

relação entre as variáveis é sempre positiva ou sempre negativa (ela não muda de direção como na relação curvilínea ou não monotônica, ilustrada no Gráfico C). A Figura 4.3 mostra um exemplo de função monotônica positiva que não é estritamente linear.

Lembre que esses são padrões gerais. Mesmo que, em geral, exista uma relação linear positiva, isso não significa que todos os que tiveram escores elevados numa variável terão escores elevados na segunda variável. Desvios individuais em relação a um padrão geral são prováveis. Além disso, para conhecer o tipo de relação predominante entre duas variáveis, é necessário conhecer também a força da relação; isto é, precisamos saber qual é a magnitude da correlação entre as variáveis. Algumas vezes, duas variáveis estão fortemente relacionadas entre si e há pouco desvio em relação ao padrão predominante. Outras vezes, as duas variáveis não estão fortemente correlacionadas, porque muitos indivíduos desviam-se do padrão predominante. Um índice numérico da força da relação entre variáveis é chamado **coeficiente de correlação**. Os coeficientes de correlação são muito importantes; precisamos saber quão fortemente uma variável está ligada a outra. Discutiremos coeficientes de correlação em detalhes nos Capítulos 5 e 12.

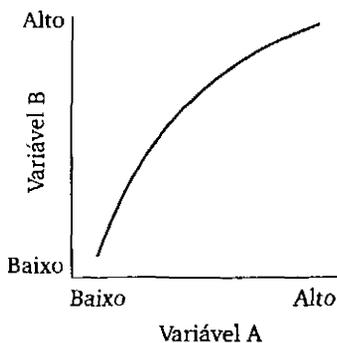


Figura 4.3 Função monotônica positiva.

## MÉTODO NÃO EXPERIMENTAL VERSUS MÉTODO EXPERIMENTAL

Como podemos determinar se existe relação entre as variáveis? Há duas abordagens gerais para estudar relações entre variáveis: o método não experimental e o método experimental. Com o método não experimental, as relações são estudadas por meio de observações ou mensuração das variáveis de interesse. (Os termos *observação* e *mensuração* serão usados alternativamente no decorrer do

texto.) Isto é, o comportamento é observado como ocorre naturalmente. Isso é feito pedindo-se às pessoas para descrever seu comportamento, observando-se o comportamento diretamente, registrando-se respostas fisiológicas ou mesmo examinando-se vários registros públicos, tais como dados de censo. Uma relação entre variáveis é estabelecida quando as duas variáveis variam em conjunto. Por exemplo, em um estudo que será descrito no Capítulo 7, Steinberg e Dornbusch (1991) relacionaram o número de horas de trabalho às médias escolares de estudantes secundaristas. As duas variáveis variaram juntas; estudantes que trabalhavam maior número de horas tendiam a ter notas mais baixas.

A segunda abordagem para estudar relações entre variáveis é chamada método experimental. O método experimental envolve manipulação direta e controle das variáveis. O pesquisador manipula a primeira variável de interesse e, então, observa a resposta. Por exemplo, Loftus (1979) usou o método experimental ao perguntar a participantes que haviam visto um filme mostrando um acidente automobilístico, se viram "um" farol dianteiro quebrado ou "o" farol dianteiro quebrado. O método de questionamento foi manipulado e as respostas dos participantes foram, então, medidas. Com esse método, as duas variáveis não variam apenas conjuntamente; uma variável é introduzida primeiro para verificar se afeta a segunda variável. Essa diferença entre os métodos tem implicações importantes.<sup>1</sup>

### Método Não Experimental

Suponha que um pesquisador esteja interessado nas relações entre exercício e ansiedade. Como poderia estudar esse problema? Usando o método não experimental, o pesquisador deveria elaborar definições operacionais para medir a quantidade de exercício realizada pelas pessoas e seu nível de ansiedade. Poderia haver diferentes maneiras de definir operacionalmente essas duas variáveis; por exemplo, poder-se-ia pedir, simplesmente, às pessoas para fazerem um auto-relato de seus padrões de exercício e de seus níveis atuais de ansiedade. O ponto importante a ser lembrado é que ambas as variáveis são medidas quando usamos o método não experimental. Agora suponha que o pesquisador colete dados sobre exercício e ansiedade de certo número de pessoas e que encontre uma relação negativa entre exercício físico e ansiedade. Isto é, quanto mais as pessoas praticam exercício físico, menor é seu nível de ansiedade. As duas variáveis variam conjuntamente, ou seja, estão relacionadas; diferenças observadas em exercício físico estão

<sup>1</sup> Sobre delineamento de experimentos e quase-experimentos sugiro também a leitura de CAMPBELL, D. T.; STANLEY, J. C. *Delineamentos experimentais e quase-experimentais de pesquisa*. São Paulo: EPU, 1979 (NT).

associadas com o grau de ansiedade. Como o método não experimental permite-nos observar a variação conjunta de variáveis, outro termo frequentemente utilizado é **método correlacional**. Com esse método, examinamos se as variáveis estão correlacionadas ou variam juntas.

O método não experimental parece ser uma abordagem razoável para estudar relações entre variáveis como exercício e ansiedade. Uma relação é estabelecida ao se verificar que as duas variáveis variam juntas – as variáveis covariam ou estão correlacionadas entre si. No entanto, percebemos uma fraqueza desse método ao formular questões sobre causa e efeito.<sup>2</sup>

Sabemos que as duas variáveis estão relacionadas, mas o que podemos dizer sobre o impacto causal de uma variável sobre a outra? Há dois problemas para fazer afirmações causais quando o método não experimental é usado: (1) a direção de causa e efeito e (2) o problema da terceira variável – isto é, a variável externa que pode estar causando a relação observada.<sup>3</sup>

A Figura 4.4 ilustra esse problema, representando as ligações causais entre variáveis por meio de setas.

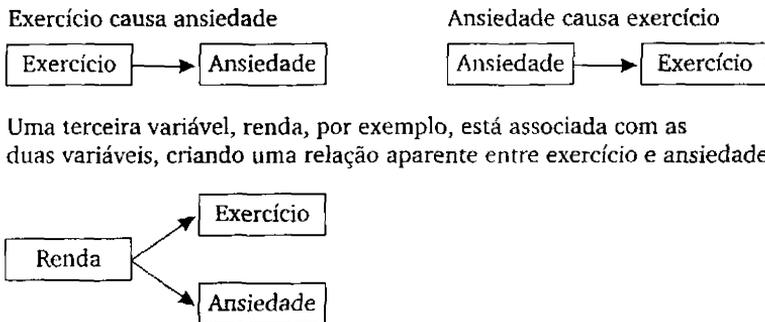


Figura 4.4 Possibilidades causais num estudo não experimental.

**Direção de causa e efeito.** O primeiro problema é o da direção de causa e efeito. Com o método não experimental é difícil determinar qual variável causa a outra. Em outras palavras, não podemos realmente dizer que o exercício físico

2 Conforme Brugger, "ciência é o conhecimento certo das coisas por suas causas principais". BRUGGER, W. *Dicionário de filosofia*. São Paulo: EPU, 1987. p. 84-85 (NT).

3 Para um aprofundamento no problema da terceira variável, recomendo Rosenberg, M. *A lógica do levantamento e análise de dados*. São Paulo: Cultrix: Edusp, 1976. Para um aprofundamento sobre a direção de causa e efeito, recomendo KENNY, D. A. *Correlation and causality*. New York: Wiley, 1979, e HEISE, D. R. *Causal analysis*. New York: Wiley, 1975 (NT).

causa uma redução da ansiedade. Embora haja razões plausíveis para esse modelo de causa e efeito, há também razões para que o oposto possa ocorrer. Talvez, o nível de ansiedade cause o exercício físico; talvez, a experiência de ansiedade elevada interfira com a capacidade de fazer exercício físico. Essa é uma diferença importante. Se os exercícios reduzem a ansiedade, então, aderir a um programa de exercício pode ser um caminho razoável para reduzir a própria ansiedade. No entanto, se a ansiedade leva as pessoas a interromper o exercício, forçar alguém a exercitar-se provavelmente não irá reduzir seu nível de ansiedade.

No entanto, o problema da direção da causa e efeito não é o pior problema do método não experimental. Por exemplo, os cientistas lembram que os astrônomos conseguem fazer previsões acuradas, embora não possam manipular variáveis experimentalmente. Além disso, a direção de causa e efeito frequentemente não é crucial, porque para alguns pares de variáveis o modelo causal pode operar nos dois sentidos. Por exemplo, parece haver dois modelos causais na relação entre as variáveis similaridade e gostar: (1) a similaridade faz com que pessoas gostem umas das outras; e (2) gostar torna as pessoas mais similares. Em geral, o problema da terceira variável é uma falha muito mais séria do método não experimental.

