2. Uma das primeiras provas que vincularam o câncer de pulmão a fumar cigarros foi publicada por Doll (1955). Ele tabulou o número médio de cigarros fumados por pessoas de 11 países, em 1930, e o número de mortes por câncer de pulmão entre homens em 1950. A contagem do número de mortes foi feita 20 anos após a avaliação do consumo de cigarros, pois parece óbvio que decorreriam anos para que fosse constatada uma relação de causa e efeito, caso existisse. Em virtude de muito poucas mulheres fumarem em 1930, também pareceu melhor relacionar a intensidade do hábito de fumar à morte de homens. A tabela no final das Questões para Discussão constitui uma adaptação dos importantes resultados obtidos por Doll.
   (a) Examine os resultados. O que as duas colunas de números parecem indicar?
   (b) Trace um gráfico relacionando as duas variáveis, como aquele mostrado na Figura 2.3. O que ele mostra?
   (c) Calcule agora a relação exata entre as duas variáveis usando a fórmula do coeficiente *r* de Pearson dada no Apêndice B. Quais são o sinal e a magnitude exatos do coeficiente de correlação que você obteve?
3. As análises que você realizou na Questão 2(c) permitem a conclusão de que o fumo causa câncer de pulmão? Se o coeficiente de correlação fosse maior, digamos 1,95, você teria mais certeza quanto à relação de causa e efeito? Caso considere que esses dados não justifiquem que o fumo cause câncer de pulmão, de que outra maneira você poderia explicar os resultados?
4. Faça uma lista de pares de variáveis que você julga possuírem uma correlação elevada (seja ela positiva ou negativa), porém entre as quais você considera haver pouca possibilidade de existir uma conexão causal. Como você poderia determinar se a correlação indica efetivamente uma relação de causa e efeito?

País	Consumo de cigarros em 1930	Mortes em 1950 por milhão*
Austrália	480	180
Canadá	500	150
Dinamarca	380	170
Finlândia	1100	350
Grã-Bretanha	1100	460
Holanda	490	240
Islândia	230	60
Noruega	250	90
Suécia	300	110
Suíça	510	250
Estados Unidos	1300	200

* De câncer de pulmão.

CAPÍTULO 3

TÉCNICAS DE PESQUISA: EXPERIMENTAÇÃO

O QUE É UM EXPERIMENTO?
Vantagens da Experimentação
Por que São Feitos Experimentos?
VARIÁVEIS
Variáveis Independentes
Variáveis Dependentes
Variáveis de Controle
Identifique as Variáveis
Mais de uma Variável Independente
Mais de uma Variável Dependente
DELINEAMENTOS EXPERIMENTAIS
Delineamentos de Grupo
Delineamentos de Sujeito Único
Delineamentos de Baixo *n*
Delineamentos Mistos
Condições de Controle
Armadilhas
Quase-Experimentos
Do Problema à Experiência
DADOS
Obtenção dos Dados
Análise dos Dados
Relato dos Dados
RESUMO
TERMOS-CHAVE
QUESTÕES PARA DISCUSSÃO

Ninguém acredita em uma hipótese, a não ser quem a formula, porém todos acreditam em um experimento, exceto quem o realiza. (W. I. B. Beveridge)

Imagine que você seja um aluno em um curso de psicologia experimental e tenha recebido a seguinte tarefa: ir à biblioteca e "apossar-se" de uma mesa, impedindo que qualquer outra pessoa sente-se a ela durante todo o tempo que você puder. Você deve empregar apenas meios não-verbais e não-violentos para executar a sua tarefa. Para isso, você poderia esperar na biblioteca lotada até que uma mesa vagasse, sentando-se rapidamente e passando a espalhar seus livros, roupas e outros pertences por toda a mesa, esperando que essa desordem possa afastar as outras pessoas. Após algum tempo, digamos, cerca de 15 minutos, alguém finalmente senta à sua mesa, encerrando sua tarefa. Você realizou um experimento?

Antes de responder a essa pergunta, vamos esboçar os principais critérios para um experimento, os quais foram discutidos resumidamente nos capítulos precedentes. Um experimento ocorre quando o ambiente é sistematicamente manipulado, de modo que o efeito causal dessa manipulação em um comportamento possa ser observado. Aspectos do ambiente que não são de interesse e, portanto, não são manipulados, permanecem constantes, de modo a não influenciar o resultado do experimento. Podemos, então, concluir que o comportamento resultou da manipulação realizada. Precisamos explicar dois termos especiais introduzidos brevemente no Capítulo 1 – variáveis *independentes* e *dependentes* –, a fim de descrever como o ambiente é manipulado e como o comportamento é observado.

▶ O QUE É UM EXPERIMENTO?

Muitos alunos ficam surpresos ao descobrir que as ações descritas em nosso exercício envolvendo uma mesa na biblioteca não constituem um experimento. Todos os experimentos requerem pelo menos duas características essenciais, as variáveis independente e dependente que acabamos de mencionar. A **variável dependente** é a medida da resposta de um experimento que é *dependente* da pessoa. Nesse caso, o tempo decorrido até que alguém se sentasse à mesa é a variável dependente ou medida da resposta. A **variável independente** é uma manipulação do ambiente controlado pelo experimentador. Nesse caso, é o ato de espalhar objetos sobre a mesa.

Porém, um experimento precisa ter pelo menos dois valores, ou **níveis**, do ambiente. Esses níveis podem diferir em um sentido quantitativo (objetos espalhados somente em uma parte da mesa *versus* objetos espalhados por toda a mesa) ou refletir uma diferença qualitativa (a pessoa que se apossa da mesa demonstra uma expressão amigável e convidativa em oposição a uma expressão rígida e de proibição). É necessário que ao menos duas condições sejam comparadas entre si para determinar se a variável independente (parte da mesa coberta ou a expressão facial) produz uma alteração no comportamento ou no resultado. Algumas vezes, esses dois níveis poderiam ser simplesmente a existência ou a ausência de uma manipulação. O exemplo da biblioteca deixa de atender a esse critério, pois envolve apenas um nível da variável independente.

Como poderíamos alterar o procedimento para obter um experimento? O modo mais simples seria sentar-se novamente, dessa vez sem espalhar nenhum objeto. Então, nossa variável independente teria os dois níveis necessários: a mesa com os objetos espalhados e a mesa desimpedida, sem objetos espalhados. Agora temos algo para comparar à primeira condição.

O experimento tem três resultados possíveis: (1) os objetos sobre a mesa resultam em um intervalo de tempo maior antes de ela ser invadida por outra pessoa; (2) o tempo até a invasão é o mesmo, com ou sem artigos espalhados sobre a mesa; e (3) espalhar objetos resulta em um menor período de tempo até que ocorra a invasão. Sem o segundo nível da variável independente (a mesa sem objetos espalhados), esses três resultados não podem ser formulados. De fato, é impossível afirmar algo sobre a eficácia dos objetos espalhados para assegurar a posse de mesas na biblioteca até que os dois níveis da variável independente sejam testados.

Quando esse experimento na biblioteca é executado adequadamente, é obtido o primeiro resultado possível. Uma mesa pode ser mais bem protegida por uma pessoa com objetos espalhados do que somente por uma pessoa.

Podemos constatar, então, que os experimentos precisam ter, no mínimo, as variáveis independente e dependente. Os métodos de pesquisa discutidos no capítulo anterior não permitem ou exigem a manipulação do ambiente, porém, antes que um **experimento** possa ser estabelecido, são necessárias variáveis independentes possuindo, no mínimo, dois níveis.

Vantagens da Experimentação

A principal vantagem dos experimentos sobre os métodos discutidos no Capítulo 2 consiste no melhor controle das variáveis externas ao foco principal da pesquisa. No experimento ideal, nenhum fator (variáveis), exceto o estudado, pode influenciar o resultado; no jargão da psicologia experimental, dizemos que esses outros fatores estão *controlados*. Se, como no caso de experimentos ideais, todos os fatores, com exceção de um (aquele sob investigação), são mantidos constantes, podemos concluir logicamente que todas as diferenças no resultado devem ser causadas pela manipulação da variável independente. À medida que são alterados os níveis da variável independente, as diferenças resultantes na variável dependente podem ocorrer somente porque a variável independente variou. Em outras palavras, variações da variável independente causam as variações observadas na variável dependente. No exemplo da biblioteca, poderíamos desejar manipular a expressão facial da pessoa que se "apossou" da mesa. Para controlar variáveis externas, teríamos de prestar muita atenção em outros fatores que poderiam prejudicar nossa possibilidade de fazer afirmativas sobre a causação. Nesse caso, poderíamos querer depender de apenas uma assistente para apossar-se da mesa, enquanto durasse o experimento, ou, então, estabelecer objetivamente que nossas assistentes são, por exemplo, igualmente atraentes. Também poderíamos decidir *controlar* o gênero, incorporando-o como uma variável independente adicional ou usando somente assistentes de pesquisa do sexo feminino (ou masculino). Criar experimentos, de modo que possa existir somente uma explicação do resultado, ocupa uma posição central no método experimental. Enquanto os métodos de pesquisa não-experimentais são limitados a descrever e correlacionar dados, os experimentos permitem afirmar sobre a causa – isto é, a variável independente A (expressão facial) faz com que a variável B (tempo decorrido até alguém se sentar à mesa) se altere. Nesse experimento, esperaríamos que o tempo decorrido fosse menor quando a assistente assumisse uma expressão amigável e convidativa do que quando a expressão da assistente fosse séria e desagradável.

Portanto, em princípio, os experimentos levam a afirmativas a respeito da causação. Na prática, essas afirmativas nem sempre são verdadeiras. Nenhum experimento alcança 100% de sucesso eliminando ou mantendo constantes todas as demais fontes de variação que acompa-

nham aquela estudada. No entanto, os experimentos eliminam mais variáveis externas ao foco principal da pesquisa do que outros métodos de pesquisa. Discutiremos posteriormente, neste capítulo, maneiras específicas pelas quais os experimentos limitam as variáveis estranhas.

Uma outra vantagem das experiências é a economia. Usar a técnica da observação natural exige que o cientista espere pacientemente até que ocorram as condições de interesse. Se você vivesse em Trondheim, na Noruega – perto do Círculo Ártico – e desejasse estudar como o calor afeta a agressão, depender do sol para que produza temperaturas elevadas exigiria paciência considerável e muito tempo. O experimentador controla a situação criando as condições de interesse (diversas intensidades de calor em um ambiente de laboratório), obtendo desse modo dados de uma maneira rápida e eficiente.

Por que São Feitos Experimentos?

As mesmas razões gerais que se aplicam à condução de qualquer pesquisa também explicam por que os psicólogos realizam experimentos. Na pesquisa básica, os experimentos são realizados para testar teorias e fornecer o banco de dados para explicações do comportamento. Esses tipos de experimentos normalmente são bem-planejados, o investigador possui uma idéia clara do resultado previsto. Os denominados **experimentos críticos** tentam colocar em oposição duas teorias que fazem previsões diferentes. Um resultado favorece a teoria A; o outro, a teoria B. Portanto, em princípio, o experimento determinará qual teoria rejeitar e qual manter. Na prática, esses experimentos críticos não proporcionam um resultado muito bom, porque os defensores da teoria rejeitada demonstram criatividade para apresentar explicações que desacreditem a interpretação desfavorável do experimento. Um exemplo de tal explicação encontra-se em um estudo de como as pessoas esquecem. As duas principais explicações para o esquecimento são (1) as lembranças diminuem ou se apagam ao longo do tempo, do mesmo modo que uma lâmpada incandescente se apaga quando a eletricidade é desligada (essa explicação é denominada "esvaecimento de traço") ou (2) as lembranças nunca se apagam; porém, por esse motivo, elas interferem entre si, causando confusão ou interferências mútuas. Em um experimento crítico simples, poderia-se variar o tempo entre a introdução de itens sucessivos na memória, mantendo constante o número de itens (Waugh e Norman, 1965). A memória deveria piorar ao longo do tempo, de acordo com os teóricos do esvaecimento de traço, por haver mais tempo para as lembranças se apagarem. Porém, em virtude de o número de itens permanecer o mesmo, independentemente da época em que são introduzidos na memória, a teoria da interferência prevê que não há diferenças no esquecimento. Quando esse experimento é realizado, não há diferença de memória; isso pareceria anular a explicação do esvaecimento de traço. No entanto, a resposta dos cientistas que adotam essa teoria é que o tempo adicional concedido entre os itens permite que as pessoas ensaiem, isto é, repitam o item para si, o que evita o esquecimento.

Os pesquisadores, com menor freqüência, realizam um experimento na falta de uma teoria convincente apenas para observar o que acontece; podemos denominá-lo um **experimento "E se...?"**. Os alunos freqüentemente criam experimentos "E se...?", pois eles não exigem conhecimento teórico ou de um conjunto de dados da literatura existente e podem ser formulados com base em experiências e observações pessoais. Alguns cientistas desaprovam os experimentos "E se...?"; para eles, a principal objeção é a ineficiência dessas tentativas. Se, conforme ocorre com freqüência, nada ocorrer de importante em um experimento "E se...?" – por exemplo, a variável independente não exerce efeito-nada se ganha com o experimento. Em contraste, se nada de relevante ocorrer em um experimento cuidadoso para o qual uma teoria prevê que algo acontecerá, constatar a inexistência de diferença pode ser útil. Devemos admitir ter tentado experimentos "E se...?". A maioria deles não deu certo, mas eram divertidos. Aconselhamos que você verifique com seu professor antes de tentar um experimento "E se...?". Ele provavelmente pode lhe fornecer uma estimativa da possibilidade de você descobrir algo ou pode até mesmo conhecer os resultados de um experimento similar que já tenha sido realizado.

