

CAPÍTULO 4

DESENHO DE PESQUISA

RESUMO:

Sendo nosso foco a causalidade, cabe perguntar: quais são as estratégias de pesquisa que cientistas políticos usam para investigar relações causais? De modo geral, o experimento com controle do pesquisador é a base para a pesquisa científica. Alguns cientistas políticos utilizam experimentos em seus trabalhos. Contudo, devido à natureza do objeto, a maioria dos cientistas políticos adota um dos dois tipos de desenhos de pesquisa “observacional” que visam imitar experimentos: o estudo observacional de corte transversal que estuda a variação entre unidades (como pessoas ou países) e o estudo observacional de corte temporal que permite examinar a variação de quantidades agregadas (como a popularidade presidencial) ao longo do tempo. O que é um “experimento” e por que ele é útil? Como estudos observacionais tentam imitar desenhos experimentais? Mais importante, quais são as forças e fraquezas de cada um desses três desenhos de pesquisa para estabelecer ou não a existência de uma relação causal entre dois conceitos? Isto é, como cada um nos ajuda a superar os quatro obstáculos causais identificados no capítulo 3? Semelhantemente, introduzimos questões concernentes à seleção de amostra de casos para estudos nos quais não somos capazes de analisar toda a população à qual nossa teoria se aplica. Esse é um assunto que terá um lugar de destaque em muitos dos capítulos subsequentes.

4.1 A COMPARAÇÃO COMO A CHAVE PARA ESTABELECEER RELAÇÕES CAUSAIS

Até o momento, você aprendeu que cientistas políticos se preocupam com relações causais. Você aprendeu que a maioria dos fenômenos que estamos interessados

em explicar possuem múltiplas causas, mas que nossa teoria tipicamente lida apenas com uma enquanto ignora as demais. Em alguns exemplos de pesquisas apresentados nos capítulos anteriores, pontuamos que a natureza multivariada do mundo pode nos fazer, em um primeiro momento, entender uma evidência de maneira enganosa. No exemplo sobre raça e participação política, em um primeiro momento parecia que a raça estava causalmente relacionada com as taxas de participação, com caucasianos participando mais do que outras raças. Mas, como argumentamos, esse primeiro olhar era bastante enganoso.

Por quê? Porque o que aparentava ser uma comparação direta entre três grupos (a taxa de participação entre caucasianos, latinos e afro-americanos) não era simples como parecia. Nossos diferentes grupos da variável X estavam longe de serem iguais em importantes fatores. Isto é, pessoas de diferentes grupos raciais (X) tinham diferentes *status* socioeconômicos (Z), o que estava correlacionado com raça (X) e também afetava o nível de participação (Y). Esse exemplo mostra que, por mais que comparações bivariadas possam parecer convincentes, elas provavelmente estão erradas.

As comparações estão no coração da ciência. Se estivermos avaliando uma teoria sobre a relação entre algum X e algum Y , nosso trabalho como cientistas é fazer tudo que for possível para termos certeza de que não existem outras influências (Z) interferindo nas comparações que utilizaremos para fazer nossas inferências sobre uma possível relação causal entre X e Y .

Os obstáculos à inferência causal que descrevemos no capítulo 3 são substanciais, mas não intransponíveis. Não sabemos, na realidade, se X causa Y . Podemos utilizar uma teoria que sugere que X , efetivamente, causa Y , mas teorias podem estar (e frequentemente estão) erradas ou incompletas. Então como cientistas em geral, e cientistas políticos em particular, testam se X causa Y ? Existem várias estratégias, ou **desenhos de pesquisa**, que pesquisadores podem utilizar para testar teorias. O objetivo de todos os tipos de desenhos de pesquisa é nos ajudar a avaliar quão bem uma teoria se sai perante os quatro obstáculos causais – isto é, responder do modo mais conclusivo possível se X causa Y . Nas próximas duas seções, focamos duas das mais comuns e efetivas estratégias usadas por cientistas políticos: **experimentos** e **estudos observacionais**¹.