**O problema da terceira variável.** Quando o método não experimental é usado, há o perigo de que não haja uma relação causal direta entre duas variáveis. É possível que os exercícios não influenciem a ansiedade e que a ansiedade não tenha efeito causal sobre os exercícios. Por outro lado, é possível que haja uma relação entre essas duas variáveis porque outra está causando ambas, exercício e ansiedade. Isso é conhecido como o problema da terceira variável. Muitas variáveis podem exercer o papel de terceira variável e ser responsáveis por uma relação observada entre duas variáveis. No exemplo do exercício e da ansiedade, uma terceira variável poderia ser o nível de renda. Talvez uma renda elevada propicie às pessoas mais tempo livre para fazer exercícios físicos (e a possibilidade de pagar uma academia) e também faça a ansiedade baixar. Se a renda é a variável determinante, não há relação de causa e efeito entre exercício e ansiedade; a relação foi causada pela terceira variável, nível de renda.

Como você pode ver, direção de causa e efeito e possibilidade de existir terças variáveis representam sérias limitações ao método não experimental. Frequentemente, elas não são consideradas nos relatos de resultados de pesquisas divulgados na mídia. Por exemplo, um periódico científico pode relatar os resultados de um estudo não experimental que encontrou uma relação positiva entre a quantidade de café consumida e a probabilidade de ocorrência de um ataque cardíaco. É fácil verificar que não há necessariamente uma relação de causa e efeito entre as duas variáveis. Numerosas terças variáveis (por exemplo, ocupação, personalidade, predisposição genética) podem causar ambos, tanto o comportamento de beber café quanto a probabilidade de um ataque cardíaco.

Em suma, os resultados desses estudos são ambíguos e devem ser encarados com ceticismo.<sup>4</sup>

### Método Experimental

O método experimental reduz essa ambigüidade na interpretação dos resultados. Com o método experimental uma variável é manipulada e a outra é, então, medida. Se um pesquisador decidiu utilizar o método experimental para estudar se exercícios físicos reduzem a ansiedade, os exercícios seriam manipulados – por exemplo, comparando um grupo de pessoas que se exercitasse todos os dias durante uma semana com outro grupo que se abstivesse de exercícios. A ansiedade poderia então ser medida. Suponha que se constate que as pessoas praticantes de exercícios físicos têm menor ansiedade em comparação com aquelas que não se exercitam. O pesquisador pode agora dizer algo sobre a direção de causa e efeito: no experimento, os exercícios vieram primeiro na seqüência de eventos e, assim, o nível de ansiedade não poderia ter influenciado a quantidade de exercícios em que uma pessoa se engaja.

Outra característica do método experimental é que ele busca eliminar a influência de todas as terceiras variáveis estranhas. Isso é denominado controle de variáveis estranhas. Geralmente se consegue esse controle mantendo todas as características do ambiente constantes, exceto a variável manipulada.<sup>5</sup>

Qualquer variável que não seja mantida constante é controlada, garantindo-se que seu efeito seja randomizado. Por meio da randomização, a influência de qualquer variável estranha é equivalente em todas as condições experimentais. Ambos os procedimentos são utilizados para garantir que quaisquer diferenças entre os grupos são devidas à variável manipulada.

**Controle experimental.** Com o controle experimental todas as variáveis estranhas são mantidas constantes. Se uma variável é mantida constante, ela não pode ser responsável pelos resultados do experimento. No experimento sobre o efeito do exercício físico, o pesquisador gostaria de assegurar que a diferença entre os grupos com e sem exercício é devida somente ao exercício. Tendo em vista que pessoas do grupo com exercício saem de sua rotina para exercitar-se, pessoas do grupo sem exercício também devem ser retiradas de sua rotina. Caso contrário, os níveis baixos de ansiedade no grupo com exercício poderiam ter resultado do “descanso” da rotina diária e não do exercício.

4 A presença de correlação espúria pode levar ao aparecimento do Paradoxo de Simpson ou Stein (<http://www.autobox.com/spur6.html>) (NT).

5 Também conhecido por Princípio *Ceteris Paribus* (NT).

Consegue-se controle experimental tratando as pessoas de todos os grupos do experimento de forma idêntica; a única diferença entre os grupos é a variável manipulada. No experimento de Loftus sobre memória, ambos os grupos testemunharam o mesmo acidente, o mesmo experimentador fez as perguntas para ambos os grupos, a iluminação e todas as outras condições foram as mesmas, e assim por diante. Nesse caso, se houver uma diferença no relato dos dois grupos, podemos estar certos de que foi devida ao método de questionamento e não a qualquer outra variável que não tenha sido mantida constante.

**Randomização.** Algumas vezes, é difícil manter uma variável constante. A mais óbvia dessas variáveis diz respeito a características dos participantes. Considere um estudo experimental em que metade dos participantes da pesquisa está na condição com exercício e metade, na condição sem exercício; os participantes nas duas condições podem diferir quanto a alguma terceira variável não controlada (covariável), como por exemplo a renda. Essa diferença pode causar uma relação aparente entre exercício e ansiedade. Como o pesquisador pode eliminar a influência de covariáveis desse tipo num experimento?

O método experimental elimina a influência dessas variáveis por meio de randomização. A randomização assegura que a covariável tenha a mesma probabilidade de afetar tanto um grupo experimental quanto o outro. Para eliminar a influência de características individuais, o pesquisador distribui os participantes pelos dois grupos de maneira randômica. De fato, isso significa que a distribuição é determinada por meio de uma lista de números randômicos. Para compreender isso, pense que os participantes de um experimento formam uma linha. Cada pessoa vai para a frente da fila e recebe um número randômico, da mesma forma que se sorteiam números randômicos na loteria. Se o número for par, o indivíduo é designado para um grupo (por exemplo, exercício), e se for ímpar, é designado para outro grupo (por exemplo, não-exercício). Usando um procedimento de designação randômica, o pesquisador confia que as características dos participantes dos dois grupos são praticamente idênticas. Nessa "loteria", por exemplo, pessoas com renda baixa, média e alta, estarão distribuídas igualmente nos dois grupos. De fato, a randomização assegura que a composição das características individuais dos dois grupos seja praticamente idêntica em todos os aspectos. A designação randômica dos participantes da pesquisa às condições experimentais é uma diferença importante entre o método experimental e o método não experimental.

Para tornar o conceito de designação randômica mais concreto, você pode tentar fazer um exercício parecido com o que eu fiz com uma caixa cheia de figurinhas velhas de jogadores de beisebol. A caixa continha figurinhas de 50 jogadores da Liga norte-americana e de 50 jogadores da Liga Nacional. Misturei completamente as figurinhas e, em seguida, selecionei 32 delas e as designei a "grupos", usando a lista de números aleatórios no Apêndice C.1. A medida que

selecionava cada figurinha, usava a seguinte regra: se o número randômico fosse par, o jogador iria para o "Grupo 1", e se o número fosse ímpar, ele iria para o "Grupo 2". Em seguida, procurei verificar se meus dois grupos diferiam em termos da representação na Liga. O Grupo 1 tinha 9 jogadores da Liga Norte-Americana e 7 jogadores da Liga Nacional, enquanto o Grupo 2 tinha igual número de jogadores de ambas as ligas. Os dois grupos eram praticamente idênticos!

Qualquer outra variável que não possa ser mantida constante é também controlada pela randomização. A propósito, muitos experimentos são realizados durante vários dias ou semanas e os participantes chegam para participar dos experimentos em diferentes horários. Nesses casos, o pesquisador usa uma ordem randômica para esquematizar a seqüência das várias condições experimentais. Esse procedimento evita uma situação em que uma condição é estudada durante os primeiros dias do experimento enquanto outra somente é estudada durante os dias posteriores. Da mesma forma, participantes de um grupo não serão estudados somente pela manhã e outros, somente à tarde.

Controle direto e randomização eliminam a influência de quaisquer covariáveis. Assim, o método experimental permite uma interpretação relativamente não ambígua dos resultados. Qualquer diferença entre os grupos na variável observada pode ser atribuída somente à influência da variável manipulada.<sup>6</sup>

## VARIÁVEL INDEPENDENTE E VARIÁVEL DEPENDENTE

Quando os pesquisadores estudam a relação entre duas variáveis, imaginam que elas têm uma relação de causa e efeito; isto é, uma variável é considerada a "causa" e a outra, o "efeito". Assim, a credibilidade do orador é vista como causa da mudança de atitude, e o exercício é visto como afetando a ansiedade. Tanto os pesquisadores que usam o método experimental quanto os que usam o método não experimental pensam as variáveis dessa forma, embora, como vimos, haja menos ambigüidade sobre a direção de causa e efeito quando o método experimental é usado. Os pesquisadores usam os termos **variável independente** e **variável dependente** quando se referem às variáveis que estão sendo estudadas. A variável considerada "causa" é a variável independente, e a variável considerada "efeito", a variável dependente.