Isso nos leva à última razão principal para realizar experimentos em pesquisa básica, o qual consiste em repetir ou reproduzir uma descoberta anterior. Um experimento único *per si* é muito menos convincente do que uma série de experimentos relacionados. A replicação mais simples consiste na repetição direta de um experimento existente, sem alterar os procedimentos. Replicações diretas são especialmente úteis quando o experimento original representava uma iniciativa realmente inovadora. No entanto, geralmente a melhor maneira de replicar um experimento consiste em *ampliar* o procedimento anterior, acrescentando algo novo e ao mesmo tempo conservando algo do antigo. Portanto, parte da replicação é uma repetição literal, porém a parte nova enriquece o conhecimento científico. Esse tipo de replicação demonstra a generalidade de um resultado, mostrando a maneira como é (ou não é) mantido nas diferentes variáveis independentes. O conceito de replicação e suas várias formas são discutidos com mais detalhes no Capítulo 11.

▶ VARIÁVEIS

Variáveis são as engrenagens que fazem os experimentos funcionarem. A seleção e a manipulação eficazes das variáveis fazem a diferença entre um bom e um mau experimento. Esta seção cobre os três tipos de variáveis que precisam ser analisados cuidadosamente antes do início de um experimento: *variáveis independentes, dependentes* e *de controle*. Concluímos discutindo experimentos que possuem mais de uma variável independente ou dependente.

Variáveis Independentes

Nos experimentos de verdade, **variáveis independentes** são as *manipuladas* pelo experimentador. O brilho de uma luz, a altura de um som, a temperatura de uma sala, o número de bolinhas de alimento dadas a um rato são todas variáveis independentes, pois o experimentador determina a qualidade e a quantidade. As variáveis independentes são selecionadas porque um experimentador acredita que elas provocarão mudanças de comportamento. Aumentar a intensidade de um som deve aumentar a velocidade com que as pessoas reagem a ele. Aumentar o número de bolinhas dadas a um rato quando pressiona uma barra deveria aumentar o número de vezes que a barra é pressionada. Quando uma variação no nível (quantidade) de uma variável independente causar uma mudança de comportamento, dizemos que o comportamento está sob controle da variável independente.

A falha de uma variável independente em controlar o comportamento, muitas vezes denominada **resultado nulo**, pode ter mais de uma interpretação. Primeiro, o experimentador pode ter intuído incorretamente que a variável independente era importante: o resultado nulo pode estar correto. A maioria dos cientistas aceitará essa interpretação apenas de modo relutante e, portanto, são comuns as explicações alternativas que se seguem ao resultado nulo. O experimentador pode não ter feito uma manipulação válida da variável independente. Digamos que você esteja realizando um experimento com crianças da segunda série e que sua variável independente seja o número de pequenos confeitos (M&Ms, balas de goma) que elas recebem após cada resposta correta. Algumas crianças recebem somente um confeito, ao passo que outras recebem dois. Você não constata uma diferença de comportamento. No entanto, se sua variável independente tivesse uma faixa maior de variação – isto é, de um a dez confeitos –, talvez você tivesse obtido uma diferença. Sua manipulação poderia não ter sido suficiente para revelar um efeito da variável independente. Ou talvez, com o seu

desconhecimento, as crianças estiveram em uma festa de aniversário um pouco antes de o experimento ter sido iniciado, e seus pequenos estômagos estavam repletos de sorvete e bolo. Nesse caso, talvez dez confeitos não tivessem nenhum efeito. É por esse motivo que, em estudos de aprendizagem animal com alimento como conseqüência, os animais são privados de alimentação antes do início do experimento.

Podemos observar que os experimentadores precisam ser cuidadosos, a fim de produzir uma grande manipulação da variável independente. A falha em conseguir isso constitui uma causa comum dos resultados nulos. Em virtude de não haver um meio para determinar se a manipulação falhou ou se os resultados nulos estão corretos, os experimentadores não podem chegar a uma conclusão a respeito do efeito da variável independente sobre a variável dependente. Outras causas comuns dos resultados nulos relacionam-se às variáveis dependentes e de controle, que passamos a estudar agora.

Variáveis Dependentes

A **variável dependente** é a medida obtida de um experimento que *depende* da reação do indivíduo à nossa manipulação do ambiente. Em outras palavras, o comportamento do indivíduo é observado e registrado pelos experimentadores e depende da variável independente. O tempo decorrido antes de uma pessoa sentar-se a uma mesa de posse de um assistente de pesquisa, a velocidade de um verme arrastando-se em um labirinto, o número de vezes que um rato pressiona uma barra são variáveis dependentes, porque dependem do modo pelo qual o experimentador manipula o ambiente. No exemplo da biblioteca, poderíamos prever que uma pessoa demonstraria maior relutância em sentar-se a uma mesa de posse de um assistente que exibisse uma expressão desagradável do que se o assistente assumisse uma expressão simpática. Nesse caso, o comportamento da pessoa é *dependente* da expressão que instruímos que seja adotada pelo assistente. O tempo decorrido até a pessoa sentar-se à mesa representa a variável dependente de interesse.

A **estabilidade** constitui nosso critério para uma boa variável dependente. Quando um experimento for repetido com exatidão – a mesma pessoa, os mesmos níveis da variável independente, e assim por diante –, a variável dependente deveria proporcionar o mesmo resultado constatado previamente. A falta de estabilidade pode ocorrer por causa de certo déficit no modo como medimos algumas variáveis dependentes. Suponha que desejamos medir os pesos de um objeto em gramas – por exemplo, uma vela – antes e após ter sido acesa durante 15 minutos. Usamos uma escala que opera por meio de uma mola que move um ponteiro. A mola se contrai quando estiver fria e se expande quando estiver quente. Desde que nossas medidas de peso sejam feitas em temperaturas constantes, elas serão confiáveis. Porém, se a temperatura variar enquanto os objetos estão sendo pesados, o mesmo objeto proporcionará leituras diferentes. Nossa variável dependente não possui estabilidade.

O resultado nulo muitas vezes pode ser causado por inadequações da variável dependente, mesmo se ela for estável. A causa mais comum é uma faixa de variação restrita ou limitada da variável dependente, de modo que ela fica "presa" ao maior ou ao menor valor de sua escala. Imagine que você esteja ensinando a um amigo um tanto desajeitado como jogar boliche pela primeira vez. Como você conhece, da Introdução à Psicologia, as conseqüências que podem aprimorar o desempenho, você se oferece para pagar uma cerveja a seu amigo cada vez que ele fizer um ponto. Seu amigo erra todas as jogadas e, então, você bebe as cervejas. Desse modo, você não pode mais oferecer as cervejas e, portanto, tem a expectativa de um desempenho inferior. Porém, como é impossível um desempenho pior do que errar todas as jogadas, você não consegue observar o resultado inferior. Seu amigo já se encontra na parte

mais baixa da escala. Isso é denominado **efeito de piso**. O problema oposto, obter 100% de acertos, chama-se **efeito de teto**. Os efeitos de piso e de teto (ver Capítulo 10) evitam que a influência de uma variável independente reflita-se, de modo preciso, na variável dependente.

Variáveis de Controle

Uma **variável de controle** é uma variável independente potencial, mantida constante durante um experimento por ser controlada pelo experimentador. Para qualquer experimento, a lista de variáveis de controle relevantes é muito grande, muito maior do que pode ser obtida na prática. Mesmo em um experimento relativamente simples – por exemplo, solicitar às pessoas que memorizem sílabas com três letras –, muitas variáveis devem ser controladas. A ordem do dia altera sua eficiência; idealmente, isso deveria ser controlado. A temperatura poderia ser importante, porque você poderia adormecer caso a sala de testes estivesse muito quente. O tempo decorrido desde sua última refeição também poderia afetar o desempenho da memória. A inteligência também está relacionada. A lista poderia ser ampliada. Na prática, um experimentador tenta controlar o maior número possível de variáveis importantes, esperando que o efeito de fatores não-controlados seja pequeno, relativamente ao efeito da variável independente. Embora seja sempre relevante exercer um controle rigoroso sobre fatores externos, isso se torna até mais decisivo quando a variável independente produz um efeito reduzido na variável dependente. Manter uma variável constante não é a única maneira de eliminar variáveis externas. Técnicas estatísticas (discutidas posteriormente no capítulo) também controlam as variáveis externas. No entanto, manter uma variável constante constitui a técnica experimental mais direta para o controle de fatores externos e, portanto, limitamos nossa definição de variáveis de controle somente a essa técnica. O resultado nulo ocorre freqüentemente em um experimento porque existe um controle insuficiente desses outros fatores, isto é, permitiu-se que variassem sistematicamente com a variável independente. Dependendo da relação entre uma variável externa e uma variável independente, essa variação não-controlada pode obscurecer ou aumentar o efeito da variável independente sobre a variável dependente de interesse. O problema das variáveis externas ocorre com mais freqüência em estudos conduzidos fora dos laboratórios, em que a habilidade para manter as variáveis de controle constantes reduz-se significativamente.

A variável INDEPENDENTE é MANIPULADA.
A variável DEPENDENTE é OBSERVADA.
A variável DE CONTROLE é mantida CONSTANTE.

Identifique as Variáveis

Em virtude de ser tão importante compreender as variáveis independentes, dependentes e de controle, incluímos alguns exemplos para que você teste sua compreensão. Identifique, em cada situação, os três tipos de variáveis. As respostas encontram-se após os exemplos. Não vale olhar antes!

1. Uma montadora de automóveis quer saber qual deve ser a intensidade do brilho do *brake light*, a fim de minimizar o tempo exigido do motorista do carro que vem atrás para perceber que o carro à frente está parando. É feito um experimento para obter a resposta. Identifique as variáveis.

2. Um pombo é treinado para bicar um disco de respostas se aparecer uma luz verde, e para não fazê-lo, caso apareça uma luz vermelha. As bicadas corretas são seguidas por acesso a grãos de alimento. Identifique as variáveis.
3. Um terapeuta tenta melhorar a imagem que um paciente possui de si mesmo. Toda vez que o paciente diz algo positivo a seu respeito, o terapeuta reconhece meneando a cabeça, sorrindo e sendo muito atencioso. Identifique as variáveis.
4. Um psicólogo social realiza um experimento para descobrir se homens ou mulheres atribuem índices menores de desconforto quando seis pessoas se aglomeram em uma cabine telefônica. Identifique as variáveis.

RESPOSTAS

1. Variável independente (manipulada): Intensidade (brilho) do *brake light*
   Variável dependente (observada): Tempo decorrido desde o aparecimento do *brake light* até o motorista do carro que vem atrás pisar no pedal de freio
   Variáveis de controle (constantes): Cor do *brake light*, forma do pedal de freio, força necessária para acionar o pedal de freio, iluminação externa etc.
2. Variável independente: Cor da luz (vermelha ou verde)
   Variável dependente: Número de bicadas no disco de respostas
   Variáveis de controle: Número de horas de privação de alimento, tamanho do disco, intensidade das luzes vermelha e verde etc.
3. Variável independente: Na realidade, esse não é um experimento, porque existe somente um nível da variável independente. Para torná-lo um experimento, precisamos de um outro nível – por exemplo, reconhecer comentários positivos sobre a sogra do paciente e desprezar os negativos Nesse caso, a variável independente seria o tipo de afirmativa elogiada.
   Variável dependente: Número (ou freqüência) de afirmativas
   Variáveis de controle: Ambiente do consultório, terapeuta
4. Variável independente: Sexo do participante[1]
   Variável dependente: Grau de desconforto
   Variáveis de controle: Dimensões da cabine telefônica, número de pessoas (seis) na cabine, tamanho das pessoas etc.

[1] O sexo é um tipo especial de variável independente denominado variável de sujeito, discutida posteriormente neste capítulo.

Mais de uma Variável Independente

É incomum deparar-se com um experimento divulgado em uma publicação de psicologia em que é usada somente uma variável independente (manipulada); o experimento típico

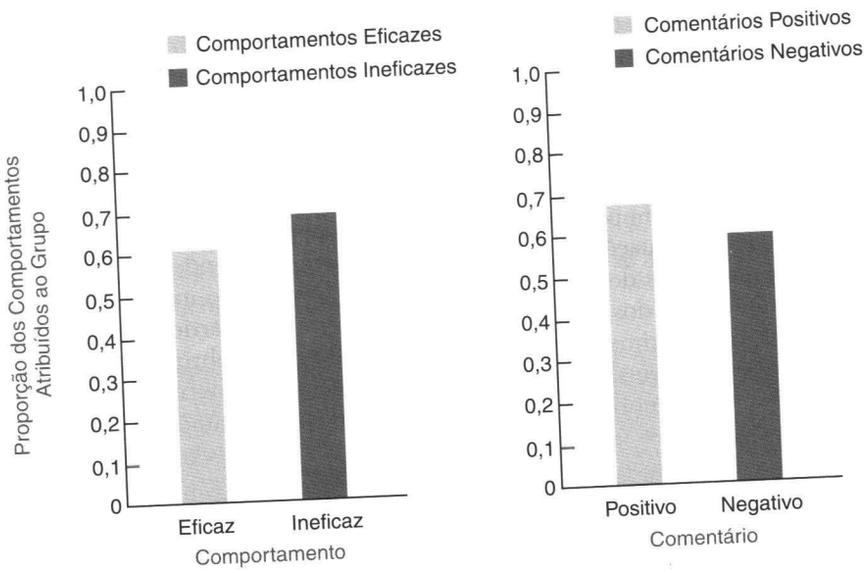
manipula de duas a quatro variáveis independentes simultaneamente. Esse procedimento possui diversas vantagens. Primeiro, com freqüência é mais eficiente conduzir um experimento, por exemplo, com três variáveis independentes do que conduzir três experimentos distintos. Segundo, o controle experimental muitas vezes é melhor, pois em um único experimento algumas variáveis de controle – hora do dia, temperatura, umidade e assim por diante – têm mais possibilidade de serem mantidas constantes do que em três experimentos distintos. Terceiro, e mais importante, os resultados generalizados – isto é, que mostraram ser válidos em diversas situações – em diversas variáveis independentes são mais valiosos do que os dados que ainda devem ser generalizados. Do mesmo modo que é importante estabelecer o caráter geral dos resultados por diversos tipos de participantes experimentais (ver Capítulo 12), os experimentadores também precisam descobrir se algum resultado é válido nos níveis das variáveis independentes. Quarto, isso nos permite estudar as interações, que são as relações entre variáveis independentes. Ilustramos essas vantagens por meio de alguns exemplos.

