dos Dados

A codificação é o processo técnico pelo qual são categorizados os dados. Através da codificação, os dados brutos