---

6 A título de aprofundamento nesse assunto, recomendo a leitura do Capítulo 26, Design of experiments, randomization, and sample size planning, de NETER, J. et al. *Applied linear statistical models*. 4. ed. San Francisco: McGraw-Hill: Irwin, 1996 (NT).

Num experimento, a variável manipulada é a variável independente, e a segunda variável, que é medida, é a variável dependente. Uma maneira de lembrar essa distinção é relacionar os termos com o que acontece com um indivíduo durante um experimento. O pesquisador cria uma situação a que expõe os participantes, como por exemplo assistir a um programa violento versus um programa não violento ou realizar exercício físico versus não realizar. Essa situação é a variável manipulada; ela é denominada variável independente porque o participante não tem controle sobre sua ocorrência. No passo seguinte do experimento, o pesquisador mede a resposta à variável manipulada. O participante responde à situação a que foi submetido, e o pesquisador supõe que seu comportamento ou aquilo que ele diz é causado ou depende do efeito da variável independente (manipulada). A variável independente, então, é a variável manipulada pelo experimentador, e a variável dependente é o comportamento medido, causado pela variável independente.

Quando se representa graficamente a relação entre uma variável independente e uma variável dependente, a independente é sempre colocada no eixo horizontal e a dependente, no eixo vertical. Examinando a Figura 4.2, o leitor verá que esse método gráfico foi usado para apresentar as quatro relações. No Gráfico B, por exemplo, a variável independente "tamanho do grupo" aparece no eixo horizontal; a variável dependente, "quantidade de barulho", no eixo vertical.

Observe que algumas pesquisas focalizam basicamente a variável independente, e o pesquisador estuda o efeito de uma única variável independente sobre numerosos comportamentos. Outros pesquisadores podem focalizar uma variável dependente específica e estudar como diversas variáveis independentes afetam esse comportamento. Para tornar essa distinção mais concreta, considere um estudo sobre o efeito do tamanho do júri no resultado de um julgamento. Um pesquisador que esteja estudando esse assunto poderia estar interessado no efeito do tamanho do grupo sobre diferentes comportamentos, incluindo decisões de um júri ou riscos assumidos por empresários. Outro pesquisador, interessado somente em decisões do júri, poderia estudar os efeitos de várias características de um julgamento, como tamanho do júri ou instruções do juiz, sobre o comportamento dos jurados. Tanto num caso como no outro podem ser derivadas pesquisas importantes.

## CAUSALIDADE

Ao descrever relações entre variáveis enfatizamos inferências sobre causa e efeito. Que significa dizer que uma variável tem um efeito causal sobre outra? Inferimos causalidade com base na informação de que dispomos sobre a variá-

vel em questão. Em primeiro lugar, precisamos conhecer alguma coisa sobre a relação temporal da causa e do efeito. A variável causal deveria vir antes na ordem temporal dos eventos e, então, ser seguida pelo efeito. Ao utilizar o método experimental, o pesquisador trata da ordem temporal manipulando primeiro a variável independente e, então, observando se tem efeito sobre a variável dependente. Em outras situações, podemos observar a ordem temporal ou concluir logicamente que uma ordem é mais plausível do que outra. Em segundo lugar, deve haver covariação das duas variáveis. Uma mudança na primeira variável deve ser acompanhada por uma mudança na segunda. Em terceiro lugar, há necessidade de eliminar explicações alternativas para a relação observada. Uma explicação alternativa baseia-se na possibilidade de alguma outra variável ser responsável pela relação observada. Essas variáveis são chamadas "terceiras" variáveis. Quando planejamos uma pesquisa, é importante eliminar explicações alternativas. O método experimental começa tentando tornar essas variáveis constantes por meio de designação randômica e controle experimental. Outras questões de controle serão discutidas em capítulos posteriores. Em geral, inferências sobre relações causais são mais fortes quando existem poucas explicações alternativas para a relação observada.

Algumas vezes, é preciso impor exigências mais severas para concluir que há uma relação causal. Alguns cientistas, filósofos e mesmo muitos estudantes argumentam que a relação de causa e efeito somente é comprovada se a causa for necessária e suficiente para a ocorrência do efeito. Suponha que você conclua que ler a matéria para um exame está relacionado com as notas obtidas no exame; estudantes que leram a matéria têm notas mais altas do que os que não leram. Para ser necessária, a causa deve estar presente para que o efeito ocorra. Para provar que ler a matéria é a causa das altas notas, é preciso demonstrar que ler a matéria precisa ocorrer para que o estudante tenha bom desempenho no exame. Para ser suficiente, a causa deverá produzir o efeito sempre. Para provar que ler a matéria é a causa, é preciso que, tendo lido a matéria, o estudante sempre obtenha notas altas no exame.

Analisemos essa situação em termos das condições necessárias e suficientes. Se estivermos falando de uma disciplina em que o exame se baseia apenas na leitura de um livro, a obtenção de uma boa nota provavelmente dependerá da leitura. Nesse ponto, alguns estudantes dirão que basta prestar atenção nas aulas e fazer um bom exame sem ler o livro. Isso pode ser verdade em algumas turmas, mas é importante delimitar condições numa discussão sobre relações de causa e efeito.

Ler a matéria é suficiente para ir bem no exame? Ou seja, ler a matéria sempre resulta em notas altas num exame? Você deve estar pensando agora quantas vezes leu a matéria, mas não foi bem num exame. Ler a matéria não é uma causa suficiente; provavelmente, é preciso prestar atenção para reter a

matéria, relacionar a informação com o que já sabe e praticar, recordando a matéria. Novamente, é mais provável que a pesquisa descubra as condições necessárias em que a matéria produz notas altas num exame.

O requisito "necessário e suficiente" para estabelecer a causa é raro em Psicologia. Sempre que psicólogos afirmam que há uma causa necessária e suficiente para o comportamento, logo surgem pesquisas mostrando que não é bem assim. Por exemplo, em certa época os psicólogos afirmavam que "frustração causa agressão"; sempre que ocorresse frustração, resultaria em agressão e sempre que ocorresse agressão, a causa precedente seria frustração. Demonstrou-se que essa afirmação era incorreta. Frustração pode levar a agressão, mas também a outras respostas (tais como esquivar passiva ou aumento de esforço para superar a frustração). Além disso, a agressão pode resultar da frustração, mas outros eventos também podem produzir agressão, tais como dor, insulto ou ataque direto.

Os cientistas do comportamento não se preocupam indevidamente com questões de causa e efeito últimas. Em vez disso, estão mais interessados em descrever cuidadosamente o comportamento, estudar como as variáveis afetam umas às outras e em desenvolver teorias que expliquem o comportamento. O consenso geral é que há poucas causas "necessárias e suficientes" do comportamento que sejam interessantes. Em vez disso, pesquisas que envolvem diferentes variáveis finalmente levam à compreensão de toda uma "rede causal" na qual diferentes variáveis estão envolvidas num padrão complexo de causa e efeito. Nosso livro não tratará dessas questões complexas, mas dos métodos usados para estudar o comportamento.

## ESCOLHA DE UM MÉTODO: VANTAGENS E DESVANTAGENS DE DIFERENTES MÉTODOS

Enfatizamos as vantagens da utilização do método experimental para estudar relações entre variáveis. No entanto, também há desvantagens em realização de experimentos e boas razões para usar outros métodos. A seguir, examinaremos algumas questões que surgem na escolha de um método.

### Artificialidade dos Experimentos

Num experimento de laboratório, manipula-se a variável independente dentro de um ambiente de laboratório cuidadosamente controlado. Esse procedimento permite inferências relativamente não ambíguas a respeito de causa e efeito e reduz a possibilidade de variáveis estranhas influenciarem os resultados.

Esse é um recurso extremamente útil para estudar muitos problemas. No entanto, o alto grau de controle e o ambiente do laboratório podem criar, algumas vezes, uma atmosfera artificial, que limita tanto as questões tratadas quanto a generalidade dos resultados. Por essa razão, os pesquisadores podem optar pelo uso de métodos não experimentais.

Outra alternativa é tentar realizar um experimento em condições de campo. Num **experimento de campo**, a variável independente é manipulada em ambiente natural. Como em qualquer experimento, o pesquisador recorre tanto a randomização quanto a controle experimental na tentativa de controlar variáveis estranhas. Como exemplo de um experimento de campo, considere o estudo de Langer e Rodin (1976) sobre os efeitos de dar aos residentes de lares de idosos maior controle sobre as decisões que afetam suas vidas. A um grupo de residentes atribuiu-se a responsabilidade de fazer escolhas no que diz respeito ao funcionamento do lar; um segundo grupo foi levado a acreditar que a equipe de profissionais era responsável por seus cuidados e suprimento de suas necessidades. Os experimentadores mediram as variáveis dependentes, entre as quais o nível de atividade e a alegria dos residentes. Os resultados mostraram que os membros do grupo ao qual se atribuiu responsabilidade aumentada foram mais ativos e felizes. Um estudo de acompanhamento revelou que também tiveram ganho de saúde física (Rodin; Langer, 1977).