Vamos supor que desejamos saber, entre dois tipos de conseqüências, qual facilita o aprendizado de geometria por alunos do curso secundário. A primeira conseqüência é oferecer dinheiro diretamente em função de problemas de geometria resolvidos corretamente; a segunda consiste em dispensar mais cedo da aula, isto é, cada solução correta permite ao aluno sair da sala cinco minutos mais cedo. Suponha que os resultados desse experimento hipotético mostrem que dispensar mais cedo seja a melhor conseqüência. Antes de tornarmos a dispensa uma regra na escola secundária, deveríamos estabelecer inicialmente sua generalidade, comparando os dois tipos de conseqüências em outras disciplinas, como história ou biologia. Nesse caso, a matéria seria uma segunda variável independente. Seria melhor reunir essas duas variáveis em um único experimento do que conduzir dois experimentos sucessivos. Isso evitaria os problemas de controle, como uma classe ser testada na semana da final de um campeonato esportivo (caso em que nenhuma dessas conseqüências melhoraria a aprendizagem) e a outra ser testada na semana após o campeonato ter acabado (quando os alunos sentiam-se mais à vontade a respeito do aprendizado).

Quando os efeitos produzidos por uma variável independente são diferentes para cada nível de uma segunda variável independente, temos uma **interação**. A busca por interações constitui uma razão importante para o uso de mais de uma variável independente por experimento. Isso pode ser mais bem demonstrado por um exemplo.

Martell e Willis (1993) estavam interessados em descobrir como os comentários do desempenho de um grupo influenciariam aquilo que as pessoas lembrariam sobre o comportamento do grupo. Antes da observação do mesmo grupo de trabalho, as pessoas ouviram comentários positivos ou negativos sobre o desempenho dele. As pessoas que ouviram comentários positivos foram informadas que um painel de especialistas havia classificado o grupo na faixa dos 20% mais elevados de todos os grupos de trabalho. As pessoas que ouviram comentários negativos receberam a informação de que o grupo se classificava na faixa dos 20% mais baixos.

Havia duas variáveis independentes. As pessoas ouviam inicialmente os comentários positivos ou negativos. Todas as pessoas assistiram ao mesmo videoteipe de um grupo de trabalho. Os membros do grupo de trabalho exteriorizaram comportamentos tanto eficazes como ineficazes, em sua tentativa de construir uma ponte formada por tábuas e cordas, a fim de atravessar uma poça de água. Todas as pessoas assistiram ao mesmo videoteipe em que o grupo apresenta comportamentos eficazes e ineficazes. Portanto, as duas variáveis independentes eram o comentário positivo ou negativo sobre o desempenho e o tipo de comportamento observado (eficaz ou ineficaz). Depois de as pessoas assistirem ao videoteipe, uma



▶ FIGURA 3.1

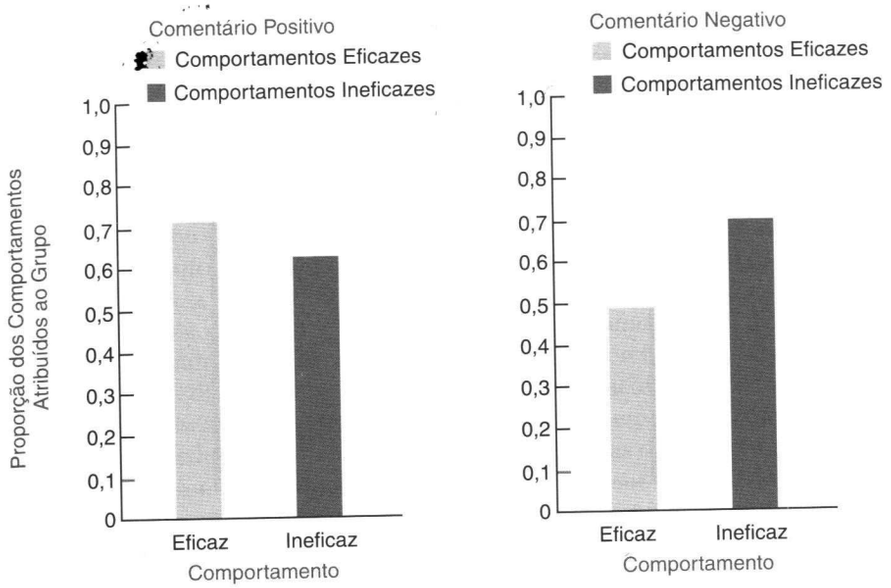
Efeito de duas variáveis independentes sobre a proporção de comportamentos atribuídos ao grupo. (Dados de Martell e Willis, 1993.)

lista de comportamentos foi mostrada e foi-lhes solicitado que informassem qual deles elas haviam visto. A variável dependente era, então, a proporção de comportamentos eficazes e ineficazes que as pessoas atribuíam ao grupo de trabalho.

Os resultados desse experimento estão indicados na Figura 3.1, mostrando cada variável independente. Quando as pessoas observaram um comportamento eficaz, foi verificada uma proporção menor dos comportamentos atribuídos ao grupo do que quando observaram comportamentos ineficazes. Quando as pessoas ouviam comentários positivos, constatou-se uma proporção maior de comportamentos atribuídos ao grupo.

A Figura 3.2 mostra, no entanto, que essa interpretação simples dos resultados pode ser enganosa. Comportamentos eficazes nem sempre produzem uma proporção menor de comportamentos atribuídos ao grupo, como se poderia concluir pela Figura 3.1. Como alternativa, o sentido do efeito do comportamento depende do tipo de comentário feito. Somente para comentários negativos os comportamentos eficazes produzem uma proporção menor de comportamentos atribuídos ao grupo. Isso não foi verdadeiro para os comentários positivos.

Lembre-se de que uma interação entre duas variáveis indica que os efeitos produzidos por uma variável independente (comentário sobre o desempenho) não são os mesmos para cada nível de uma segunda variável independente (comportamento no trabalho eficaz *versus* ineficaz). É exatamente isso que observamos na Figura 3.1. As pessoas que ouvem comentários positivos sobre o grupo de trabalho reconheceram uma maior proporção de comportamentos eficazes em relação aos ineficazes. Em contraste, as pessoas que ouviram comentários negativos responderam de modo contrário, reconhecendo uma maior proporção de comportamentos ineficazes em relação aos eficazes no grupo de trabalho. Esse julgamento representa uma ilusão de memória criada pela exposição a um comentário positivo ou negativo. Apesar do fato de as pessoas nas duas condições terem visto efetivamente o mesmo



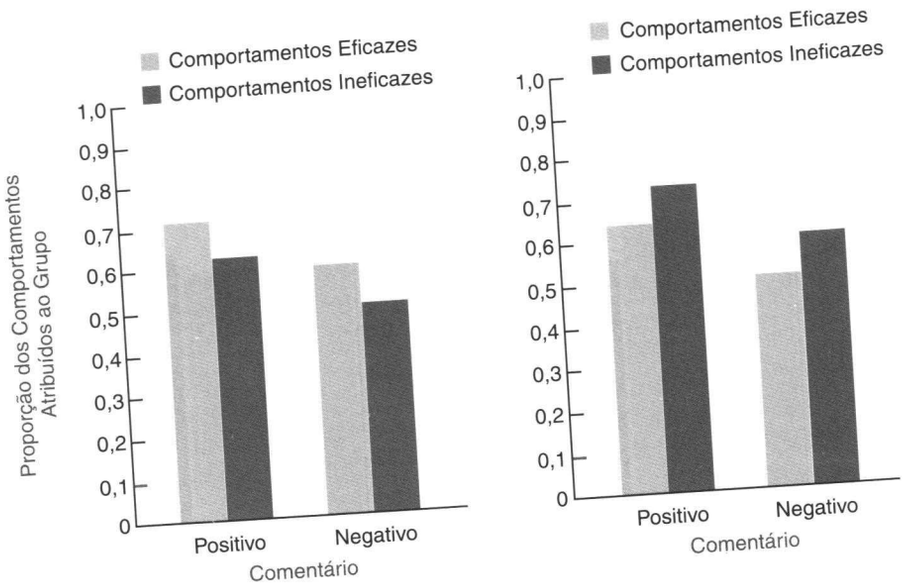
▶ FIGURA 3.2

Efeito do comentário sobre o desempenho na proporção dos comportamentos atribuídos ao grupo. Os efeitos dos comentários sobre o desempenho (positivos ou negativos) são mostrados separadamente para os comportamentos eficazes e ineficazes. (Dados de Martell e Willis, 1993.)

videoteipe, a memória que possuíam dos comportamentos do grupo estava distorcida pelas suas expectativas anteriores.

A Figura 3.2 mostra as duas condições que envolvem os comentários indicados no mesmo gráfico. Nesse caso, é mais fácil ver a natureza da interação; os efeitos produzidos pela manipulação do comentário sobre o desempenho dependem de os comentários serem eficazes ou ineficazes, com as pessoas reconhecendo uma maior proporção de comportamentos que são coerentes com o comentário (por exemplo, uma maior proporção de comportamentos eficazes associados a comentários positivos). Se os comentários sobre os desempenhos produziram resultados similares para comportamentos eficazes e ineficazes, os resultados deveriam ser parecidos com um dos gráficos de barra indicados na Figura 3.3, com ambos os grupos que ouvem comentários reconhecendo comportamentos mais eficazes (gráfico à esquerda) ou comportamentos mais ineficazes (gráfico à direita) do grupo de trabalho.

Vamos agora imaginar que esse experimento foi realizado com as duas condições de comentários, mas os experimentadores decidiram desprezar (ou deixaram de considerar) a distinção entre comportamentos eficazes e ineficazes no trabalho. Em outras palavras, os experimentadores investigaram o efeito do comentário sobre o desempenho em geral. A Figura 3.4 demonstra os perigos de não dar importância a variáveis importantes quando se planejam experimentos. A partir dos resultados mostrados nesse gráfico de barras, os experimentadores teriam sido levados a concluir que o comentário sobre o desempenho produz pouco ou nenhum efeito na proporção de comportamentos atribuídos ao grupo de trabalho. Você pode observar que, realizando o experimento desse modo, os experimentadores teriam, no final, resultados nulos e, como conseqüência, não teriam conseguido constatar um fato



▶ FIGURA 3.3

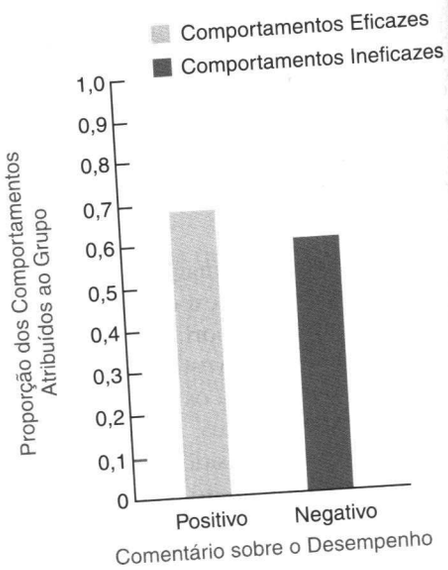
Dados hipotéticos indicando a ausência de interação do comentário sobre o desempenho e o tipo de comportamento no trabalho em relação à proporção de comportamentos atribuídos ao grupo.

importante a respeito dos efeitos dos comentários sobre o desempenho nas avaliações do comportamento no trabalho. (Os resultados na Figura 3.4 são simplesmente a média dos resultados dos dois tipos de comportamento mostrados na Figura 3.1 ou na Figura 3.2.)

A Figura 3.2, que contém os resultados efetivos, indica uma interação. Os efeitos de uma variável independente dependem do nível da outra variável independente. As pessoas

▶ FIGURA 3.4

Efeito do comentário sobre o desempenho no comportamento durante o trabalho em geral (isto é, com os dados agrupados dos comportamentos eficazes e ineficazes). (Dados de Martell e Willis, 1993.)



tinham maior possibilidade de endossar ou "lembrar-se" de comportamentos no trabalho que fossem coerentes com os comentários que ouviram a respeito do grupo de trabalho.

Muitos experimentos incluem duas ou mais variáveis independentes; isso significa que os resultados podem conter uma interação. Em virtude da freqüência com que se tem possibilidade de constatar interações, apresentamos um outro exemplo de um experimento com duas variáveis para ajudá-lo na interpretação dos resultados de experimentos complexos.

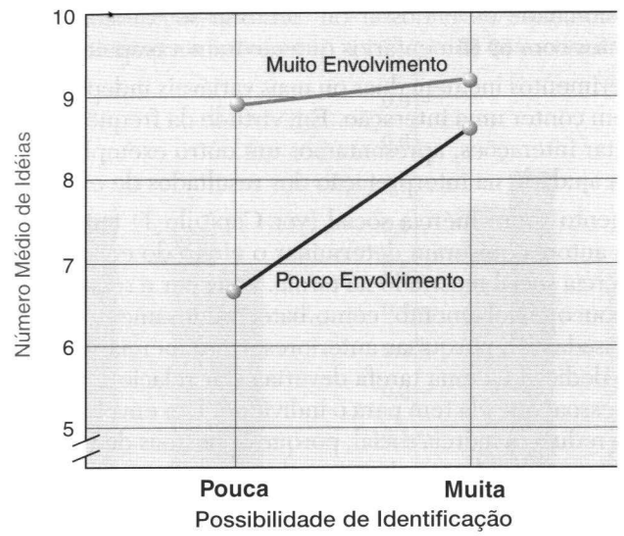
No experimento sobre inércia social (ver Capítulo 1) feito por Brickner, Harkins e Ostrom (1986), os autores quiseram determinar o efeito do envolvimento pessoal na tarefa sobre o grau de inércia social mostrado na tarefa. Brickner e seus colaboradores observaram que tarefas com pouco envolvimento, como bater ruidosamente e inventar usos para uma faca, haviam sido usadas em pesquisas anteriores sobre inércia social. Os autores raciocinaram que o esforço dedicado a uma tarefa deveria estar relacionado à importância intrínseca ou ao significado pessoal que ela tem para o indivíduo. Um envolvimento pessoal elevado em uma tarefa deveria reduzir a inércia social, porque as pessoas dedicariam muito esforço a tais tarefas, independentemente de seu desempenho pessoal ser controlado. Portanto, os pesquisadores variaram o envolvimento dos participantes na tarefa e também variaram o grau do esforço individual que poderia ser avaliado. Se o raciocínio dos pesquisadores estivesse correto, deveria ocorrer uma interação: o pouco envolvimento deveria resultar em inércia social (esforço reduzido quando o esforço individual não pode ser avaliado), porém um grande envolvimento deveria resultar no mesmo grau aproximado de esforço, se o indivíduo pudesse ou não ser identificado.