4.2 DESENHOS DE PESQUISA EXPERIMENTAIS

Suponha que você seja candidato a um cargo político e aparente estar no caminho certo. Seu orçamento de campanha tem dinheiro suficiente para o restante da campanha e você está decidindo se deve comprar algum espaço de propaganda na televisão para exibir um comercial que contraste o seu histórico com o do seu concorrente – um ato que alguns considerarão como uma propaganda negativa e de ataque ao oponente. O coordenador da campanha contratou uma empresa de relações públicas para

¹ Neste livro utilizaremos o termo “experimento” da mesma forma que os pesquisadores da área de saúde utilizam o termo “ensaio clínico aleatório”.

trabalhar no comercial e mostrou o resultado para você em uma das reuniões de estratégia. Você gostou, mas quer saber a opinião da sua equipe e pergunta diretamente: “O comercial funcionará com o eleitorado?”. No final, você tem duas opções: exibir o comercial de ataque ou não fazer nada.

Esperamos que você esteja se acostumando a detectar a questão causal embutida em um cenário como o supracitado. A exposição a uma propaganda com conteúdo negativo sobre um candidato (X) pode, ou não, afetar a probabilidade do eleitor votar no candidato-alvo da propaganda (Y). É importante salientar que a afirmação causal possui um componente direcional específico; isto é, a exposição a propagandas aumentará as chances de o eleitor escolher o candidato que fez a propaganda².

Como podemos avaliar tal afirmação causal? Os viciados em campanhas políticas estão provavelmente pensando que a campanha poderia realizar um grupo focal para observar como alguns eleitores reagem ao anúncio. Essa não é uma má ideia. Deixemos definir informalmente um grupo focal como um grupo de indivíduos selecionado que é exposto a uma ideia (como uma nova faca de cozinha ou uma propaganda eleitoral de TV), em que se tenta detectar a maneira como os indivíduos respondem à ideia. Porém, existe um problema com o grupo focal, especialmente no caso do anúncio de TV do candidato: o que os indivíduos teriam dito sobre o candidato caso eles *não* fossem expostos no anúncio? Não existe nada que possa ser usado para comparação.

É muito importante, e nem um pouco surpreendente, pensar que eleitores podem votar a favor ou contra você por uma variedade de razões (Z s) que não tem a ver com a exposição à propaganda – *status* socioeconômico, ideologia e afiliação partidária podem ser causas para um eleitor votar a favor de um candidato ou de outro. Portanto, como podemos estabelecer se, entre outras influências (Z), o anúncio (X) também faz com que os eleitores estejam mais propícios a votar em você (Y)?

Podemos fazer algo melhor que utilizar um grupo focal para estudar o problema? Qual seria uma abordagem mais científica? Como a introdução deste capítulo tenta mostrar, precisamos de algum tipo de comparação e queremos utilizar tal comparação para mensurar qualquer efeito potencial diferente que o anúncio tem sobre a chance de uma pessoa votar em você de forma isolada.

A abordagem-padrão para uma situação como essa nas ciências físicas e médicas é que necessitaríamos conduzir um experimento. Como a palavra “experimento” possui amplo uso na linguagem convencional, seu significado científico é frequentemente mal entendido. Um experimento *não* é simplesmente qualquer tipo de análise de natureza quantitativa; nem é exclusivamente algo que se realiza dentro de um laboratório por cientistas vestidos de jalecos brancos com proteções. Nossa definição de experimento é: *um experimento é um desenho de pesquisa no qual o pesquisador tem tanto o controle quanto a capacidade de atribuir randomicamente valores da variável independente aos participantes.*

² Existe uma substancial literatura na ciência política sobre o efeito que propagandas negativas têm tanto sobre o comparecimento eleitoral como sobre o voto. Para visões contrastantes sobre o efeito de propagandas negativas, ver Ansolabehere e Iyengar (1997), Wattenberg e Brian (1999) e Geer (2006).