Muitos outros experimentos de campo acontecem em espaços públicos, tais como esquinas, centros comerciais e estacionamentos. Ruback e Juieng (1997) mediram quanto tempo motoristas levavam para deixar sua vaga num estacionamento, sob duas condições: (1) quando outro carro estava esperando pela vaga ou (2) quando não havia outro carro esperando. Como poderíamos esperar, os motoristas levavam mais tempo para tirar seus carros quando um carro estava esperando pela vaga. Aparentemente, a motivação para proteger um território temporário era mais forte do que a motivação para sair o mais rapidamente possível! A vantagem do experimento de campo é a investigação da variável independente num contexto natural. A desvantagem está na impossibilidade de o pesquisador controlar diretamente muitos aspectos da situação. O experimento de laboratório permite aos pesquisadores manter constantes, com mais facilidade, as covariáveis, eliminando assim sua influência no resultado do experimento. De fato, é exatamente esse controle que torna a investigação de laboratório mais artificial. Felizmente, experimentos realizados em situação de laboratório e de campo geralmente produzem resultados semelhantes (Anderson; Lindsay; Bushman, 1999).

### Considerações Éticas e Práticas

Algumas vezes, o método experimental não é uma alternativa viável, porque a experimentação pode ser antiética ou impraticável. A manipulação de

práticas de cuidado infantil por meio do método experimental, por exemplo, pode ser inviável. Além disso, mesmo que fosse possível designar pais randomicamente para duas condições de cuidados, tais como retirada de amor *versus* punição física, a manipulação seria antiética. Em lugar de manipular variáveis, como a das técnicas de criação de crianças, os pesquisadores usualmente estudam como elas ocorrem em ambiente natural. Muitas áreas importantes de pesquisa apresentam problemas semelhantes – por exemplo, estudos dos efeitos do alcoolismo, do divórcio e suas conseqüências, ou do impacto do trabalho da mãe fora de casa sobre as crianças. Esses problemas precisam ser estudados e, geralmente, as únicas técnicas possíveis são não experimentais.

Ao estudar essas variáveis, os pesquisadores costumam classificar as pessoas em grupos com base em sua experiência. Num estudo sobre os efeitos do emprego materno, um grupo poderia ser composto por mães que trabalham fora e outro por mães que não trabalham fora. Isso é algumas vezes denominado um delimitamento *ex post facto*. *Ex post facto* significa “após o fato” – o termo foi criado para descrever pesquisas nas quais os grupos são formados com base em alguma diferença real, em vez de serem criados por designação randômica, como num experimento. É extremamente importante estudar essas diferenças. No entanto, é importante reconhecer que essa é uma pesquisa não experimental porque não há designação randômica dos grupos.

### Variáveis do Participante

As variáveis do participante (também denominadas variáveis do sujeito) são características dos indivíduos, tais como idade, sexo, personalidade ou estado civil. Por definição, essas variáveis são não experimentais e, portanto, podem ser apenas medidas. Por exemplo, para estudar uma característica de personalidade, como a extroversão, o leitor pode submeter pessoas a um teste de personalidade planejado para medir essa variável. Essas variáveis podem ser estudadas em experimentos juntamente com variáveis independentes manipuladas (veja Capítulo 10).

### Descrição do Comportamento

Um objetivo principal da ciência é fornecer uma descrição acurada dos eventos. Assim, o objetivo de muitas pesquisas é descrever comportamento; as questões tratadas por experimentos não são relevantes para os objetivos das pesquisas. Um exemplo clássico de pesquisa descritiva em psicologia pode ser encontrado no trabalho de Jean Piaget, que cuidadosamente observou o comportamento de seus próprios filhos, ao longo de seu desenvolvimento, e descreveu, em detalhes, as mudanças em sua forma de pensar e responder ao ambien-

te (Piaget, 1952). As descrições de Piaget e suas interpretações das observações feitas resultaram numa importante teoria do desenvolvimento cognitivo, que aumentou muito nossa compreensão do tema. A teoria de Piaget teve um impacto importante sobre a Psicologia, que persiste até hoje (Flavell, 1996).

### **Previsões Bem-Sucedidas do Comportamento Futuro**

Em muitas situações de vida real, uma preocupação essencial é fazer uma previsão bem-sucedida do comportamento futuro de uma pessoa – por exemplo, sucesso na escola, capacidade para aprender um novo trabalho ou grandes áreas de interesse na universidade. Nessas circunstâncias, questões de causa e efeito podem não ser objeto de preocupação. É possível planejar medidas que aumentem a acuracidade da previsão do comportamento futuro. Orientadores educacionais podem aplicar testes para decidir se estudantes devem ser colocados em classes “avançadas”; empregadores podem testar candidatos para decidir se devem ou não ser contratados; e universitários submetem-se a testes em busca de auxílio para sua opção por uma área de estudo. Esses tipos de medidas podem ajudar muitas pessoas a tomar decisões melhores. Quando desenvolvem medidas para prever comportamento futuro, os pesquisadores precisam realizar pesquisas para demonstrar que a medida, de fato, relaciona-se com o comportamento em questão. Esse tipo de pesquisa será discutido no Capítulo 5.

### **Vantagens do Uso de Múltiplos Métodos**

Talvez o mais importante seja reconhecer que a compreensão completa de qualquer fenômeno requer a utilização de múltiplos métodos de estudo, tanto experimentais quanto não experimentais. Nenhum método é perfeito e nenhum estudo isoladamente é definitivo. Considere, a título de exemplo, uma hipótese desenvolvida por Frank e Gilovich (1988). Eles ficaram intrigados com a observação de que a cor preta representa maldade e morte em muitas culturas e em diferentes épocas e resolveram verificar se isso influencia nosso comportamento. Observaram que muitos times esportivos profissionais usam uniformes pretos e levantaram a hipótese de que esses times poderiam ser mais agressivos em comparação com outros que não usam uniformes pretos.

Primeiro, eles precisavam de uma definição operacional de uniformes “pretos” e “não pretos”; decidiram que uniforme preto era aquele que tinha 50% ou mais da cor preta. Usando essa definição, cinco times NFL e cinco times NHL<sup>7</sup>

<sup>7</sup> O Autor utilizou a sigla NFL para referir-se à National Football League e a sigla NHL para referir-se à National Hockey League (NT).

tinham uniformes pretos. Começaram pedindo a pessoas que não tinham conhecimento da NFL ou da NHL para examinar cada uniforme e, então, avaliar os times usando adjetivos “malévolos”, tais como “mesquinho” e “agressivo.” Em geral, as avaliações para os times de uniformes pretos mostraram que eles foram percebidos como mais malévolos. Em seguida, compararam a distância em jardas das faltas dos times do NFL preto e não preto e a duração das mesmas em minutos dos times do NHL. Nos dois casos, os times pretos sofreram mais faltas – mas pode-se dizer que esse é um modelo causal? Frank e Gilovich descobriram que dois times NHL que usavam uniformes não pretos passaram a usar uniformes pretos e compararam as durações das faltas antes e depois da troca. O resultado foi consistente com a hipótese: as faltas aumentaram para os dois times. Também examinaram as durações das faltas de um terceiro time que trocou um uniforme não preto por outro não preto e não encontraram qualquer mudança. Note que nenhum desses estudos usou o método experimental. Num experimento para testar a hipótese de que as pessoas percebem times que usam uniformes pretos como mais agressivos, os pesquisadores exibiram a estudantes vídeos de dois jogos de futebol simulados, em que a defesa estava usando uniforme branco ou preto. Nos dois jogos, a defesa realizou uma ação agressiva. Nesses jogos, os estudantes tendiam a penalizar mais o time que estava usando uniforme preto do que o time que estava usando uniforme não preto. Num último experimento, as pessoas eram trazidas ao laboratório em grupos de três, para verificar se pertencer ao time de uniforme preto poderia aumentar a agressividade. Dizia-se aos grupos que eles formavam um “time” que iria competir com outro. Os membros do time recebiam camisas brancas ou pretas que deveriam vestir para a competição. Em seguida, eles deviam escolher um jogo para a competição. Alguns jogos eram agressivos (“duelo de dardos”) e outros não eram (“jogos de arremesso”). Como esperado, o time de uniforme preto escolheu jogos mais agressivos.