Brickner e seus colaboradores solicitaram que alunos de curso superior gerassem o maior número de idéias que pudessem, durante um período de 12 minutos, a respeito de uma proposta para implementar exames avançados abrangentes a que um aluno teria de se submeter, a fim de concluir o curso. Na condição de muito envolvimento, os alunos foram levados a acreditar que a proposta seria instituída na faculdade que cursavam antes de eles se formarem.

Portanto, a inclusão de exames abrangentes como um pré-requisito para a graduação deveria ter grande relevância pessoal. Na condição de pouco envolvimento pessoal, os alunos foram levados a acreditar que os exames seriam instituídos posteriormente, em uma outra faculdade. A condição da possibilidade de identificação do esforço individual também foi manipulada por instruções. Os participantes escreveram cada uma de suas idéias sobre exames abrangentes em uma folha de papel. Na condição de pouca possibilidade de identificação, as pessoas foram informadas de que suas idéias seriam agrupadas com as de outras, porque o comitê de avaliação das idéias desejava analisar a variedade de opiniões para o grupo como um todo. Na condição de muita identificação, as pessoas foram informadas de que suas opiniões seriam analisadas separadamente das demais, porque o comitê encarregado desejava avaliar as respostas individuais.

Resumindo, a variável dependente era o número de idéias geradas nas quatro condições: pouca possibilidade de identificação e pouco envolvimento; pouca possibilidade de identificação e grande envolvimento; muita possibilidade de identificação e pouco envolvimento; muita possibilidade de identificação e grande envolvimento.

Os resultados são mostrados na Figura 3.5, que indica o número de idéias geradas e a possibilidade de identificação para as duas condições de envolvimento. Pesquisas anteriores sobre inércia social são reproduzidas na condição de pouco envolvimento: foram geradas menos idéias quando as pessoas acreditavam que seu desempenho individual não estava sendo avaliado. Examine agora os resultados quando houve um grande envolvimento: o número de idéias era aproximadamente o mesmo, independentemente da possibilidade de



► **FIGURA 3.5**

Resultados do experimento realizado por Brickner, Harkins e Ostrom (1986), mostrando uma interação entre variáveis. Inércia social (baixo número de idéias que acompanham a baixa possibilidade de identificação), em contraste à alta, que ocorre em tarefas de baixo, mas não de alto envolvimento.

identificação. Portanto, as variáveis interagem: os efeitos da possibilidade de identificação dependem do nível de envolvimento na tarefa. Em outras palavras, a inércia social e, portanto, a difusão de responsabilidade têm menos possibilidade de ocorrer quando um indivíduo se confronta com uma tarefa que o envolve pessoalmente do que quando a tarefa não possui muito interesse intrínseco.

Em resumo, ocorre uma interação quando os níveis de uma variável independente forem diferentemente afetados pelos níveis de outras variáveis independentes. Quando há interações, não faz sentido discutir separadamente os efeitos de cada variável independente. Tendo em vista que os efeitos de uma variável também dependem dos níveis das outras variáveis, somos forçados a discutir as variáveis interagindo juntas.

Mais de uma Variável Dependente

A variável dependente (observada) é empregada como um índice de comportamento. Ela indica como é o desempenho da pessoa, permitindo ao experimentador avaliar o comportamento. O experimentador precisa decidir que aspectos do comportamento são relevantes para o experimento em foco. Embora algumas variáveis tenham sido usadas tradicionalmente, isso não significa que elas sejam os únicos índices de comportamento ou mesmo os melhores. Considere, por exemplo, o comportamento de um rato que pressiona uma barra ou de um pombo que bica um disco, respostas que são usadas em estudos de aprendizagem animal. A variável dependente mais comum é o número de pressões ou bicadas registradas. Porém, a força com que um disco é bicado também pode levar a descobertas interessantes

(ver Notterman e Mintz, 1965; Neuringer, 2002, p. 680), do mesmo modo que a latência (o tempo decorrido para responder). Os pesquisadores geralmente podem utilizar diversas variáveis dependentes que possam ser apropriadas. Digamos que desejamos estudar a legibilidade dos tipos gráficos que você está lendo neste momento. Evidentemente, não podemos observar a "legibilidade". Que variáveis dependentes poderíamos observar? Eis algumas que foram usadas no passado: retenção de informações significativas após a leitura de texto, tempo necessário para ler um número fixo de palavras, número de erros ao reconhecer letras individuais, velocidade na transcrição ou na redigitação do texto, batimentos cardíacos durante a leitura e tensão muscular durante a leitura, a lista é muito mais extensa.

Razões de economia justificam o maior número de medições simultâneas da variável dependente que seja viável. Apesar disso, o experimento típico usa somente uma ou, no máximo, duas variáveis dependentes simultaneamente. Isso é lamentável: do mesmo modo que a generalidade de um experimento é ampliada empregando mais de uma variável independente, ela também se amplia com diversas variáveis dependentes. A razão pela qual não são usadas mais variáveis dependentes é provavelmente por ser difícil analisar estatisticamente, ao mesmo tempo, diversas variáveis dependentes. Embora as técnicas modernas de computação tornem os cálculos muito viáveis, muitos psicólogos experimentais não foram bem treinados nesses procedimentos estatísticos multivariados e, portanto, hesitam em utilizá-los. Poderiam ser feitas análises separadas para cada variável dependente *per si*, porém isso acarreta a perda de informações de modo quase idêntico que uma análise separada das variáveis independentes não leva em conta as interações. A análise multivariada é complexa; no entanto, você deve ficar ciente de que, muitas vezes, é vantajoso usar mais de uma variável dependente em um experimento.

► DELINEAMENTOS EXPERIMENTAIS

O propósito do delineamento experimental consiste em minimizar os efeitos de variáveis externas ou incontroladas, aumentando, desse modo, a possibilidade de um experimento produzir resultados válidos e coerentes. Livros inteiros foram escritos sobre delineamentos experimentais. Nesta seção, cobrimos uma amostra de algumas das técnicas mais comuns usadas para melhorar o delineamento dos experimentos.

Uma das primeiras decisões relativas a delineamento que o experimentador precisa tomar consiste em como designar participantes para os vários níveis das variáveis independentes. As duas principais possibilidades são designar apenas alguns sujeitos para cada nível ou designar cada sujeito para todos os níveis. A primeira possibilidade denomina-se **delineamento de grupo** e a segunda, **delineamento de sujeito único**. A diferença pode ser mostrada por meio de um exemplo simples. Trinta alunos do curso de Introdução à Psicologia se apresentaram para participar de um experimento que você está conduzindo, a fim de testar a habilidade para se lembrar de palavras absurdas. Sua variável independente é o número de vezes que você dirá cada palavra: uma ou cinco vezes. Você espera que uma palavra apresentada cinco vezes seja mais bem gravada do que uma apresentada somente uma vez. O delineamento de grupo requer que você divida o grupo completo por dois, isto é, dois grupos de 15 alunos cada, com um grupo ouvindo cada palavra cinco vezes e o outro, uma vez. (Discutiremos brevemente como distribuir os sujeitos em cada grupo.) O delineamento de sujeito único engloba todas as 30 pessoas aprendendo com ambos os níveis da variável independente, isto é, cada um é testado com uma citação e também repetindo a palavra cinco vezes. (Como determinar a ordem pela

qual cada sujeito expõe-se a essas duas situações também é discutido posteriormente.) Que delineamento você deveria usar?

Delineamentos de Grupo

O delineamento de grupo é conservador. Não há possibilidade de que um tratamento continue a interferir no outro, porque cada pessoa recebe apenas um tratamento (uma citação ou cinco repetições, mas não ambas). Uma desvantagem, contudo, é que o delineamento de grupo precisa lidar com diferenças entre pessoas e isso diminui sua eficiência, isto é, sua habilidade para detectar diferenças reais entre uma e cinco repetições das palavras a serem memorizadas.

Em todo delineamento de grupo, o experimentador precisa tentar minimizar as diferenças entre os indivíduos em dois ou mais grupos de tratamento. De modo claro, se escolhêssemos as cinco pessoas com melhor memória e as colocássemos deliberadamente no grupo de uma citação e os cinco piores no grupo de cinco repetições, poderíamos, ao final, não obter resultados diferentes, mesmo que o grupo de uma citação tenha um desempenho melhor. A fim de evitar esse resultado, o experimentador precisa assegurar que ambos os grupos sejam equivalentes no início do experimento.

Grupos Equivalentes Um modo para assegurar a equivalência seria aplicar um teste de memória para os 30 indivíduos antes de o experimento ter início, a fim de obter uma medida **de linha de base** relativa à habilidade das pessoas para memorizar palavras absurdas. Os resultados da linha de base relativos aos sujeitos poderiam ser usados, então, para formar pares de sujeitos que possuíssem resultados iguais ou muito similares. Um membro de cada par seria designado aleatoriamente a um grupo, e outro membro, ao segundo grupo. Essa técnica é denominada **pareamento por igualação**. Uma dificuldade do pareamento por igualação é que o experimentador não pode combinar sujeitos em termos de toda característica possível. Portanto, sempre existe a possibilidade de que os grupos, muito embora pareados por igualação de acordo com alguma(s) característica(s), difiram em alguma outra característica que pode ser relevante (o pareamento por igualação é discutido em mais detalhes posteriormente neste capítulo).

Uma técnica mais comum usada para assegurar que grupos equivalentes sejam formados é a **randomização**. Randomização significa que cada pessoa participante de um experimento possui igual probabilidade de ser colocada em um grupo específico. Em nosso experimento de memorização, um modo para formar dois grupos por meio da randomização consiste em sortear nomes de um saco. Ou poderíamos solicitar a cada pessoa para que viesse e, então, jogasse um dado. Um resultado par corresponderia a um grupo e um resultado ímpar, a outro. Se não possuíssemos dados, uma tabela de números aleatórios poderia ser empregada para gerar dígitos pares e ímpares. Esse método de designar indivíduos de acordo com as condições experimentais não possui viés, pois não leva em consideração todas as características das pessoas; temos a expectativa de que os grupos formados dessa maneira sejam equivalentes sob todos os aspectos relevantes. A randomização, no entanto, não garante que os grupos serão sempre iguais. Um número maior de indivíduos com melhor memória poderia vir a fazer parte de um dos grupos por acaso. A possibilidade de isso ocorrer pode ser calculada pelos métodos da teoria da probabilidade, conforme aplicados à estatística (ver Apêndice B). Essa é uma razão pela qual os delineamentos experimentais e a estatística muitas vezes são tratados como o mesmo tópico. No entanto, o delineamento experimental preocupa-se com a lógica de planejar experimentos, ao passo que a estatística lida com o cálculo de possibilidades, probabilidades e outras quantidades matemáticas.

O pareamento por igualação é preferível à randomização, se todos os aspectos relevantes tiverem sido levados em consideração. Porém, como poucas vezes temos certeza, a randomização é usada com mais freqüência.

Delineamentos de Sujeito Único

Muitos psicólogos experimentais podem preferir o delineamento de sujeito único no qual todos os 30 indivíduos são testados com uma citação e novamente com cinco repetições (ou vice-versa). Ele é mais eficiente, pois cada indivíduo é comparado consigo mesmo. Qualquer diferença resultante de uma citação *versus* cinco repetições não pode ser o resultado de diferenças entre as pessoas nos dois grupos, como poderia ser o caso para o delineamento de grupo.

Efeitos de Prática Geral Existe um risco, no entanto, mesmo no delineamento de sujeito único mais eficiente. Imagine que todos os 30 indivíduos gravem primeiro um grande número de palavras por meio de cinco repetições e, então, gravem com uma citação. Na ocasião em que os indivíduos iniciarem o tratamento de uma citação, eles podem ter se tornado mais hábeis no aprendizado das palavras absurdas ou podem estar sentindo algum tédio ou cansaço em relação à tarefa. Ambas as possibilidades são denominadas **efeitos de prática geral**. Assume-se geralmente que esses efeitos sejam os mesmos para todas as condições de tratamento, de modo a não importar se os indivíduos aprenderam com uma citação seguida por cinco repetições ou com cinco repetições seguidas por uma citação. Em virtude de os efeitos de prática geral serem os mesmos para todas as condições de tratamento, eles podem ser controlados, em grande parte, por meio de **contrabalanço**, que permite ao experimentador enfrentar a dificuldade de determinar a ordem pela qual os tratamentos devem ser apresentados aos sujeitos. Novamente, uma solução consiste em usar a randomização, sorteando a ordem dos tratamentos, usando uma tabela de números aleatórios ou utilizando um computador para dispor as condições aleatoriamente. A lógica por trás desse procedimento foi discutida anteriormente. No entanto, embora contrabalançar os tratamentos por meio da randomização produza ordens equivalentes a longo prazo, é menos provável que seja adequado quando houver somente um número reduzido de tratamentos. Na maioria dos experimentos, o número de participantes excede o número de tratamentos e, então, a randomização constitui uma boa técnica para designar sujeitos aos tratamentos.