Note os componentes gêmeos da definição do experimento: o pesquisador tem tanto o controle dos valores da variável independente – ou *X*, como temos chamado – quanto a capacidade de atribuir *randomicamente* esses valores aos participantes do experimento. Juntas, essas duas características formam uma definição completa de um experimento, o que significa que não existem outras características essenciais em um experimento além dessas duas.

O que significa dizer que um pesquisador tem o “controle” do valor da variável independente que os participantes recebem? Significa que os valores da variável independente que os participantes recebem *não são* determinados pelos participantes ou pela natureza. No nosso exemplo do anúncio de campanha de TV, esse requisito significa que não podemos comparar pessoas que, por escolha própria, já tenham sido expostas ao anúncio de TV (talvez porque sejam aficionadas de política e assistam a muitos programas de notícias na televisão, em que tais anúncios são bastante veiculados). Significa que nós, os pesquisadores, temos que decidir quais dos participantes de nosso experimento assistirão e quais não assistirão ao anúncio.

Mas a definição de um experimento tem outro componente essencial. Nós, os pesquisadores, devemos não apenas controlar o valor da nossa variável independente, mas *devemos também atribuir esses valores aos participantes randomicamente*. No nosso exemplo do anúncio de campanha, isso significa que devemos jogar uma moeda para cima, retirar um número de um chapéu, usar um gerador de números randômicos ou algum outro mecanismo desse tipo para dividir nossos participantes em um **grupo de tratamento** (os que assistirão ao anúncio negativo) e um **grupo de controle** (os que não assistirão ao anúncio, mas algo diferente do tratamento, que nas ciências médicas é chamado de um **placebo**).

Qual é o problema aqui? Por que a atribuição randômica dos indivíduos ao grupo de tratamento é importante? Quais benefícios científicos surgem da atribuição randômica de pessoas ao grupo de tratamento? Para ver por que isso é tão crucial, lembre-se que temos enfatizado que toda ciência se refere a comparações e que, também, todo fenômeno interessante que merece ser explorado (toda variável independente interessante) é causado por muitos fatores, não apenas um. A atribuição randômica do grupo de tratamento assegura que a comparação que fazemos entre o grupo de tratamento e o grupo de controle é tão pura quanto possível e que outras causas (*Z*) da variável dependente não poluirão a comparação. Ao, inicialmente, pegarmos um grupo de participantes e, então, separá-lo *randomicamente* em dois grupos, utilizando o lançamento de uma moeda como critério, o que estamos assegurando é que os participantes não serão sistematicamente diferentes uns dos outros. De fato, desde que o número de participantes seja razoavelmente grande, a atribuição aleatória de participantes ao grupo de tratamento assegura que os grupos, no seu todo, sejam *idênticos*. Se os dois grupos são idênticos, exceto pelo lançamento da moeda, então podemos ter certeza de que qualquer diferença que observarmos nos grupos deve se dar em razão da variável independente que atribuímos.

Retornando ao nosso exemplo da propaganda de uma campanha política. Um experimento sobre nosso novo anúncio envolveria encontrar um grupo de pessoas e,

então, atribuir randomicamente se elas assistirão ao nosso novo anúncio ou algo não relacionado à campanha (como desenhos ou anúncios de serviços públicos). Estamos totalmente convencidos de que existem outras causas para o comportamento eleitoral e que nosso experimento não nega esses fatores. De fato, nosso experimento nada terá a dizer sobre essas outras causas. O que ele fará, e bem, é determinar se nosso anúncio tem um efeito positivo ou negativo, ou nenhum, sobre as preferências dos eleitores.