O ponto importante aqui é constatar que nenhum estudo isoladamente é um teste perfeito de uma hipótese. No entanto, quando diferentes estudos que usam múltiplos métodos levam à mesma conclusão, aumenta muito nossa confiança nas descobertas e nossa compreensão do fenômeno.

## AVALIAÇÃO DE PESQUISAS: TRÊS VALIDADES

**Validade** refere-se à “verdade” e à representação correta da informação. As pesquisas podem ser descritas e avaliadas em termos de três tipos de validade: validade de **construto**, validade **interna** e validade **externa**. Cada uma delas nos dá uma perspectiva diferente sobre qualquer pesquisa particular.

## Validade de Construto

**Validade de construto** refere-se à adequação de uma definição operacional de uma variável – a definição de uma variável reflete realmente o verdadeiro significado teórico de uma variável? Muitas variáveis são “construtos” abstratos, tais como auto-estima, credibilidade do locutor ou vadiagem social. A medida da auto-estima precisa ser desenvolvida para avaliar o construto teórico; essa é uma definição operacional da variável. A medida tem validade de construto se medir o construto de auto-estima e não alguma outra variável, como dominância, por exemplo. Da mesma forma, uma manipulação da credibilidade do orador, comparando oradores com credibilidade alta e baixa, deve envolver apenas credibilidade e não alguma outra coisa, como atratividade, por exemplo. Como há muitas formas de manipular e medir essas variáveis, nunca há uma definição operacional perfeita de uma variável. Com o passar do tempo, muitos pesquisadores usarão múltiplos métodos para definir operacionalmente qualquer variável. Continuaremos tratando desse assunto em capítulos posteriores.

## Validade Interna

**Validade interna** refere-se à capacidade para tirar conclusões sobre relações de causa e efeito de nossos dados. Um estudo tem alta validade interna quando podemos fazer inferências fortes de que uma variável causou outra. Vimos que, em geral, é possível fazer inferências causais fortes com maior facilidade quando se usa o método experimental. A validade interna aumenta quando é possível aplicar à pesquisa as considerações de causa e efeito já discutidas. No Capítulo 8, discutiremos a validade interna de vários delineamentos experimentais.

## Validade Externa

Finalmente, **validade externa** de um estudo é o grau em que os resultados podem ser generalizados para outras populações ou situações. É possível replicar os resultados com outras definições operacionais das variáveis, com participantes diferentes, em situações diferentes? Neste capítulo, preocupações relativas à artificialidade do laboratório foram colocadas; essa é uma questão de validade externa. Observe que o objetivo de atingir elevada validade interna algumas vezes pode conflitar com o objetivo da validade externa. Experimentos de campo são uma forma de os pesquisadores tentarem aumentar a validade externa de seus experimentos. A questão da validade externa é complexa e será mais discutida no Capítulo 14.

Nesse momento, você deve estar-se perguntando como os pesquisadores escolhem uma para estudar um problema. Existem muitos métodos disponíveis, cada um deles com vantagens e desvantagens. Os pesquisadores selecionam o método que lhes parece mais adequado para responder às questões formuladas. Nenhum método é inerentemente superior a outro. A opção por um método é feita em função do tipo de problema investigado, dos custos envolvidos, do tempo gasto e de questões relativas aos três tipos de validade. No restante deste livro, muitos métodos específicos serão discutidos, todos úteis em diferentes circunstâncias. De fato, todos são necessários para compreender a grande variedade de comportamentos que interessam aos cientistas do comportamento. A compreensão completa de qualquer problema ou assunto requer o uso de uma variedade de abordagens metodológicas.

## Termos Estudados

---

Causa necessária  
Causa suficiente  
Controle experimental  
Definição operacional  
Experimento de campo  
Método experimental  
Método não experimental (método correlacional)  
Problema da terceira variável  
Randomização  
Relação curvilínea  
Relação linear negativa  
Relação linear positiva  
Validade de construto  
Validade externa  
Validade interna  
Variável  
Variável dependente  
Variável de resposta  
Variável independente  
Variável situacional  
Variável do participante (sujeito)

O delineamento experimental clássico, descrito no Capítulo 8, consiste em distribuir os participantes aleatoriamente pelas condições da variável independente e em medir uma variável dependente. Comparam-se as respostas na medida dependente, para determinar se a variável independente teve efeito. Como todas as outras variáveis são mantidas constantes, as diferenças encontradas na variável dependente devem ser devidas ao efeito da variável independente. Esse delineamento tem alta validade interna – confiamos que a variável independente tenha causado as respostas observadas na variável dependente. O leitor irá encontrar freqüentemente esse delineamento experimental ao examinar pesquisas realizadas nas ciências do comportamento. No entanto, outros delineamentos de pesquisa foram desenvolvidos para resolver problemas de pesquisa especiais. Este capítulo trata de delineamentos quase-experimentais, de delineamentos usados na área de desenvolvimento e de experimentos com sujeito único – abordagens de pesquisa que apresentam desafios interessantes no que diz respeito à determinação da validade interna.

## DELINEAMENTOS QUASE-EXPERIMENTAIS

Como vimos no Capítulo 1, os pesquisadores freqüentemente defrontam com questões de pesquisa aplicada e de avaliação do comportamento. Os delineamentos quase-experimentais surgiram quando da necessidade de realizar pesquisas aplicadas, em situações em que não é possível atingir o mesmo grau de controle que nos delineamentos experimentais propriamente ditos. Os **delineamentos quase-experimentais** tentam atingir um grau de controle próximo aos dos delineamentos experimentais, para inferir que dado tratamento teve o efeito pretendido. Embora os delineamentos quase-experimentais possam ser usados para resolver muitos problemas de pesquisa, é útil considerá-los no contexto das pesquisas de avaliação de programas.

### Avaliação de Programa

No Capítulo 1 encontra-se uma breve descrição da **avaliação de programas**. Trata-se de pesquisas que propõem e implantam programas para atingir algum efeito positivo sobre um grupo de indivíduos. Esses programas podem ser implantados em escolas, situações de trabalho ou mesmo numa comunidade toda. Um exemplo de programa implantado em escolas é o ERAD (Educação para Resistência ao Abuso de Drogas), destinado a reduzir o uso de drogas. É desenvolvido em conjunto com departamentos de polícia locais e tem sido muito utilizado desde que surgiu, no início da década de 80. A avaliação de programas aplica-se a muitas abordagens de pesquisa para investigar programas desse tipo.

Donald Campbell (1969) estimulou o surgimento de uma cultura de avaliação, em que todos esses programas são honestamente avaliados para determinar sua eficácia. Assim, o foco inicial da pesquisa de avaliação foi a “avaliação de resultado”: O programa obteve realmente os resultados positivos que pretendia obter (por exemplo, diminuição do abuso de drogas, aumento de notas, redução de faltas ou diminuição de reincidência)? No entanto, à medida que a área de avaliação de programa desenvolveu-se, desde o artigo original de Campbell em 1969, a pesquisa de avaliação passou a abranger mais do que a avaliação de resultados (Rossi; Freeman; Lipsey, 1999).

Rossi et al. (1999) identificam cinco tipos de avaliações, cada uma das quais procurando responder a uma questão diferente sobre o programa. Eles aparecem na Figura 11.1 como as cinco fases do processo de avaliação. A primeira é a avaliação da necessidade. Estudos de *avaliação da necessidade* questionam se de fato existem problemas merecedores de atenção numa população-alvo. Por exemplo, há abuso de drogas entre crianças e adolescentes na comunidade? Se há, quais são as drogas usadas? Que serviços os sem-teto necessitam com maior urgência? Delinquentes juvenis reincidentes têm problemas pessoais e familiares particulares que devem ser tratados por um programa de intervenção? Uma vez que se tenha estabelecido uma necessidade, é possível planejar programas voltados para ela. Os dados para a avaliação de necessidades podem ser provenientes de levantamentos, entrevistas e dados estatísticos de órgãos de saúde pública, de justiça criminal e outras agências.

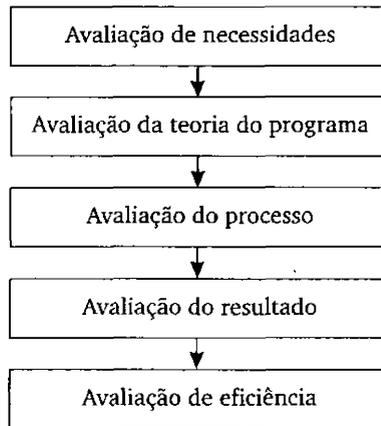


Figura 11.1 Fases da pesquisa de avaliação de um programa.