O *contrabalanço completo* assegura que sejam usadas todas as ordens possíveis de tratamento. No experimento das citações, isso é fácil, pois há somente duas ordens: uma e cinco citações, cinco e uma citações. Metade do número de indivíduos ouviria uma citação seguida por cinco repetições e a outra metade teria a ordem inversa. À medida que aumenta o número de tratamentos, o número de ordens torna-se realmente grande. Três tratamentos possuem seis ordens diferentes; quatro tratamentos, 24 ordens diferentes; cinco tratamentos, 120 ordens diferentes, e assim por diante. À medida que os níveis de uma variável independente aumentam, o contrabalanço completo logo se torna impraticável.

O contrabalanço não elimina os efeitos da ordem. Ele permite aos experimentadores avaliar os seus possíveis efeitos. Caso tais efeitos estejam presentes e, especialmente, se tiverem interações com outras variáveis independentes, é necessário tomar providências para corrigir o delineamento. O experimentador poderia decidir pela repetição do experimento usando um delineamento de grupo, a fim de evitar efeitos na ordem. Alternativamente,

o experimento original poderia ser reanalisado como um delineamento de grupo, examinando o comportamento precisamente na condição inicial pela qual cada indivíduo passou.

Efeitos de Histórias Diferenciais Os efeitos de histórias diferenciais representam um problema mais sério do que os efeitos de prática geral. No caso dos **efeitos de histórias diferenciais**, o efeito da parte inicial do experimento na parte posterior do experimento varia, dependendo de qual tratamento ocorre primeiro. Imagine que todas as 30 pessoas memorizem inicialmente as palavras com cinco repetições, e então memorizem com uma citação. Como resultado de seu experimento anterior com cinco repetições, elas poderiam decidir repetir para si mesmas, outras quatro vezes, a palavra que foi apresentada somente uma vez. Isso eliminaria quaisquer diferenças entre os dois níveis da variável independente. Esse é um exemplo do efeito de história diferencial, tendo em vista que o efeito do primeiro tratamento sobre o segundo difere, dependendo de qual tratamento ocorreu primeiro. Esse não foi o caso dos efeitos de prática geral nos quais os sujeitos consideraram o segundo tratamento do mesmo modo (isto é, com maior aptidão, tédio ou cansaço), independentemente do tratamento que receberam na primeira fase do experimento. Os efeitos de histórias diferenciais podem ser muito reduzidos por meio do contrabalanço, porém o contrabalanço não pode eliminar totalmente esses efeitos. Caso exista razão para esperar os efeitos de histórias diferenciais, temos duas opções, além do contrabalanço: usar o delineamento de grupo ou acrescentar um intervalo de tempo suficiente entre os dois tratamentos. Em virtude de o delineamento de grupo ser menos eficiente, ele exigirá que um número muito maior de pessoas seja testado; porém, isso é preferível a conduzir um experimento seriamente imperfeito. Se decidimos acrescentar um intervalo de tempo entre os dois tratamentos, precisamos identificar uma duração de tempo suficiente para eliminar a possibilidade dos efeitos de histórias diferenciais.

Delineamentos de Baixo *n*

Antes de passarmos a discutir os delineamentos mistos, gostaríamos de mencionar uma variante do delineamento de sujeito como seu próprio controle tradicional – o **delineamento de baixo *n***. Esses delineamentos apresentam os níveis de variáveis independentes ou de tratamentos a um número reduzido de indivíduos ou a um único indivíduo. Como poucas pessoas são testadas, um número substancial de observações é registrado para cada uma delas, resultando em um experimento muito econômico e com alto grau de controle. Os experimentos com delineamentos de baixo *n* são comuns em pesquisas psicofísicas, clínicas e de condicionamento operante. De modo análogo ao delineamento de sujeito único, o experimentador precisa ser cuidadoso para contrabalançar tratamentos e prever todos os problemas associados à aplicação de tratamentos múltiplos para cada pessoa. Os delineamentos de baixo *n* são discutidos detalhadamente no Capítulo 9 deste livro e no Capítulo 9 da obra de Elmes, Kantowitz e Roediger (2003).

Delineamentos Mistos

Os experimentadores não precisam especializar-se exclusivamente no delineamento de grupo ou no de sujeito único. Muitas vezes, é conveniente e aconselhável tentar algumas variáveis independentes como de grupo e outras como de sujeito único no mesmo experimento (supondo, evidentemente, que o experimento tenha mais de uma variável independente). Se uma variável – por exemplo, a administração de um fármaco – parece ter possibilidade de afetar outras, essa pode ser feita em uma variável de grupo, enquanto as restantes

são manipuladas entre os mesmos sujeitos. Quando as tentativas de realizar uma tarefa ou a prática repetida dela for de interesse, ela é necessariamente uma variável que deve ser lidada por delineamentos de sujeito único. Freqüentemente é usado um delineamento misto, no qual alguma variável é tratada entre sujeitos para constatar seu efeito em uma segunda variável tratada em cada sujeito individualmente. Esse foi o caso no estudo de Martell e Willis (1993), descrito anteriormente. Os experimentadores designaram os comentários sobre o desempenho (positivo *versus* negativo) como uma variável tratada entre sujeitos, com base na expectativa de que os pesquisados demonstrariam relutância em acreditar no comentário negativo após ouvir um comentário positivo (ou vice-versa) a respeito do grupo. Em contraste, parecia razoável que cada pessoa observasse o grupo que tivesse comportamentos eficazes e ineficazes – uma variável intra-sujeitos. Esse tipo de delineamento conciliatório **(delineamento misto)** não é tão eficiente ou econômico como um delineamento de sujeito único puro, porém muitas vezes é mais seguro.

Condições de Controle

As variáveis independentes precisam ser variadas (ou manipuladas) pelo experimentador. Isso implica que toda variável independente precisa variar em quantidade (variação quantitativa) ou em espécie (variação qualitativa) no contexto do experimento. Por exemplo, se a conseqüência dada a um rato for uma variável independente, a quantidade escolhida pelo experimentador poderia ser uma ou quatro bolinhas de alimento. Alternativamente, poderíamos oferecer diferentes tipos de conseqüência, tais como alimento e bebida. O termo técnico para um único tratamento ou condição de uma variável independente é **nível**. Diríamos que os níveis da variável independente são um e quatro bolinhas de alimento, no primeiro exemplo, e alimento e água, no segundo exemplo.

Muitos experimentos incluem, além das variáveis independentes, algum **grupo de controle** (delineamento de grupo) ou alguma **condição de controle** (delineamento de sujeito único). O grupo de controle (ou, simplesmente, grupo controle), em sua forma mais simples, não possui os níveis de interesse da variável independente. No exemplo que acabamos de descrever, um grupo controle de ratos não receberia conseqüência. Ou digamos que um experimentador esteja interessado no efeito do ruído sobre o estudo. Usando um delineamento de grupo, o experimentador exporia um grupo de sujeitos a um ruído elevado por meia hora enquanto estivesse estudando; esse é o nível de interesse da variável independente. Um grupo controle estudaria o mesmo material durante meia hora em um ambiente silencioso (um nível muito baixo de ruído). Os dois grupos, então, seriam testados sobre o material. Qualquer diferença obtida no teste entre os dois grupos seria atribuída ao efeito do ruído. A característica importante de uma condição de controle é o fato de ela proporcionar uma linha de base sobre a qual alguma variável de interesse possa ser comparada. Um exemplo freqüente ocorre em pesquisas de memória, nas quais se solicita a um grupo de indivíduos que memorize duas listas diferentes de palavras. O experimentador está interessado no modo como memorizar uma lista interfere na memorização de outra. O grupo experimental (recebendo o nível de interesse da variável independente) memoriza primeiro a lista A, posteriormente a B e, então, é novamente testado sobre a lista A. O experimentador gostaria de mostrar que a memorização da lista B interfere na da lista A. Antes de se chegar a uma conclusão desse tipo, exige-se uma condição de controle para fins de comparação. É meramente insuficiente comparar o teste final da lista A com o primeiro teste, porque os indivíduos poderiam ter um desempenho pior no último teste da lista A simplesmente por estarem cansados ou poderiam ter desempenho melhor, porque obtiveram uma prática adicional. Uma condição de controle sem tratamento teria um grupo

Grupo Experimental	Memorizar a Lista A	Memorizar a Lista B	Testar a Lista A
Grupo Controle	Memorizar a Lista A	Fazer Contas	Testar a Lista A

▶ FIGURA 3.6

Exemplos de grupos experimental e de controle para a memorização de listas.

controle memorizando a lista A que, então, ficaria parado durante o tempo que levasse para o grupo experimental memorizar a lista B e, então, seria testado novamente na lista A. Porém, essa seria uma condição de controle deficiente, porque os indivíduos poderiam praticar ou repetir subvocalmente a lista A enquanto estivessem sem fazer nada. Isso melhoraria o desempenho final que teriam na última lista A e faria com que parecesse incorretamente que, no grupo experimental, a lista B interferia na lista A mais do que realmente ocorria. Uma condição de linha de base apropriada ocuparia o grupo controle durante o tempo em que o grupo experimental estivesse memorizando a lista B; talvez o experimentador solicitasse que eles fizessem cálculos ou alguma outra "tarefa que leva tempo" e evitasse qualquer treino excedente (ver Figura 3.6).

Algumas vezes, a condição de controle está contida implicitamente no experimento. Recorde-se do experimento sobre a memória discutido anteriormente, no qual a variável independente era o número de citações de uma palavra: uma ou cinco. Nenhum experimentador se preocuparia em incluir um grupo ou condição de controle com nenhuma citação, pois nenhum aprendizado poderia ocorrer sob essa condição peculiar. A condição de controle é implícita, pois cinco citações podem ser comparadas a uma, e vice-versa. Em virtude de o experimentador poder estar interessado tanto nos efeitos de uma única citação como nos de cinco repetições, provavelmente não designaríamos explicitamente o nível de uma citação uma condição de controle. Porém, ela proporciona uma linha de base para comparação – e então, por esse motivo, isso ocorra com a condição de cinco repetições, pois os resultados de uma citação podem ser-lhe comparados.

Diversos tipos de experimentos exigem mais de uma linha de base. Por exemplo, nas pesquisas fisiológicas e de medicamentos, é necessário um controle do trauma provocado por uma cirurgia ou injeção. Portanto, uma pessoa pode submeter-se a uma operação fictícia ou à ingestão de uma substância neutra (um placebo) na condição de controle; elas também seriam comparadas a outras pessoas de controle que não estiveram sujeitas a uma operação ou injeção.

Armadilhas

Infelizmente é muito fácil delinear inadequadamente um experimento; a maioria dos psicólogos experimentais relegou erros desse tipo a um arquivo morto. Nesta seção, discutimos somente uma pequena amostra de erros de delineamento, os que são tão comuns que você precisa conhecê-los.

Características de Demanda Os experimentos de laboratório tentam entender o comportamento do modo como é realmente influenciado pela variável independente. Algumas vezes, o próprio ambiente do laboratório ou o conhecimento de que um

experimento está se realizando alteram os padrões de comportamento. Muitas vezes, os participantes em pesquisa criam espontaneamente hipóteses ou suposições a respeito da finalidade do experimentador ao conduzir o experimento ou, então, comportam-se ou reagem de um modo que satisfará essa "finalidade". Tente esta demonstração simples para convencer-se de que tais efeitos ocorrem. Diga a cinco de seus amigos que você está conduzindo um experimento em sua aula de psicologia e gostaria de obter a colaboração deles como pesquisados. Caso eles concordem, solicite-lhes para segurar três ou quatro cubos de gelo nas mãos. Observe quantos seguram os cubos de gelo até derreterem. Peça a cinco outros amigos para que segurem os cubos de gelo sem mencionar nada a respeito do experimento. Em vez de segurar os cubos de gelo até que derretam, eles considerarão seu pedido um tanto estranho e logo o informarão a respeito disso. Há algo incomum a respeito da pronta aceitação dos amigos que sabiam estar participando de um experimento: um maior número deles estava disposto a segurar os cubos de gelo por um período de tempo mais longo. Os psicólogos denominam **características de demanda** as indicações dadas às pessoas que lhes permitem determinar a finalidade do experimento, ou o que é esperado pelo experimentador. Na medida em que o comportamento dos participantes na pesquisa é controlado pelas características de demanda, em vez de pelas variáveis independentes, os experimentos não são válidos e não podem ser generalizados além da situação de teste.

Um exemplo bem conhecido de uma característica de demanda é o **efeito Hawthorne**, assim chamado por causa da fábrica da Western Electric Company, onde foi observado inicialmente. A companhia estava interessada em melhorar o moral e a produtividade dos empregados e conduziu diversos experimentos (como melhorar a iluminação) para aperfeiçoar o ambiente dos empregados. A produtividade dos empregados aumentou independentemente da variação experimental tentada. Os empregados sabiam que faziam parte de um grupo "especial" e, portanto, sempre tentaram ter o melhor desempenho (ver Parsons, 1974, e Bramel e Friend, 1981, para interpretações alternativas desses resultados). As características de demanda eram mais importantes para determinar a produtividade dos empregados do que as manipulações experimentais. Embora a expressão *efeito Hawthorne* seja amplamente usada para descrever experimentos de campo nos quais a produtividade aumenta por causa da participação no estudo, ocorreram diversas revisões detalhadas do experimento em Hawthorne, sugerindo que a conclusão original baseou-se em provas frágeis (Wickström e Bendix, 2000; Brannigan e Zwerman, 2001). Apesar disso, a expressão permanece amplamente utilizada.