Agora contraste a comparação que resulta de um experimento com uma que emerge de um não experimento. Discutiremos desenhos de não experimentos na próxima seção. Suponha que não façamos um experimento; que façamos apenas o anúncio e então gastemos o dinheiro da nossa campanha conduzindo um *survey* para perguntar às pessoas se elas viram nosso anúncio e em quem elas planejam votar. Assumimos também que, no *survey*, utilizamos uma amostra aleatória dos cidadãos do distrito onde a eleição acontecerá. Se analisarmos os resultados do *survey* e descobirmos que, como esperado, pessoas que disseram ter assistido ao anúncio estão mais propensas a votar em nós do que pessoas que não viram o anúncio, isso significa que o anúncio *causa* – viu a palavra novamente? – uma mudança na opinião das pessoas a nosso favor? Não. Por que não? Porque as pessoas que viram nosso anúncio e as pessoas que não assistiram podem ser *sistematicamente diferentes* uma das outras. O que isso significa? Significa que as pessoas que voluntariamente assistem mais política na TV são (obviamente) mais interessadas em política do que as que preferem assistir aos demais programas da TV. Nesse caso, o nível de interesse em política de uma pessoa pode ser uma importante variável *Z*. Interesse em política poderia muito bem estar associado com a probabilidade de uma pessoa votar em você. Isso significa que a simples comparação em um não experimento entre os que assistem e os que não assistem a um anúncio é potencialmente enganosa porque se deixa confundir por outros fatores, como o interesse em política. Então um maior apoio é resultado do anúncio ou do fato de que pessoas que assistem ao anúncio são pessoas com alto interesse em política?

Como esse particular desenho de pesquisa não experimental não responde a essa pergunta, não é claro que superamos o quarto obstáculo causal. É impossível saber se foi o anúncio que causou o apoio dos eleitores a você. Em um desenho não experimental como o descrito, por existirem outros fatores que influenciam o apoio a um candidato e, criticamente, por esses outros fatores também serem relacionados ao fato de as pessoas assistirem ou não à propaganda, é muito difícil dizer conclusivamente que a variável independente (exposição ao anúncio) causa a variável dependente (intensão de voto). A Figura 4.1 mostra isso graficamente.

É neste ponto que experimentos diferem drasticamente de qualquer outro tipo de desenho de pesquisa. O desenho de pesquisa experimental realiza, por meio da atribuição randômica ao grupo de tratamento, a descontaminação da comparação entre o tratamento e o grupo de controle removendo qualquer outra influência. Antes que qualquer estímulo (como o tratamento ou o placebo) seja administrado, todos os participantes pertencem ao mesmo grupo. Os pesquisadores os dividem utilizando algum fator randômico, como o lançamento de uma moeda para cima, e essa é a única coisa que divide os dois grupos.

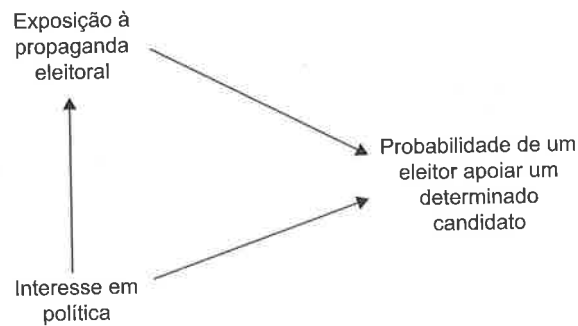


Figura 4.1 – Os possíveis efeitos de confusão do interesse político na relação entre exposição à propaganda eleitoral e intenção de voto.

Pense de outra forma. O modo como as variáveis colineares na Figura 4.1 são correlacionadas com a variável independente é altamente improvável em um experimento. Por quê? Porque se X é determinado aleatoriamente, como pelo lançamento de uma moeda, então (pela definição de randômico) é extremamente improvável que seja correlacionado com qualquer outra coisa (incluindo variáveis colineares Z). Quando pesquisadores controlam e atribuem os valores de X randomicamente, a comparação entre os diferentes grupos não será afetada pelo fato de que outros fatores certamente causam Y , a variável dependente. Em um experimento, então, como X é causado unicamente pela aleatoriedade, significa que podemos eliminar a conexão entre Z e X da Figura 4.1. E, recordando nossa definição de variável colinear, se Z não é correlacionada com X , ela não gera ruído na mensuração da relação entre X e Y .