O segundo tipo de questão de avaliação diz respeito à teoria do programa. Depois de identificar necessidades, pode-se planejar um programa voltado para elas. Rossi et al. (1999) enfatizam que ele deve basear-se em suposições válidas sobre as causas dos problemas e racional do programa proposto. A *avaliação da teoria do programa* pode envolver a colaboração de pesquisadores, de prestadores de serviço e de futuros clientes, para determinar que o programa proposto de fato atende de forma adequada às necessidades da população-alvo. Rossi et al. descrevem um estudo que avaliou a necessidade de moradores de rua de ambos os sexos, na cidade de New York (Herman; Struening; Barrow, 1994). As necessidades gerais mais importantes foram ajuda para encontrar um lugar onde morar, encontrar um emprego e aumentar a qualificação para o trabalho. Os homens em particular necessitaram de ajuda em relação a alcoolismo e uso de drogas, a uso de dinheiro e a convivência. A tendência entre as mulheres foi a de necessitar de ajuda em relação a problemas médicos e de saúde. Um programa destinado a atender a essas necessidades deveria levar em conta essa informação e ter uma racional para justificar o benefício dele resultante para os sem-teto.

O terceiro tipo de questão de avaliação diz respeito à *avaliação do processo* ou monitoramento do programa. Enquanto um programa está sendo implantado, os pesquisadores que fazem a avaliação o monitoram, para determinar se está atingindo a população-alvo, se está atraindo clientes suficientes e se a equipe de profissionais está oferecendo os serviços planejados. A equipe pode não ter recebido treinamento adequado ou os serviços podem estar sendo oferecidos num local indesejável ou de difícil acesso. Em resumo, o pesquisador deseja garantir que o programa está atingindo seus objetivos. Essa pesquisa é extremamente importante, porque não desejamos concluir que um programa é ineficiente, quando de fato o problema reside em sua *implementação*. Essa pesquisa pode envolver questionários e entrevistas, estudos de observação e análise de registros mantidos pela equipe do programa.

A quarta questão diz respeito à *avaliação dos resultados* ou avaliação de impacto: os resultados pretendidos pelo programa estão sendo atingidos? O objetivo – redução do uso de drogas, aumento da alfabetização, diminuição da reincidência por delinqüentes juvenis ou aumento de qualificação para o trabalho – está sendo atingido? Para determinar se está, o pesquisador que faz a avaliação deve elaborar uma forma de medir os resultados e, então, estudar o impacto do programa sobre essa medida. Precisamos saber como estão os participantes do programa e como estariam se não tivessem passado por ele. Idealmente, um delineamento experimental com atribuição aleatória dos participantes às condições experimentais seria planejado para responder a questões sobre resultados. No entanto, outras abordagens de pesquisa, como o delineamento quase-experimental e o delineamento com sujeito único, descritos neste capítulo, são formas muito úteis de avaliar o impacto de um programa de intervenção.

A questão final de avaliação de um programa diz respeito à *avaliação da eficiência*. Tendo demonstrado que um programa tem o efeito pretendido, os pesquisadores devem determinar se ele “se justifica”. Custos e benefícios devem ser ponderados. Além disso, os pesquisadores devem determinar se os recursos usados para implantar o programa poderiam ser mais bem empregados de outra forma.

Como já foi dito, a pesquisa de avaliação pode ser realizada por meio de todas as abordagens discutidas nos capítulos anteriores: levantamentos, observação naturalística e experimentos propriamente ditos. No entanto, frequentemente as situações e as circunstâncias de realização da pesquisa impedem a implementação de um delineamento realmente experimental. Nessas situações deve ser usada uma alternativa a um delineamento experimental. Uma alternativa é o uso de um delineamento quase-experimental. Vamos tratar aqui apenas de alguns dos delineamentos quase-experimentais mais comumente utilizados. Discussões mais detalhadas podem ser encontradas em Campbell (1968, 1969), Cook e Campbell (1979) e Campbell e Stanley (1966). Em seguida, iremos examinar delineamentos de sujeito único que podem ser usados na avaliação de programas e em muitas outras áreas de pesquisa.<sup>1</sup>

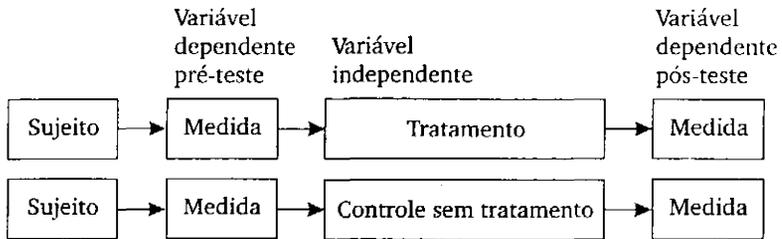
### **Delineamento pré-teste/pós-teste com grupo controle não equivalente**

O delineamento de pós-teste apenas com grupo controle não equivalente foi descrito no Capítulo 8. Esse delineamento tem um grupo experimental e um grupo controle. No entanto, os grupos podem não ser equivalentes, porque os participantes não são aleatoriamente distribuídos pelas condições. Assim, qualquer diferença encontrada na medida dependente no pós-teste pode ser devida a diferenças de seleção. É possível aprimorar muito o delineamento por meio da aplicação de um pré-teste. Procedendo dessa forma, temos um **delineamento pré-teste/pós-teste com grupo controle não equivalente**, que é um dos deli-

1 Há algumas instituições que revisam *studies of interventions*:

- Cochrane Collaboration (<<http://www.cochrane.org>>): prepara, mantém e promove acesso a revisões sistemáticas dos efeitos de intervenções na área da saúde.
- International Campbell Collaboration (C2) (<<http://www.campbellcollaboration.org>>) é uma organização sem fins lucrativos que visa auxiliar as pessoas a tomar decisões bem informadas sobre efeitos de intervenções nos cenários social, comportamental e educacional. Os objetivos da C2 são preparar, manter e disseminar revisões sistemáticas de estudos de intervenção. Obtém e promove acesso a informações a respeito de testes de intervenção. A C2 elabora sumários e brochuras eletrônicas de revisões e relatórios de testes para estrategistas, profissionais liberais, pesquisadores e público em geral (NT).

neamentos quase-experimentais mais úteis. Pode ser diagramado da seguinte forma:



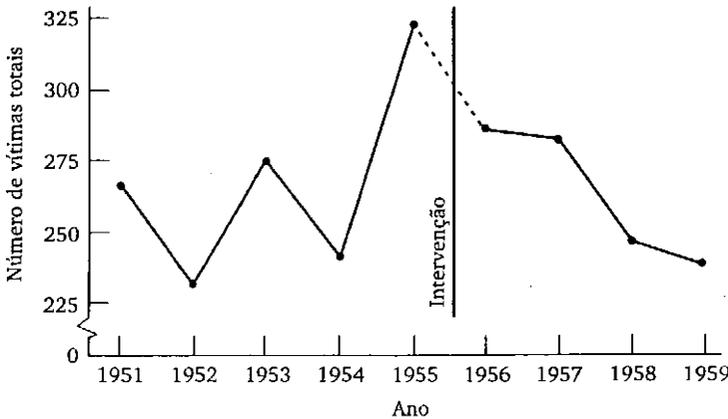
Como a distribuição dos participantes pelos grupos não é aleatória, não se pode dizer que o delineamento seja realmente experimental; é possível que os dois grupos não sejam equivalentes. No entanto, temos a vantagem de conhecer os escores do pré-teste. Assim, podemos verificar se os grupos são equivalentes no pré-teste. Mesmo que não sejam equivalentes, podemos examinar as alterações nos escores do pré-teste para o pós-teste. Se a variável independente tem efeito, o grupo experimental deveria apresentar uma alteração maior que o grupo controle (veja Kenny, 1979).

Joy, Kimball e Zabrack (1986) usaram um delineamento pré-teste/pós-teste com grupo controle não equivalente, para estudar o efeito da televisão sobre o comportamento agressivo de crianças. O foco do estudo foi uma cidade canadense que não tinha televisão até 1974 (chamada "Semtel" por Joy et al.). Os pesquisadores registraram agressões físicas e verbais de crianças, antes e depois da introdução da televisão em Semtel. Simultaneamente registraram a ocorrência de agressão em duas cidades semelhantes, uma das quais recebia apenas uma estação canadense ("Unicanal") e outra que recebia cadeias de emissoras, tanto do Canadá quanto dos Estados Unidos ("Multicanal"). Assim, foi possível comparar a alteração nos níveis de agressão em Semtel com aquela ocorrida nas comunidades de controle Unicanal e Multicanal. Os resultados do estudo mostraram que houve maior aumento de agressão em Semtel que em Unicanal ou Multicanal.