As características de demanda e o efeito Hawthorne precisam ser avaliados cuidadosamente. Um estudo recente (Fostervold, Buckmann e Lie, 2001) incluía condições de controle especiais para avaliar os efeitos dos filtros VDU (*visual display unit*) nas telas dos computadores. Na primeira parte do estudo, um grupo de participantes possuía filtros (grupo com filtros) e um outro grupo de controle, não. A comparação dos resultados para os dois grupos mostrou vários benefícios para o grupo com filtros. No entanto, os pesquisadores também incluíram uma segunda fase, na qual o grupo controle recebia um filtro, enquanto o outro grupo continuava com o mesmo filtro. Só foram observadas pequenas alterações para o grupo de controle inicial. Além disso, os benefícios iniciais para o grupo com filtro diminuíram durante a segunda fase. Portanto, os resultados na primeira fase eram causados pelas características de demanda e não por benefícios associados aos filtros VDU. Caso os experimentadores tivessem conduzido somente a primeira fase de seu estudo, um benefício falso dos filtros, na realidade causado pelas características de demanda, poderia ter sido considerado de modo incorreto.

Efeitos do Experimentador Uma armadilha intimamente relacionada às características de demanda é o **efeito do experimentador**, que influencia o resultado acidentalmente ao fornecer aos participantes algumas indicações quanto às suas expectativas. Por exemplo, um experimentador poderia não estar ciente de que move a cabeça aprovando, quando uma resposta correta é dada, e fecha a expressão após os erros. O sexo, a raça e a origem étnica do pesquisador também são efeitos potenciais do experimentador. As características do experimentador têm maior probabilidade de distorcerem os resultados de um experimento de pesquisa que focalize temas relacionados a essas características – por exemplo, a raça de um experimentador que estiver conduzindo um experimento a respeito da cor da pele nos índices de desempenho no trabalho.

Esses efeitos não são limitados a experimentos com seres humanos. O efeito do experimentador também pode ocorrer em experimentos aparentemente objetivos com animais. Rosenthal e Fode (1963) disseram a alunos experimentadores que os ratos a serem testados por eles em um labirinto eram de linhagens especiais: ágeis ou vagarosos em labirintos. Na realidade, os ratos originaram-se da mesma linhagem e tinham a mesma procedência. Apesar disso, os ratos denominados ágeis em labirintos cometiam menos erros do que os denominados vagarosos, e essa diferença era estatisticamente atestada. Os alunos experimentadores foram observados enquanto testavam os ratos: eles não enganavam ou faziam algo premeditado que distorcesse os resultados. Parece razoável que os alunos com sorte a quem foram entregues ratos supostamente espertos estivessem mais motivados para realizar o experimento do que os infelizes que tinham de ensinar aos ratos estúpidos como percorrer o labirinto. Isso afetou de alguma forma os resultados do experimento, talvez porque os experimentadores lidaram de modo diferente com os dois grupos de ratos.

A melhor maneira para eliminar esse tipo de efeito do experimentador consiste em não revelar a ele a condição do experimento, baseando-se na premissa de que os experimentadores não podem comunicar aquilo que não conhecem. Esse procedimento é denominado **experimento duplo-cego**, porque nem o experimentador nem o participante da pesquisa sabem que pessoas encontram-se em cada condição de tratamento. Tal procedimento foi usado, por exemplo, em um estudo dos efeitos da poluição do ar sobre o comportamento. As pessoas respiravam ar puro ou o ar de uma estrada movimentada. O ar estava contido em tanques; o experimentador não sabia qual tanque continha ar puro e qual armazenava ar poluído. O pior desempenho das pessoas quando respiravam ar poluído não podia ser atribuído ao experimentador que, de modo inadvertido, informava a qualidade do ar às pessoas ou tratava-as diferentemente.

Os efeitos do experimentador nem sempre são assim tão sutis. Um dos autores uma vez esteve envolvido em um experimento a respeito do comportamento de piscar os olhos. Diversos experimentadores ajudaram a conduzir o mesmo experimento, e logo foi observado que um deles obtinha resultados muito diferentes daqueles do restante de nós. Seus pesquisados iniciaram as sessões experimentais exibindo um piscar de olhos rápido e incessante. A causa desse comportamento estranho foi facilmente descoberta. Para registrar o piscar de olhos, o experimentador precisava colocar uma minúscula haste de metal na pálpebra da pessoa usando uma fita especial – normalmente, um procedimento indolor. No entanto, o experimentador em questão tinha um polegar muito grosso e não era capaz de colocar a haste sem irritar o olho da pessoa, causando-lhe um estranho piscar incessante.

Quando um experimentador suspeita que algum aspecto de sua aparência ou de suas características (por exemplo, sexo, raça, grupo racial) possa alterar o padrão de comportamento dos participantes, então uma solução possível consiste em incorporar isso como uma variável independente ou uma variável de controle adicional no delineamento experimental.

Se um experimentador afro-americano estiver conduzindo uma pesquisa sobre cor da pele e índices de desempenho no trabalho, ele poderia solicitar a um colega ou assistente de pesquisa branco para testar metade do número de pessoas e, então, comparar os efeitos da cor da pele nas duas condições da raça do experimentador.

Automação dos Experimentos Os efeitos do experimentador podem ser eliminados ou muito reduzidos fazendo com que computadores ou outros equipamentos conduzam o experimento, de modo que a pessoa não seja tocada por mãos humanas. Em muitos laboratórios, o sujeito entra em uma cabine de teste e vê uma mensagem em uma tela que lhe diz para pressionar um botão, a fim de começar. Pressionar um botão faz com que as instruções do experimento apareçam na tela. Todo o experimento é, então, conduzido por um computador. O experimentador aparece ao final da coleta de dados para prestar esclarecimentos ao participante, informando-lhe os objetivos do estudo e explicando como ajudou o avanço da ciência. Até então, o experimentador simplesmente monitora o equipamento e a pessoa, a fim de assegurar que ela está seguindo as instruções e que nada desfavorável aconteça. Tal automação obviamente reduz os perigos da distorção pelo experimentador.

Quase-Experimentos

Seja qual for a razão, muitas variáveis não podem ser manipuladas diretamente. Um fator que impede a manipulação de variáveis em experimentos são as considerações éticas que todos os cientistas precisam ter (ver Capítulo 4). É ético pesquisar ou, então, observar o uso de drogas por alunos de faculdade, desde que seja obtida permissão. No entanto, por maior que seja a imaginação, não seria ético criar um grupo de dependentes de drogas e comparar suas atividades com um grupo não consumidor de drogas também formado por nós. Uma segunda barreira à manipulação é a Mãe Natureza. Algumas variáveis, como o sexo de nossos pesquisados, não podem ser alteradas pelo experimentador (exceto em circunstâncias muito raras e controvertidas); outras variáveis, como catástrofes naturais (tornados, furacões) ou não-naturais (guerras, queda de aviões), são difíceis de implementar física e moralmente. Podemos realizar experimentos que se relacionam a esses fenômenos? Afinal de contas, tais variáveis e outras como elas são cativantes e podem desempenhar uma parte importante na experiência humana.

A resposta à pergunta (supondo que você seja um cientista ético) é a seguinte: Você pode e não pode. Não estamos sendo tolos, mas enfatizando o fato de que você não pode realizar experimentos reais com fenômenos como os que acabamos de apresentar. Você pode, no entanto, conduzir **quase-experimentos**. A técnica, nesse caso, é similar à análise *ex post facto* na análise de correlação, exceto que dois ou mais níveis da variável de interesse são examinados, em vez de ser correlacionados. Esperamos que a Mãe Natureza faça o seu trabalho e, então, comparamos os efeitos dessa "variável independente" aos efeitos que ocorrem quando a variável não estiver presente ou, de algum modo, diferir. Se compararmos a habilidade de leitura dos homens com a das mulheres ou a velocidade das pessoas com conhecimento de leitura dinâmica com a dos adultos médios, teremos conduzido um quase experimento.

As vantagens dos quase-experimentos são óbvias: eles usam variáveis independentes que ocorrem naturalmente, a maioria das quais possui um grau elevado de interesse intrínseco e implicações práticas importantes. O quase-experimento típico possui uma **variável de sujeito** como variável independente. Se desejarmos obter informações sobre quase toda variável inerente de sujeito (idade, sexo, raça, grupo étnico), atributo do participante causado

socialmente (classe social, região de residência) ou atributo do participante relacionado à doença ou incapacidade física (perda de um membro, doença mental, dano cerebral, efeitos de acidentes), teremos de selecionar, em vez de manipular nossas variáveis independentes, a não ser que seja possível realizar experimentos diretamente em organismos infra-humanos. Embora os quase-experimentos sejam interessantes e possam contribuir com pesquisas muito importantes, devemos alertá-lo, neste ponto, que as vantagens dos quase-experimentos são obtidas sacrificando-se o controle. Quando o pesquisador tem de aceitar aquilo que se encontra disponível, o que está à disposição pode incluir diversas variáveis importantes que causam confusão.

Em virtude de muitas pesquisas em psicologia preocuparem-se com variáveis de sujeito e pelo fato de que quase-experimentos usando essas variáveis têm possibilidade de ser objeto de confusão, examinaremos a seguir os problemas e as soluções possíveis.

Um experimentador não pode manipular uma variável de sujeito e, ao mesmo tempo, manter outros fatores constantes; ele pode selecionar somente sujeitos que já possuam a característica em algum grau variado e, então, compará-los com base no comportamento de interesse. Se os sujeitos nos diferentes grupos (por exemplo, QI elevado, médio e reduzido) diferirem no comportamento, não podemos concluir que a diferença relativa à variável de sujeito tenha produzido ou seja responsável pela diferença de comportamento. A razão é que outros fatores podem variar simultaneamente e confundir-se com a variável de sujeito. Se as pessoas com QI elevado realizam alguma tarefa melhor do que pessoas com QI reduzido, não podemos afirmar que o QI produziu ou causou a diferença, porque os diferentes grupos de pessoas têm possibilidade de apresentar variação em termos de outros parâmetros relevantes, como motivação, educação etc. Quando são investigadas as variáveis de sujeito, não podemos atribuir com segurança, a essa variável, as diferenças de comportamento, do modo como podemos fazer com variáveis experimentais reais. Então, tais delineamentos produzem, essencialmente, correlações entre variáveis. Podemos afirmar que as variáveis estão relacionadas, mas não podemos dizer que aquela variável específica produz ou causa o efeito na outra variável.

Esse é um aspecto muito importante; vamos considerar um exemplo. Suponha que um investigador esteja interessado no funcionamento intelectual (ou na falta dele) das pessoas que sofrem de esquizofrenia. As pessoas diagnosticadas como pertencentes a esse grupo são submetidas a numerosos testes, cuja finalidade consiste em avaliar diversas habilidades mentais. O pesquisador também aplica esses testes a um outro grupo de pessoas, denominadas normais. Ele descobre que os esquizofrênicos têm desempenho pior em relação às pessoas normais nos testes que envolvem aspectos semânticos do idioma, como aqueles relativos ao significado das palavras ou à compreensão de trechos de prosa. O investigador conclui que os esquizofrênicos possuem um desempenho pior nesses testes porque são portadores dessa doença e que sua habilidade para empregar bem o idioma na comunicação é provavelmente uma causa que contribui para a esquizofrenia.

Estudos como esses são comuns em algumas áreas da psicologia. Apesar do fato de que conclusões similares a essa muitas vezes são obtidas desses estudos, elas são totalmente injustificadas. Ambas as conclusões baseiam-se em correlações, e outros fatores poderiam muito bem ser os essenciais. Esquizofrênicos podem ter desempenho pior do que as pessoas normais por diversas razões. Eles podem não ser tão inteligentes, motivados, educados ou preparados ao se submeter a testes. Pode ser simplesmente que eles estiveram internados em uma instituição durante um longo período, tendo como resultado uma interação social e intelectual reduzida. Portanto, não podemos concluir que a razão pela qual os dois grupos diferem nos testes verbais seja a esquizofrenia ou sua inexistência nos grupos. Mesmo se

pudéssemos chegar a essa conclusão, ela certamente não implicaria a outra condição, a de que os problemas com a língua contribuem para causar esquizofrenia. Nesse caso, teríamos novamente, uma correlação entre essas duas variáveis, sem ter idéia da possibilidade ou do modo como as duas estão relacionadas causalmente.

O uso de variáveis de sujeito é muito comum em todas as pesquisas de psicologia, porém é absolutamente importante em áreas como psicologia clínica e do desenvolvimento. Portanto, os problemas relacionados à obtenção de inferências a partir de tais pesquisas devem ser cuidadosamente considerados. Uma variável fundamental na psicologia do desenvolvimento é a idade, uma variável de sujeito; isso significa que muitas pesquisas nesse campo possuem natureza correlacional. Em geral, o problema das diferenças individuais entre participantes em psicologia muitas vezes é desprezado, embora freqüentemente existam apelos para considerá-lo como primordial (ver Underwood, 1975). Dedicaremos um capítulo posteriormente às diferenças individuais (Capítulo 12). Vamos considerar neste ponto uma maneira de tentar inferências mais robustas a partir dos experimentos que empregam variáveis de sujeito.

Pareando por Igualação Novamente O problema básico na investigação de variáveis de sujeito e em outras pesquisas *ex post facto* é o fato de que todas as diferenças observadas no comportamento podem ser causadas por variáveis indiscriminadas que as acompanham. Uma maneira para tentar evitar esse problema é pareando os participantes por igualação em termos das outras variáveis relevantes. Na comparação entre pessoas esquizofrênicas e normais, observamos que os dois grupos também tinham possibilidade de diferir em outras características, como QI, educação, motivação, permanência em uma instituição e talvez até mesmo a idade. Em vez de simplesmente comparar as pessoas esquizofrênicas com as normais, poderíamos tentar compará-las com outro grupo que combinasse mais de perto nesses três parâmetros, de modo que, esperamos, a principal diferença entre os grupos seria a presença ou ausência de esquizofrenia. Por exemplo, poderíamos usar um grupo de pacientes que, em média, são similares aos esquizofrênicos em termos de idade, QI, período de tempo de permanência em uma instituição, sexo e alguma medida da motivação. Quando os dois grupos tiverem sido pareados por igualação em todas essas características, então podemos atribuir mais confiantemente qualquer diferença no desempenho entre eles ao fator de interesse, ou seja, a esquizofrenia. Ao parear por igualação, os investigadores tentam introduzir a principal característica da experimentação – ser capaz de manter constantes os fatores externos ao principal foco da pesquisa, a fim de evitar indiscriminação – naquilo que é essencialmente uma observação correlacional. O objetivo consiste em permitir inferir que a variável de interesse (esquizofrenia) produza o efeito observado.