Isso nos conecta à nossa discussão do capítulo 3 sobre como pesquisadores tentam superar os quatro obstáculos em seus esforços de estabelecer se algum X causa Y . Como veremos, experimentos não são o único método que ajudam pesquisadores a superar os quatro obstáculos causais, mas eles são os únicos capazes de realizar algumas importantes partes dessa tarefa. Considere cada um dos obstáculos por vez. Primeiro, devemos avaliar se existe um mecanismo causal crível antes de decidirmos realizar um experimento. É preciso observar que superar esse obstáculo não é nem mais fácil, nem mais difícil em experimentos do que em não experimentos. Ter um cenário causal crível que liga X a Y aumenta nossa confiança na teoria, não nos dados ou no desenho de pesquisa.

Segundo, em um experimento, é impossível para um Y causar X – o segundo obstáculo causal – por duas razões. Primeiro, a atribuição de X ocorre antes que Y seja mensurado, o que torna impossível para Y causar X . Mais importante, porém, como notado anteriormente, se X é gerado apenas pela aleatoriedade, então nada (incluindo Y) pode causá-lo. Então, na Figura 4.1, poderíamos eliminar qualquer possível causalidade reversa (Y causa X).

Terceiro, estabelecer se X e Y estão correlacionados é simples independentemente do tipo de desenho de pesquisa escolhido (como veremos no capítulo 7). E nosso quarto obstáculo causal? Controlamos por todas as variáveis colineares Z que podem

tornar a associação entre X e Y espúria? Experimentos são desenhos de pesquisa mais adequados para nos ajudar a responder a essa questão de uma maneira definitiva. Um experimento não elimina, de modo algum, a possibilidade de uma variedade de outras variáveis (que podemos chamar de Z) também afetar Y (como também X). O que um experimento faz, por meio do processo de atribuição randômica ao grupo de tratamento, é equiparar o grupo de tratamento e o grupo de controle em todos os fatores possíveis. Em toda variável possível, se ela é ou não relacionada com X , com Y , com ambas ou com nenhuma, os grupos de tratamento e de controle devem, em teoria, ser idênticos. Isso faz a comparação entre dois valores de X limpa de qualquer possível variável Z , porque esperamos que os grupos sejam equivalentes para todos os valores de Z .

Notavelmente, a habilidade dos experimentos em controlar os efeitos de outras variáveis (Z) aplica-se a todas as possíveis variáveis colineares, mesmo quando os pesquisadores não estão conscientes delas. Deixe-nos apresentar um exemplo claramente absurdo. Vamos imaginar que, daqui a vinte anos, outro time de cientistas descubra que os lóbulos da orelha (grudados ou desgrudados) causam diferentes tipos de comportamento eleitoral. Isso ameaça a inferência a que chegamos a partir do nosso experimento com o anúncio de campanha? Não, nem um pouco. Por que não? Porque, estejamos conscientes disso ou não, nossa atribuição randômica dos participantes ao grupo de tratamento significa que, prestando atenção ou não a isso, esperaríamos que nossos grupos de tratamento e controle tivessem igual número de pessoas com lóbulos grudados e desgrudados. O elemento-chave do desenho de pesquisa experimental – a atribuição randômica aos indivíduos de diferentes valores de X , a variável independente – controla todos os Z no universo, estejamos ou não conscientes deles.

Em resumo, se pensamos nos nossos quatro obstáculos causais do capítulo anterior, toda preparação adequada de um experimento começa com o *scoreboard* [$? s ? s$]. A habilidade do desenho de pesquisa experimental de responder de maneira limpa e definitiva ao quarto obstáculo causal com um “sim” – controlamos por todas as variáveis colineares Z que podem fazer a associação entre X e Y espúria? – é uma vantagem massiva³. Tudo o que resta para o estabelecimento de uma relação causal é responder de maneira clara ao primeiro obstáculo – existe algum mecanismo crível que conecta X a Y ? – e ao terceiro obstáculo – existe covariação entre X e Y ? A dificuldade em superar o primeiro obstáculo é inalterável, mas o terceiro obstáculo é muito mais fácil de ser superado porque precisamos apenas fazer uma avaliação estatística da relação entre X e Y . Como veremos no capítulo 7, tais avaliações são bastante diretas, especialmente quando comparadas a testes estatísticos que envolvem controlar outras variáveis (Z).