### Delineamento com Sucessões Cronológicas Interrompidas

Campbell (1969) discute detalhadamente a avaliação de uma reforma legal específica: a penalização rigorosa de excessos de velocidade em Connecticut, em 1955. Embora aparentemente o evento tenha acontecido num passado distante,

o exemplo ainda ilustra bem uma questão metodológica importante. A penalização rigorosa foi decidida em virtude do elevado número de acidentes de trânsito com vítimas fatais, ocorrido em 1955. A forma mais simples de avaliar o impacto da mudança na legislação é comparar o número de acidentes de trânsito com vítimas fatais em 1955 (antes das penas rigorosas) e em 1956 (depois). De fato, o número de mortes em decorrência de acidentes de trânsito diminuiu de 324, em 1955, para 284, em 1956. Essa comparação isolada constitui um delineamento pré-teste/pós-teste de um grupo, com todos os problemas de validade interna desse tipo de procedimento; muitas outras razões poderiam explicar o decréscimo das mortes por acidente de trânsito. Uma alternativa é usar um **delineamento de séries temporais interrompidas**, examinando as taxas de acidentes fatais de trânsito ao longo de um período extenso de tempo, antes e depois da mudança na legislação. A Figura 11.2 mostra essa informação para o período de 1951 a 1959. Campbell (1969) argumenta que a queda de 1955 para 1956 não parece particularmente notável, considerando as grandes flutuações ocorridas nos anos anteriores, mas aponta uma tendência constante de decréscimo após a introdução de uma legislação mais rigorosa. Mesmo em relação a isso, no entanto, Campbell vê um problema de interpretação. O decréscimo poderia ser devido a regressão estatística: como 1955 foi um ano de pico, é provável que houvesse uma queda na ausência de qualquer interferência. No entanto,



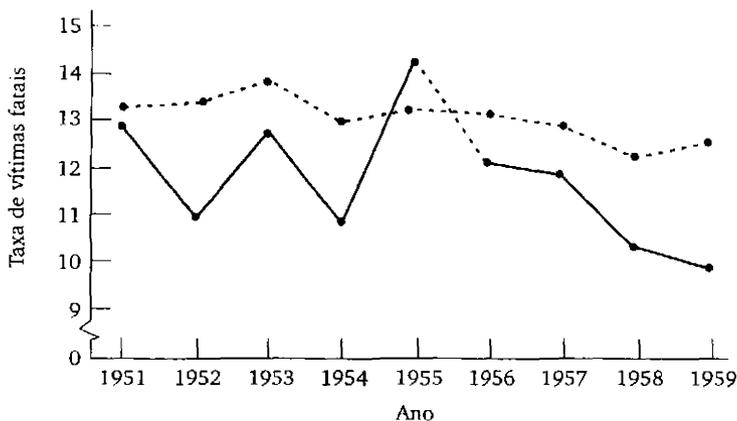
Fonte: CAMPBELL, D. T. Reformas como experimentos. *American Psychologist*, 24, p. 409-429, 1969. Direitos autorais da Associação de Psicólogos Americanos, 1969. Reproduzida mediante permissão.

Figura 11.2 Acidentes de trânsito com vítimas fatais em Connecticut: 1951-1959.

a consideração de um período de tempo mais extenso, incluindo tanto anos anteriores quanto posteriores à mudança na legislação, permite uma interpretação menos ambígua do que seria possível dispondo-se unicamente de dados relativos a 1955 e a 1956.

### Delineamento com Sucessões Controle

Uma forma de aprimorar o delineamento de *sucessões* interrompidas é encontrar algum tipo de grupo controle – um **delineamento com sucessões controle**. No exemplo de Connecticut isso foi possível porque outros Estados não aumentaram o rigor da pena para excesso de velocidade. A Figura 11.3 mostra comparativamente os dados de vítimas fatais em acidentes de trânsito em Connecticut e em quatro Estados, durante os mesmos anos. Campell concluiu que a mudança na legislação realmente teve algum efeito, porque as taxas de acidentes com vítimas fatais nos Estados de controle permaneceram relativamente constantes, enquanto em Connecticut diminuíram consistentemente.



Fonte: CAMPBELL, D. T. Reformas como experimentos. *American Psychologist*, 24, p. 409-429, 1969. Direitos autorais da Associação de Psicólogos Americanos 1969. Reproduzida mediante permissão.

Figura 11.3 *Delineamento com séries controle comparando a taxa de vítimas fatais em acidentes de trânsito ocorridos em Connecticut (linha cheia) com aquela de acidentes de trânsito ocorridos em quatro Estados comparáveis (linha interrompida).*

E a avaliação do programa ERAD? Muitos pesquisadores realizaram estudos para avaliação de resultados, usando delineamentos quase-experimentais, para examinar efeitos tanto a curto quanto a longo prazos. A maioria dos estudos comparou estudantes de escolas vinculadas ou não ao programa ERAD. A conclusão geral é que o ERAD tem pouco efeito sobre os participantes (Ennett; Tobler; Ringwalt; Flewelling, 1994). Além disso, estudos que examinaram efeitos a longo prazo concluíram que o ERAD não traz benefícios duradouros (Rosenbaum; Hanson, 1998). Não obstante isso, o programa é muito popular entre professores e diretores de escolas (Donnermeyer; Wurschmidt, 1997).

## EXPERIMENTOS COM SUJEITO ÚNICO

**Experimentos com sujeito único**, algumas vezes denominados delineamentos  $N = 1$ , costumam ser realizados no contexto de pesquisa sobre reforçamento. Essa tradição de pesquisa tem sua origem no trabalho de B. F. Skinner (1953) sobre esquemas de reforçamento e é usada frequentemente em contextos aplicados e clínicos, em que se empregam técnicas de modificação de comportamento. No entanto, as técnicas e a lógica dos experimentos com sujeito único podem ser facilmente aplicadas também a outras áreas de pesquisa.

Os delineamentos de sujeito único desenvolveram-se da necessidade de determinar se uma manipulação experimental teve efeito num participante individual de uma pesquisa. Num delineamento de sujeito único, registra-se o comportamento do sujeito durante um período de controle **de linha de base**. Introduce-se a manipulação durante um período de tratamento em que o comportamento do sujeito continua sendo observado. Uma mudança no comportamento do sujeito da linha de base para os períodos de tratamento indica a eficácia da manipulação. O problema, no entanto, é que pode haver muitas outras razões para a mudança além do tratamento experimental. Por exemplo, algum outro evento pode ter coincidido com a introdução do tratamento. Os delineamentos de sujeito único, descritos a seguir, tratam desse problema.

### Delineamentos com Reversão

Nos experimentos com sujeito único, a questão básica é como garantir que a manipulação da variável independente teve algum efeito. Um método usado é demonstrar a reversibilidade da manipulação. Um **delineamento com reversão simples** assume a seguinte forma:

A (linha de base) → B (tratamento) → A (linha de base)

Esse delineamento, denominado delineamento ABA, requer que o comportamento seja observado durante o período de controle de linha de base (A), durante o período de tratamento (B) e também durante um segundo período de linha de base (A), em que não se usa o tratamento experimental. (Algumas vezes usa-se a denominação *delineamento de retirada*, salientando o fato de o tratamento ser removido ou retirado). Por exemplo, pode-se determinar o efeito de um método de reforçamento sobre o desempenho acadêmico de uma criança com delineamento ABA. Pode-se determinar diariamente, durante a linha de base, o número de problemas corretamente resolvidos em casa. Introduce-se, em seguida, um período de reforçamento, em que a criança ganha estrelas pela solução correta de problemas e pode trocar as estrelas acumuladas por brinquedos ou doces. Posteriormente, durante o segundo período de linha de base (A), o tratamento é interrompido. A Figura 11.4 mostra dados hipotéticos de um experimento desse tipo. O fato de o comportamento mudar com a introdução do tratamento e de ser revertido com sua interrupção indica a eficácia do tratamento.

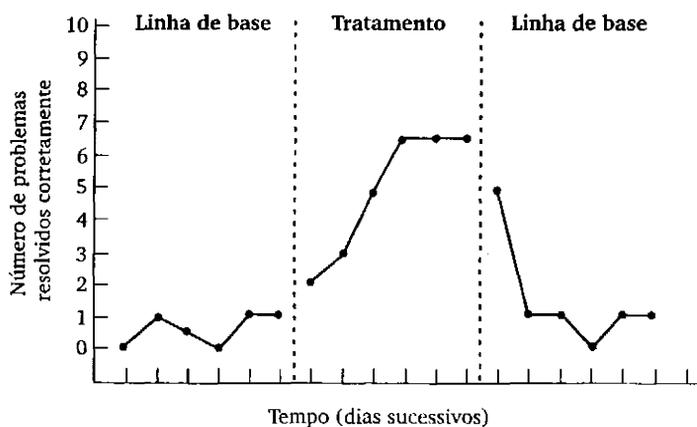


Figura 11.4 Dados hipotéticos de um delineamento com reversão ABA.