Diversos problemas relativamente sérios encontram-se associados ao pareamento por igualação. Um dos motivos é que isso muitas vezes requer um grande esforço, porque algumas das variáveis relevantes podem ser bem difíceis de medir. Mesmo quando um pesquisador se dá ao trabalho de tomar as medidas adicionais necessárias, ainda pode ser impossível parear os grupos por igualação, especialmente se poucas pessoas estão envolvidas antes de se tentar o pareamento por igualação. Mesmo quando o pareamento por igualação é bem-sucedido, ele muitas vezes reduz consideravelmente o tamanho da amostra na qual as observações são feitas. Possuímos, então, menos confiança em nossas observações, porque podem não ser estáveis e reproduzíveis.

O pareamento por igualação muitas vezes é difícil, porque diferenças importantes entre sujeitos podem provocar efeitos sutis. Além disso, os efeitos de uma diferença podem interagir com outra. Portanto, *interações sutis* entre variáveis pareadas por igualação

podem indiscriminar os resultados. Para ilustrar essas dificuldades, vamos considerar uma parte do trabalho realizado por Brazelton e seus colaboradores sobre comportamento neonatal (Lester e Brazelton, 1982).

O principal interesse de Brazelton reside nas diferenças culturais existentes no comportamento neonatal, conforme medidas pela Escala de Avaliação do Comportamento Neonatal de Brazelton. A estratégia geral consiste em comparar recém-nascidos de várias culturas e grupos étnicos com recém-nascidos dos Estados Unidos. Nesses quase-experimentos, a cultura ou o grupo étnico, que é uma variável de sujeito, é a variável quase-independente. As tentativas geralmente são feitas para parear por igualação os bebês de culturas diferentes em termos de diversos parâmetros, como peso ao nascimento, altura ao nascimento e risco obstétrico (incluindo a informação se a mãe recebeu medicação durante o parto, se o bebê era prematuro etc.). Lester e Brazelton demonstram que existe uma relação sinérgica entre esses fatores. A **sinergia** em um contexto médico significa que os efeitos combinados de duas ou mais variáveis não se somam: o efeito combinado é maior do que a soma dos componentes individuais. Isso significa que as variáveis interagem.

O modo pelo qual as características neonatais e o risco obstétrico interagem encontra-se descrito a seguir. Estudos têm mostrado que o comportamento (conforme avaliado pela escala de Brazelton) de crianças ligeiramente abaixo do peso é mais intensamente influenciado (de modo negativo) por pequenas doses de medicação tomadas pela mãe do que o comportamento de recém-nascidos que estão com o peso mais próximo da média. Muito embora os recém-nascidos sejam selecionados cuidadosamente, efeitos sutis e interagentes das variáveis pareadas por igualação podem influenciar os resultados. Esse é um problema especialmente difícil no trabalho de Brazelton, porque grande parte de suas pesquisas analisou recém-nascidos de culturas pobres, nas quais o peso dos bebês que nascem é reduzido e o risco obstétrico, muito elevado. Em termos gerais, você deve lembrar que variáveis pareadas por igualação raramente encontram-se sob controle direto, o que significa a possibilidade de efeitos inter-relacionados de variáveis indiscriminadas estarem sempre presentes.

Um outro problema com o pareamento por igualação envolve a introdução do temível **dispositivo de regressão**. Isso é discutido no Capítulo 12; contudo, o explicamos brevemente nesta seção. Sob certas condições, em muitos tipos de medições ocorre um fenômeno estatístico conhecido como **regressão para a média**. A média de um grupo de resultados é o que a maioria das pessoas considera como a média aritmética: o total de todas as observações dividido pelo número de observações. Por exemplo, a altura média em uma amostra de 60 pessoas é a soma de todas as suas alturas dividida por 60. Normalmente, se as pessoas que tiveram medições nos extremos (isto é, pessoas muito altas ou muito baixas) em alguma característica são retestadas, a segunda medição será mais próxima da média de todo o grupo do que suas medições originais. Considere um exemplo. Submetemos 200 pessoas a um teste padrão de raciocínio matemático, para o qual existem duas formas equivalentes ou duas versões do teste que sabemos ser equivalentes. O resultado médio (a média aritmética) do teste é 60 entre 100 possíveis pontos. Consideramos as 15 pessoas com a maior pontuação e as 15 com a menor. A média desses grupos é, por exemplo, 95 e 30, respectivamente. Em seguida, testamos-as novamente, dessa vez usando a outra versão do teste. Poderíamos agora constatar que as médias dos dois grupos são 87 e 35. No segundo teste, os resultados desses grupos extremos tendem para a média; o grupo com medições maiores tem desempenho pior, e o grupo com medições menores tem desempenho um pouco melhor. Isso acontece basicamente para o grupo com medições maiores, porque algumas pessoas cujos "resultados verdadeiros" são um pouco menores do que os realmente testados tiveram sorte e obtiveram um resultado maior do que deveriam obter no teste. Quando testadas novamente, as pessoas com medição

muito elevada tendem a obter resultados menores, próximos de seu resultado verdadeiro. A situação se inverte para o grupo com resultados menores, isto é, algumas das pessoas tiveram um resultado abaixo de seus "resultados verdadeiros" no primeiro teste; testar novamente faz com que elas obtenham um resultado maior ou próximo do resultado verdadeiro.

Essa regressão para a média sempre é observada nas condições em que existe uma correlação menos do que perfeita entre duas medições. Quanto mais extrema a seleção dos resultados, maior a regressão para a média. Isso também ocorre em todos os tipos de situações envolvendo medições. Se pais anormalmente altos ou baixos têm um filho, provavelmente ele estará mais próximo da altura média da população do que a dos pais. Conforme ocorre com a maioria dos fenômenos estatísticos, a regressão para a média é válida para grupos de observações, sendo probabilística (isto é, pode ocorrer sempre). Por exemplo, alguns indivíduos podem se afastar da média no segundo teste de raciocínio matemático, porém a tendência do grupo será em direção à média.

De que modo a regressão para a média afeta os quase-experimentos nos quais as pessoas são iguais em termos de alguma variável? Considere novamente um exemplo. Esse, como grande parte das pesquisas *ex post facto* realizadas em relação a problemas aplicados, possui implicações importantes. Vamos supor que temos um programa educacional que, acreditamos, será especialmente vantajoso para elevar a capacidade de leitura das crianças afro-americanas. Isso é especialmente importante, porque a capacidade de leitura das crianças afro-americanas normalmente é menor que a das crianças brancas, presumivelmente por causa dos ambientes culturais diferentes. Escolhemos dois grupos de crianças, um de afro-americanas e outro de brancas, e os passaremos por igualação em diversos critérios, incluindo sexo, idade e, o mais importante, desempenho inicial na leitura. Proporcionamos a ambos os grupos de crianças um programa de melhoria de leitura e, então, testamos os resultados da leitura após o programa. Descobrimos, surpreendentemente, que as crianças afro-americanas na realidade têm um desempenho pior após o programa de leitura do que antes dele e que as crianças brancas têm um desempenho melhor. Concluímos, evidentemente, que o programa ajudou as crianças brancas, mas prejudicou efetivamente as crianças afro-americanas, apesar do fato de ter sido criado especialmente para estas últimas.

Essa conclusão, muito embora possa parecer razoável para você, é quase certamente errônea nesse caso, por causa dos dispositivos de regressão. Considere o que aconteceu quando as crianças afro-americanas e brancas estavam em pé de igualdade relativamente aos resultados iniciais da leitura. Como no princípio as populações diferiam, com os afro-americanos alcançando menor pontuação que os brancos, tornou-se necessário, a fim de parear por igualação as duas amostras, selecionar os alunos afro-americanos com uma pontuação maior que a média de seu grupo e os alunos brancos tendo uma pontuação menor que a média de seu grupo. Tendo escolhido esses grupos extremos, preveríamos (por causa da regressão para a média) que, quando testadas novamente, em média as crianças afro-americanas teriam obtido resultados piores e as crianças brancas, resultados melhores, mesmo se o programa de melhoria da leitura não tivesse exercido nenhum efeito! As crianças afro-americanas com desempenho excepcionalmente elevado tenderiam a se aproximar da média de seu grupo, e as brancas com desempenho ruim tenderiam a aproximar-se da média de seu grupo. O mesmo teria acontecido se não houvesse um programa e as crianças simplesmente tivessem sido testadas novamente.

O mesmo resultado provavelmente teria sido obtido se as crianças tivessem sido pareadas por igualação em termos de QI, em vez dos resultados da leitura, pois os dois provavelmente possuem correlação positiva. Portanto, simplesmente identificar uma outra variável que resulte em pareamento de igualação pode não ser uma solução. Uma solução

seria combinar amostras muito grandes de crianças afro-americanas e brancas e, então, dividir cada grupo, oferecendo o programa de leitura a um subgrupo, e não a outro. Todas as crianças seriam testadas novamente no final da participação de um subgrupo no programa. (Evidentemente, a designação das crianças afro-americanas e brancas aos subgrupos seria aleatória.) A regressão para a média deveria ocorrer em ambos os subgrupos, porém o efeito do programa de leitura poderia ser avaliado em relação ao grupo para o qual foi oferecido o programa. Talvez as crianças afro-americanas que tiveram o programa de leitura teriam demonstrado uma diminuição muito menor (regressão para a média) do que as que não o tiveram, indicando que o programa realmente exerceu um efeito positivo.

Em virtude de a pesquisa envolvendo quase-experimentos com variáveis de sujeito ser muitas vezes conduzida para avaliar os programas educacionais, seus praticantes precisam estar conscientes dos muitos problemas espinhosos associados a seu uso. Pode não se ter condições de fazer muitas afirmações a respeito dos resultados ou de tirar conclusões importantes por causa dos fatores não-controlados que podem afetar os resultados. Parear por igualação ajuda a diminuir esse problema em alguns casos em que seu uso é possível, mas, então, introduz-se a possibilidade dos dispositivos de regressão. E muitos pesquisadores parecem não ter conhecimento desse problema. Um erro grosseiro famoso em tal pesquisa de avaliação, muito similar ao estudo hipotético delineado nesta seção, é discutido no Capítulo 12.

Quando o pareamento por igualação é uma possibilidade prática e quando os dispositivos de regressão são avaliados, podemos nos sentir um tanto mais confiantes com as conclusões de nossos resultados. Porém, deveríamos nos lembrar de que aquilo que possuímos ainda é apenas uma correlação, embora muito controlada. O pareamento por igualação algumas vezes é útil, mas não é um remédio para todos os males. Em nosso exemplo anterior, comparando pessoas esquizofrênicas com outras pessoas nos testes de desempenho mental, se os esquizofrênicos ainda tivessem um desempenho pior do que o novo grupo controle pareado por igualação, poderíamos concluir, então, que a esquizofrenia resultava em um uso inferior do idioma? Não poderíamos. Ainda poderia ser algo diferente, alguma outra diferença entre os dois grupos. Nunca podemos estar absolutamente seguros de que pareamos por igualação as variáveis relevantes.

O estudo de delineamento experimental é complexo. Na maioria dos capítulos, incluímos uma seção denominada Do Problema à Experiência, que mostra como transformar algum tema ou questão em um experimento real. Descreveremos essa situação em seguida.

DO PROBLEMA À EXPERIÊNCIA

OS ELEMENTOS ESSENCIAIS

Problema *Condução de uma Experiência*

Diversas decisões tomadas em relação a uma experiência não se encontram explicadas claramente nos relatos de pesquisa publicados. Embora uma parte dessa concisão possa ser atribuída à economia de espaço imposta pelos editores de publicações que apreciam artigos curtos, uma parte maior baseia-se na suposição de que os psicólogos experimentais, ou realmente os psicólogos que pesquisam uma determinada área específica, partilham um conhecimento básico comum. Isso é válido para todos os ramos da ciência. Por exemplo, um físico escrevendo em uma publicação presume que os leitores já saibam que uma dina é uma unidade de força e não se dará ao trabalho de explicar esse termo. De modo similar, os psicólogos presumem geralmente que o leitor sabe o que os termos estímulo e resposta significam, embora mesmo assim possam ser definidos. Uma finalidade deste livro consiste em proporcionar-lhe uma parte do vocabulário necessário, caso você deseje ler ou escrever sobre pesquisa em psicologia.

Um outro problema para o novo pesquisador relaciona-se ao "folclore de laboratório". "Todos" sabem que existem certas maneiras "óbvias" para realizar certos tipos de pesquisa. Essas maneiras diferem em cada área, mas são bem conhecidas em cada categoria. Elas são tão bem conhecidas a ponto de os pesquisadores poucas vezes preocuparem-se em explicá-las e realmente ficam muito surpresos quando os novos pesquisadores não conhecem esses truques e técnicas "óbvios". Os pesquisadores de animais muitas vezes os privam de alimento durante diversas horas antes do início do experimento ou mantêm seus pombos com uma certa porcentagem do peso que teriam se possuíssem alimento disponível continuamente. Embora as razões para isso sejam óbvias para o pesquisador, elas podem não ser óbvias para você. Como um experimentador sabe qual é o número de palavras a serem usadas em um experimento de memória? Quanto deve demorar um experimento? Por que uma variável dependente é selecionada de um conjunto que aparenta possuir variáveis dependentes igualmente válidas? Quantas pessoas devem participar de um experimento? As seções intituladas Do Problema ao Experimento nos capítulos da Parte II respondem a perguntas "óbvias" como essas.