Conjuntamente, isso tudo significa que experimentos trazem consigo uma confiança particularmente forte nas inferências causais feitas a partir de sua análise. Em linguagem

³ Afinal, mesmo o mais bem desenhado e executado desenho de pesquisa não experimental deve manter aberta a possibilidade de que, em algum lugar, existe uma variável Z que não foi considerada e controlada.

científica, isso é chamado de **validade interna**. Se um desenho de pesquisa produz altos níveis de confiança em suas conclusões sobre causalidade, é dito que ele tem alta validade interna. Contrariamente, dizemos que desenhos de pesquisa que não permitem extrair conclusões definitivas sobre *X* causar *Y* possuem baixo grau de validade interna.

4.2.1 “ATRIBUIÇÃO RANDÔMICA” VERSUS “AMOSTRA ALEATÓRIA”

É fundamental que você não confunda o processo de atribuição randômica dos casos ao grupo de tratamento em um experimento com o processo de seleção aleatória de casos. Eles são totalmente diferentes e, de fato, a única coisa que possuem em comum é o princípio da aleatoriedade no nome do termo. São, contudo, frequentemente confundidos. **Atribuição randômica** aos grupos de tratamento e controle ocorre quando os participantes de um experimento são atribuídos randomicamente a diferentes valores de *X*, a variável dependente. É importante observar que essa definição não diz nada a respeito de como os indivíduos são selecionados para participar do experimento. Já a **amostra aleatória**, essencialmente, relaciona-se a como pesquisadores selecionam casos para os seus estudos – eles são selecionados aleatoriamente, o que significa que cada membro da **população** em questão tem a mesma probabilidade de ser selecionado. (Esse é um método comum de seleção em pesquisas de *survey*, por exemplo.)

Confundir esses dois conceitos críticos produzirá uma boa dose de confusão. Em particular, confundir amostra aleatória com atribuição randômica ao grupo de tratamento significará que a distinção entre experimentos e não experimentos terá se perdido, e essa diferença está entre as mais importantes de toda a ciência. Para entender como a ciência funciona, mantenha esses dois importantíssimos conceitos separados.

4.2.2 VARIEDADES DE EXPERIMENTOS E QUASE EXPERIMENTOS

Nem todos os experimentos acontecem em laboratórios com cientistas vestindo jalecos brancos. Alguns experimentos em ciências sociais são conduzidos utilizando *surveys* que usam amostras aleatórias (ver acima detalhes sobre a amostragem aleatória). Desde os anos 1990, tem ocorrido um crescente movimento no campo da pesquisa de *survey* – que tradicionalmente utiliza amostragem aleatória da população – de utilização de computadores no processo de entrevista para incluir a randomização experimental de variações das perguntas do *survey*, uma técnica chamada **experimento de survey**. Esses projetos tentam aproveitar os benefícios da atribuição aleatória ao grupo tratamento e, dessa forma, aumentar a validade interna e os benefícios da amostragem aleatória, aumentando a **validade externa**⁴. Experimentos de *survey* podem ser conduzidos por telefone ou pela internet (meio cada vez mais utilizado).

Outra configuração possível para um experimento é realizá-lo no mundo real. Um **experimento de campo** ocorre nas configurações naturais em que os indivíduos

⁴ Ver Piazza, Sniderman e Tetlock (1990) e Sniderman e Piazza (1993).

normalmente vivem suas vidas. A atribuição aleatória ao grupo de tratamento tem permitido aos pesquisadores em ciências sociais estudar indivíduos que inicialmente não poderiam ser estudados por meio de experimentos. Economistas têm, há muito tempo, buscado evidências sobre a efetividade (ou a falta dela) das políticas de desenvolvimento econômico. Por exemplo, subsídios do governo para fertilizantes (X) afetam o produto da agricultura (Y)? Com o objetivo de responder a essa pergunta, Duflo, Kremer e Robinson (2011) realizaram um experimento em uma região do oeste do Quênia no qual o subsídio de entrega gratuita de fertilizantes foi oferecido somente a alguns fazendeiros escolhidos randomicamente.