O tratamento apresentado na Figura 11.4 teve um impacto relativamente dramático sobre o comportamento. Alguns tratamentos produzem uma mudança imediata no comportamento, mas muitas outras variáveis podem demandar mais tempo para produzir efeito. Segundo Dermer e Hoch (1999), os delineamentos com sujeito único são apropriados no caso dessas variáveis, que envolvem um período longo de tratamento.

Pode-se aperfeiçoar muito o delineamento ABA transformando-o num delineamento ABAB, em que se introduz o tratamento experimental pela segunda

vez, ou mesmo num delineamento ABABAB, que permite testar o efeito do tratamento pela terceira vez. O delineamento de reversão ABA trata de dois problemas. Primeiro, uma só reversão não representa evidência muito forte da eficácia do tratamento. A reversão observada poderia ter sido devida a uma flutuação casual no comportamento da criança. Ou talvez o tratamento tenha coincidido por acaso com algum outro evento, como o aniversário da criança, que causou a mudança (e a reversão, no dia posterior ao aniversário). Essas possibilidades tornam-se bem menos prováveis quando o tratamento se revela eficaz duas vezes ou mais; é pouco provável que ambas as reversões sejam devidas a coincidências ou a eventos aleatórios. O segundo problema é ético. Como apontam Barlow e Hersen (1984), não parece correto interromper o delineamento com a retirada de um tratamento, que pode ser muito benéfico para o participante. O delineamento ABAB permite observar uma segunda reversão, quando o tratamento é reintroduzido. A seqüência termina com o tratamento e não com sua remoção.

Num delineamento que envolve reversão, também é possível usar um grupo controle. Por exemplo, um grupo de pesquisadores usou reforçamento para aumentar o uso de ônibus num *campus* universitário (Everett; Hayward; Meyers, 1974). Contou-se o número de passageiros em dois ônibus diferentes, durante 36 dias. A manipulação experimental consistiu em dar uma ficha a cada passageiro em um dos ônibus especialmente marcados; as fichas podiam ser trocadas por mercadorias e serviços no comércio local. No ônibus de controle os passageiros não recebiam fichas. Usava-se um delineamento ABA. Os primeiros 16 dias foram um período de linha de base e, durante essa fase, o número de passageiros por ônibus foi de aproximadamente 250 pessoas por dia. A manipulação experimental foi introduzida nos dias 17-24. Durante esse período, o número de passageiros no ônibus experimental aumentou para aproximadamente 400 por dia; não houve alteração do número de passageiros no ônibus de controle. O sistema de fichas foi interrompido no dia 25 e o número de passageiros foi monitorado nos dois tipos de ônibus até o dia 36. Durante esse período, o número de passageiros no ônibus experimental voltou aos níveis de linha de base, não se diferenciando mais daquele do ônibus de controle.

### **Delineamentos com Linha de Base Múltipla**

A reversão de alguns comportamentos pode ser impossível ou antiética. Por exemplo, não é ético reverter um tratamento que reduz comportamentos perigosos ou ilegais, como exibicionismo ou alcoolismo, mesmo que a reintrodução do tratamento possa resultar em outra mudança. Outros tratamentos podem produzir uma mudança comportamental duradoura que não é reversível. Nesses casos, são feitas medidas múltiplas no tempo, antes e depois da manipulação. Se a manipulação for efetiva, uma mudança comportamental será imedia-

tamente observada e a mudança continuará refletindo-se em outras medidas do comportamento. Num **delineamento com linha de base múltipla**, demonstra-se a eficácia do tratamento quando um comportamento muda apenas após a introdução da manipulação. Para demonstrar a eficácia do tratamento, a mudança deve ser observada sob *múltiplas* circunstâncias, para excluir a possibilidade de que outros eventos sejam responsáveis.

Existem muitas variantes do delineamento com linha de base múltipla (Barlow; Hersen, 1984). Na linha de base múltipla *entre* sujeitos, registra-se o comportamento de vários *sujeitos* ao longo do tempo; para cada sujeito, no entanto, a manipulação é introduzida num momento diferente. A Figura 11.5 mostra dados de um experimento hipotético com 3 sujeitos, para redução do comportamento de fumar. Note que a introdução da manipulação foi seguida de mudança comportamental em cada um deles. No entanto, como essa mudança ocorreu em todos os indivíduos e a manipulação foi introduzida em momentos diferentes para cada um deles, podemos excluir interpretações baseadas em acaso, eventos históricos e assim por diante.

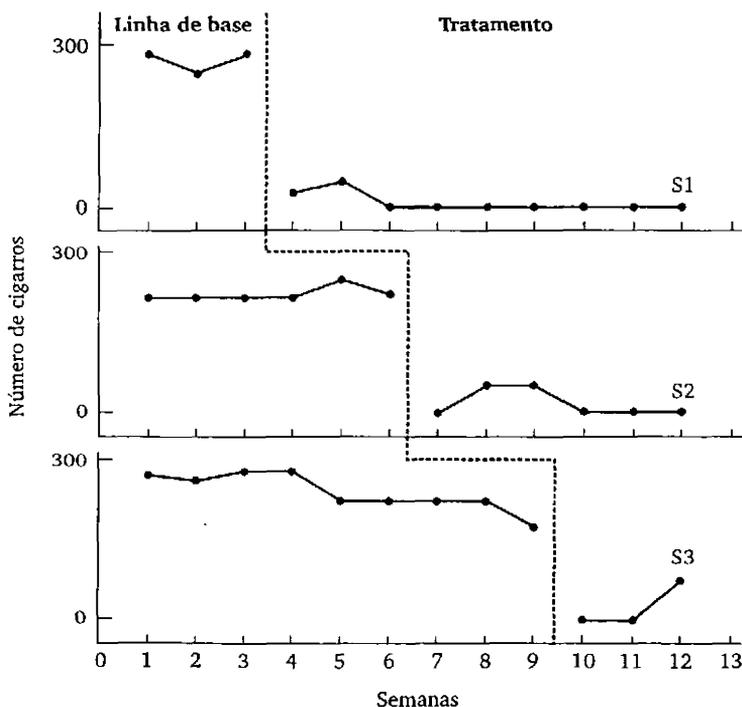


Figura 11.5 *Dados hipotéticos de um delineamento com linha de base múltipla entre sujeitos.*

Numa linha de base múltipla *entre comportamentos*, vários comportamentos diferentes de um único sujeito são registrados ao longo do tempo. Em diferentes momentos, aplica-se a mesma manipulação a cada um dos comportamentos. Por exemplo, introduz-se um sistema de recompensa para aumentar a socialização, a limpeza e o hábito de leitura de um paciente psiquiátrico. O sistema de recompensa é aplicado a cada um desses comportamentos em diferentes momentos. A eficácia da manipulação é indicada pelo aumento de cada comportamento, com a aplicação do sistema de recompensa.

A terceira variante é a linha de base múltipla *entre situações*, em que o mesmo comportamento é registrado em diferentes contextos, em casa e no trabalho, por exemplo. Novamente, uma manipulação é introduzida num momento diferente em cada contexto, com a expectativa de que ocorra uma mudança no comportamento em cada contexto após a manipulação.

### Réplicas em Delineamentos com Sujeito Único

Os procedimentos usados com um sujeito único podem, de fato, ser replicados com outros sujeitos, aumentando muito a possibilidade de generalização dos resultados. Em geral, os relatos de pesquisa que empregam procedimentos experimentais com sujeito único apresentam os resultados de vários sujeitos (e freqüentemente em vários contextos). A tradição na pesquisa que utiliza sujeito único tem sido apresentar os resultados de cada sujeito individualmente, em lugar de agrupar os dados e apresentar médias grupais. Segundo Sidman (1960), um expoente dessa tradição, o agrupamento de dados de vários sujeitos, por meio da média, pode algumas vezes fornecer um quadro enganoso das respostas individuais à manipulação. Por exemplo, a manipulação pode ser eficaz em modificar o comportamento de alguns sujeitos, mas não de outros. Isso aconteceu num estudo sobre o uso do cinto de segurança (Berry; Geller, 1991). Foram estudadas diferentes condições de sinalização (por exemplo, apresentação de um segundo sinal quando o sujeito não afivelava o cinto após um sinal inicial). Entre 13 sujeitos, 6 sempre usaram o cinto, independentemente da condição, e 3 nunca usaram. Para os outros 4 sujeitos, as condições de sinalização tiveram efeito. Como o estudo focalizava os sujeitos individualmente, o padrão de resultados logo foi identificado.

Delineamentos que envolvem sujeito único são úteis para investigar muitos problemas de pesquisa, especialmente úteis quando se aplica alguma técnica de modificação de comportamento num ambiente natural – por exemplo, um professor tenta uma nova técnica em sala de aula. Além disso, não há necessidade de aplicar técnicas estatísticas complexas quando se usam delineamentos com sujeito único.