Do Problema à Experiência

Toda pesquisa tem por objetivo resolver um problema que pode ser abstrato e teórico ou concreto e aplicado. O problema pode surgir de uma observação feita mais ou menos casualmente, como as pessoas serem mais agressivas durante o verão. Nesse caso, o problema pode ser formulado como "Por que o calor do verão causa agressividade?" ou, de modo ainda mais cético, como "A temperatura elevada causa agressividade?". Um problema pode surgir de uma descoberta acidental em um laboratório, como a descoberta de bolor em um pedaço de pão. Solucionar esse problema – por que o bolor está aumentando aqui? – levou à descoberta da penicilina. Finalmente, um problema pode surgir diretamente de um modelo teórico – por exemplo, quando perguntamos "Por que o reforço aumenta a probabilidade da ocorrência do comportamento que o precedeu?".

O primeiro passo que o experimentador precisa seguir consiste em transformar o problema em uma hipótese que possa ser testada. A hipótese precisa ser, então, transformada em um experimento com variáveis independentes, dependentes e de controle.

Do Problema à Hipótese Um problema é aproximadamente uma proposição vaga que precisa ser verificada ou uma pergunta que deve ser respondida. A não ser que cada uma dessas situações seja indicada de modo específico ou preciso, ela não pode ser testada experimentalmente. Toda hipótese é uma posição específica, originária de um problema e, muitas vezes, indicada da seguinte forma: se existir A, então ocorre B. A principal distinção entre um problema e uma hipótese é que esta pode ser testada diretamente, ao passo que um problema, não. Um teste experimental precisa ser capaz de tornar uma hipótese nula.

A finalidade de todo experimento consiste em testar hipóteses a respeito dos efeitos de uma ou mais variáveis independentes na variável dependente. Para fazer isso, precisamos coletar **dados**. Uma vez obtidos, esses dados precisam ser analisados e, então, divulgados. Discutimos brevemente esses aspectos a seguir.

▶ DADOS

Obtenção dos Dados

Planejar um delineamento experimental não estabelece todas as condições necessárias para a obtenção de dados. Embora o delineamento lhe explique como designar pessoas para participar de experimentos, não informa como consegui-las e, sem elas, não há dados.

Os psicólogos que investigam o comportamento animal possuem um controle muito maior sobre a seleção dos sujeitos experimentais do que os que estudam seres humanos. Embora os psicólogos que estudam animais precisem arcar com a despesa adicional de obter abrigo e alimentação para seus sujeitos experimentais, eles podem selecionar a linhagem que desejam adquirir e sempre ter sujeitos disponíveis, evitando algum fiasco.

A pesquisa com seres humanos utiliza, com muita freqüência, alunos de faculdade matriculados na disciplina Introdução à Psicologia como sujeitos experimentais. Desde que essa participação seja usada como uma oportunidade de aprendizagem para o aluno, ela é considerada ética e adequada (American Psychological Association, APA, 1987). Se o experimento não constituir uma oportunidade de aprendizagem, o experimentador deve pagar aos sujeitos. Em virtude de os alunos de faculdade representarem uma população selecionada, os experimentadores precisam ser cuidadosos em relação a generalizar resultados a outras populações de participantes. Por exemplo, técnicas de um sistema de ensino programado e individualizado criadas para ensinar química inorgânica podem não ter sucesso no ensino de hidráulica.

Seleção aleatória significa que qualquer membro de uma população possui a mesma probabilidade de ser selecionado como participante. Além disso, cada seleção é independente de outras seleções e, portanto, escolher uma pessoa não afeta a probabilidade de selecionar uma outra. Algumas vezes, em um experimento típico de psicologia, pode ser difícil especificar a população da qual se obtêm amostras (Gigerenzer, 1993). Mesmo se os participantes puderem ser selecionados aleatoriamente, qual população um conjunto de alunos da universidade representa exatamente? Não está claro se a população de alunos que cursam as disciplinas exigidas de psicologia é representativa de todos os alunos da universidade. Em virtude de a população de alunos ser agora tão diversificada, representando pessoas com idades e formações muito diferentes, os pesquisadores precisam ser cuidadosos a respeito da extrapolação de resultados da amostra de teste para outras populações.

Designação aleatória significa que cada sujeito experimental é designado aleatoriamente aos experimentos (Holland, 1993). Essa é uma técnica prudente, porque aumenta nossa possibilidade de fazer inferências causais a partir dos resultados experimentais. As implicações estatísticas da amostragem são discutidas no Apêndice B.

Após sua amostra ter sido selecionada e seu delineamento estabelecido, permanece uma decisão importante a ser tomada. Você deveria testar os participantes um por vez ou em um grupo? Ambos os procedimentos possuem vantagens e desvantagens. A maior vantagem de testar um grupo é a economia. Leva somente uma hora para testar 30 participantes por uma hora como um grupo, ao passo que demora 30 horas para testá-los individualmente. Assim, na igualdade dos demais fatores, é mais rápido e, portanto, melhor testar participantes em grupo. Porém, há muitos casos em que os demais fatores afastam-se muito da igualdade. Por exemplo, considere um experimento sobre audição em que palavras distintas são apresentadas aos ouvidos esquerdo e direito. Uma doutoranda apressada decidiu poupar tempo e testar os participantes como um grupo. Ela se esqueceu de que, a não ser que os participantes estivessem posicionados exatamente entre os dois alto-falantes, uma mensagem alcançaria um ouvido antes de a outra chegar ao outro ouvido. Isso tornou inválida a variável independente. Evidentemente, teria sido aceitável testar os participantes em um grupo, se cada pessoa tivesse fones de ouvido, evitando, assim, essa dificuldade. O outro problema no teste de grupo é a possibilidade de os participantes influenciarem-se mutuamente, influindo, desse modo, nos dados. Talvez um participante pudesse enganar e copiar respostas de outro ou a composição por sexo do grupo poderia alterar a motivação. Algumas vezes, esses problemas podem ser evitados colocando os participantes em cabines individuais que impedem a interação social.

Análise dos Dados

O resultado imediato de um experimento é uma grande série de números que representam o comportamento sob diferentes condições. Conforme Sidman (1960) descreve com humor, os cientistas acreditam que todos os dados já nascem falhos. Os dados pertencem ao Acaso ou à Ciência, porém nunca a ambos. Antes de o psicólogo poder estar seguro de que os dados pertencem à Ciência, o demônio Acaso precisa ser exorcizado. Isso é feito por meio de um ritual denominado *análise estatística inferencial*.

Após a análise estatística informar-lhe quais dados são confiáveis (não ocorreram por acaso), você ainda tem de decidir quais deles são importantes. Nenhum cálculo matemático pode indicar que hipóteses estão sendo testadas, aquilo que é previsto pelas teorias, e assim por diante. As estatísticas nunca substituem o pensamento. A análise estatística é um procedimento teoricamente neutro que atende à teoria e ao teste das hipóteses. Exceto no caso de um experimento "E se...?", as teorias e hipóteses precedem as estatísticas.

Em virtude de ser virtualmente impossível entender o significado do grande conjunto de números que um experimento produz, os dados geralmente são condensados pela estatística descritiva. A média e o desvio padrão são os mais comuns. Como parte da análise de dados, médias são calculadas para cada nível de cada variável independente, bem como para combinações de variáveis independentes para mostrar interações.

Relato dos Dados

Os dados são apresentados em tabelas ou gráficos. Gráficos geralmente são mais fáceis de compreender. A Figura 3.2 constitui um exemplo típico de como os resultados de um experimento são apresentados. A variável dependente é indicada no eixo das **ordenadas** – o eixo vertical. A variável independente é indicada no eixo das **abscissas** – o eixo horizontal. Mais de uma variável independente pode ser indicada no mesmo gráfico, usando linhas contínuas e pontilhadas e/ou símbolos de formato diferente para cada variável independente.

Dados brutos (não analisados) dificilmente chegam a ser apresentados. Como alternativa, algum parâmetro estatístico descritivo, como a média, é usado para resumir os dados. Outros parâmetros estatísticos muitas vezes acompanham os dados, para informar o leitor a respeito da confiabilidade desses dados.

Diversos estilos e formatos diferentes podem ser usados para divulgar os dados. Recomendamos o formato indicado no *Publication Manual of the American Psychological Association*, que se tornou a referência padrão em psicologia e em diversos outros campos das ciências sociais. Esse livro informará mais do que você gostaria de saber a respeito de todo aspecto relacionado ao preparo de um relato de pesquisa.

▶ RESUMO

1. Um experimento é um procedimento controlado para investigar os efeitos de uma ou mais variáveis independentes sobre uma ou mais variáveis dependentes. A *variável independente* é manipulada pelo experimentador, ao passo que a *variável dependente* é observada e registrada. Os experimentos oferecem ao investigador a melhor oportunidade para eliminar ou minimizar os efeitos de variáveis externas e indesejáveis. Os experimentos são realizados para testar teorias, para reproduzir e ampliar descobertas anteriores ou para mostrar que as pesquisas anteriores não podem ser confirmadas. Apenas raramente os experimentos são realizados para ver o que poderia acontecer.
2. As variáveis independentes são escolhidas porque um experimentador julga que elas controlarão o comportamento. Caso não o controlem, isso pode significar que a manipulação foi inadequada ou que o experimentador estava errado. Variáveis dependentes precisam ser *estáveis*, isto é, elas precisam produzir constantemente os mesmos resultados sob as mesmas condições. Os *efeitos de teto* e *de piso* resultam de uma variação inadequada da variável dependente. *Variáveis de controle* são variáveis independentes potenciais que não são manipuladas durante um experimento.
3. A maioria dos experimentadores testa mais de uma variável independente a cada vez. Além de proporcionar economia, isso permite ao experimentador obter informações importantes sobre interações. As *interações* ocorrem quando os efeitos de uma variável independente não forem os mesmos para diferentes níveis de uma outra variável independente. Ocasionalmente, os experimentadores usam mais de uma variável dependente.
4. O delineamento experimental designa participantes a condições diferentes, de modo que resulte na minimização de variáveis externas e indesejadas. Em um *delineamento de grupo*, grupos diferentes de participantes são sujeitos a tratamentos diferentes. Em um *delineamento de sujeito único*, todos os sujeitos são submetidos a todos os tratamentos. O delineamento de grupo é mais seguro, mas o delineamento de sujeito único é mais eficiente. *Delineamentos mistos* possuem algumas variáveis independentes tratadas entre diferentes sujeitos e outras tratadas intra-sujeitos. Nos delineamentos de grupo, os grupos equivalentes são formados por pareamento por igualação e por *randomização*. *Efeitos de prática geral* e *efeitos de histórias diferenciais* em delineamentos de sujeito único são avaliados, porém não eliminados por procedimento de contrabalanço. As condições de controle proporcionam uma linha de base com a qual pode ser comparada a condição (ou as condições) de interesse.
5. Há muitas armadilhas em delineamentos experimentais. As *características de demanda* resultam do conhecimento da pessoa de que está participando de um experimento. Os *efeitos do experimentador* são dispositivos introduzidos acidentalmente quando o experimentador (por meio do comportamento ou de características individuais) fornece informações a respeito da finalidade do experimento ou influencia sistematicamente o sujeito. Os efeitos do experimentador podem ser minimizados pelo uso de equipamentos, a fim de evitar diferenças sutis no comportamento do experimentador.
6. A seleção de sujeitos de alguma população é denominada **amostragem**. *Amostragem aleatória* significa que cada membro da população tem possibilidade igual de ser selecionado. É mais eficiente testar pessoas em grupos, porém deve-se tomar cuidado para evitar prejuízos ao experimento.

7. Quase-experimentos na psicologia muitas vezes empregam variáveis de sujeito. Essas variáveis são parâmetros como idade, QI, saúde mental, altura, cor do cabelo e inúmeras outras características que diferem de uma pessoa para outra. Tais variáveis são determinadas após o fato, pois as pessoas muitas vezes possuem traços herdados (ou, pelo menos, as pessoas participam do estudo psicológico com a variável já determinada). Por não ser possível designar pessoas aleatoriamente às condições de interesse, estudos que usam variáveis de sujeito são inerentemente correlacionais por natureza.
8. Para tentar fazer proposições do tipo causa e efeito a partir da manipulação das variáveis de sujeito, os pesquisadores muitas vezes pareiam por igualação os sujeitos com base em pareamentos por igualação de outras variáveis. Portanto, se um pesquisador estivesse interessado nos efeitos da cor do cabelo sobre o desempenho em alguma tarefa ou na reação das demais pessoas em alguma situação, ele tentaria controlar o maior número possível de outras variáveis, a fim de assegurar que a cor do cabelo fosse o único aspecto no qual as pessoas nas várias condições diferissem. O pareamento por igualação muitas vezes é uma ferramenta útil para essas finalidades, porém, deve-se ter certeza de que a possibilidade de haver dispositivos de regressão não atrapalhe as conclusões.
9. *Regressão para a média* refere-se ao fato de que, quando um subgrupo com resultados extremos é obtido de um grupo maior e testado novamente, os membros tenderão a se localizar mais próximos da média de todo o grupo no segundo teste. Se, ao parear por igualação dois grupos com base em um primeiro teste, o pesquisador estiver obtendo bons resultados de um grupo que geralmente tem mau desempenho e resultados ruins de um grupo que geralmente possui um bom desempenho, então, mesmo se os grupos não forem tratados diferencialmente em um experimento, o pesquisador pode esperar que as pessoas apresentem resultados diferentes em um segundo teste, simplesmente por causa da regressão para a média. Esse problema é conhecido como um *dispositivo de regressão*.

▶ TERMOS-CHAVE

abscissa
amostragem
características de demanda
condição de controle
contrabalanço
dados
delineamentos de baixo *n*
delineamentos de grupo
delineamentos de sujeito único
delineamentos mistos
designação aleatória
dispositivo de regressão
efeito Hawthorne
efeito de piso
efeito de teto
efeitos de histórias diferenciais
efeitos de prática geral
efeitos do experimentador
estabilidade
experimento
experimento crítico
experimento "E se...?"
experimento duplo-cego
grupo controle
interação
linha de base
nível
ordenada
pareamento por igualação
quase-experimentos