Experimentos de campo podem também acontecer no contexto de políticas públicas, algumas vezes com compreensíveis controvérsias. A decisão policial de prender ou não um homem acusado de violência doméstica (X) afeta a chance de casos de violência se repetirem no mesmo endereço nos meses subsequentes (Y)? Sherman e Berk (1984) conduziram um experimento de campo em Minneapolis, randomizando se o homem seria preso automaticamente (ou não) quando a polícia chegasse à casa.

Em algumas ocasiões, situações na natureza que não são propriamente definidas como experimentos – porque os valores de X não foram controlados e atribuídos pelo pesquisador – são, entretanto, semelhantes a experimentos em características-chave. Em um **experimento natural** – que, enfatizamos, não atende à nossa definição de experimento –, os valores da variável independente são naturalmente dados de maneira que pareça que uma verdadeira atribuição aleatória tenha ocorrido. Por exemplo, o tamanho de um grupo étnico dentro de uma população (X) afeta o conflito intergrupos ou a cooperação (Y)? Posner (2004) investigou por que os povos Chewa e Tumbuka são aliados na Zâmbia, mas adversários no Malawi. O autor desenhou a pesquisa dessa forma porque o tamanho da população dos grupos nos diferentes países tinha características da atribuição randômica, e a comparação foi feita *como se* o tamanho das respectivas populações fosse randomicamente atribuído pelo pesquisador, quando (obviamente) não era esse o caso.

4.2.3 EXISTEM LIMITAÇÕES AO DESENHO DE PESQUISA EXPERIMENTAL?

Experimentos, como vimos, têm a habilidade única de fazer cientistas sociais superarem os nossos quatro obstáculos para estabelecer se X causa Y . Mas isso não significa que eles não tenham desvantagens. Muitas dessas desvantagens estão relacionadas às diferenças entre as ciências física e médica, por um lado, e as ciências sociais, por outro. Agora discutiremos quatro desvantagens dos experimentos.

Primeiro, especialmente nas ciências sociais, nem toda variável independente (X) é controlável e sujeita a manipulação experimental. Suponha, por exemplo, que desejamos estudar os efeitos dos diferentes gêneros na participação política. Homens contribuem com mais dinheiro, votam mais, se voluntariam mais em campanhas do que mulheres? Existe uma variedade de modos não experimentais de estudar essa relação,

mas é impossível manipular experimentalmente o gênero do indivíduo. Lembre-se que a definição de um experimento é que o pesquisador tenha controle, atribuindo randomicamente os valores da variável independente. Nesse caso, a causa presumível (a variável independente) é o gênero da pessoa. Comparado com drogas e placebos, a atribuição do gênero de um participante é um assunto inteiramente diferente. Essa manipulação é impossível. Pessoas que participam de um experimento são homens ou mulheres; não está entre as capacidades do experimento atribuir se um participante é homem ou mulher.

Isso é verdade em muitos e muitos exemplos de ciência política. Existe simplesmente uma miríade de problemas substantivos que são impossíveis de serem estudados de uma maneira experimental. Como a preferência partidária de um indivíduo (X) afeta sua opinião (Y)? Como a renda de uma pessoa (X) afeta o quanto ela contribui para campanhas (Y)? Como o nível de democratização de um país (X) afeta sua abertura para o comércio internacional (Y)? Como o nível de gasto militar na Índia (X) afeta o nível do gasto militar no Paquistão (Y), e vice-versa? Como a cobertura da mídia (X) de uma campanha eleitoral afeta as prioridades do eleitor (Y)? Ocupar uma cadeira no parlamento britânico (X) faz o parlamentar mais rico (Y)? Em cada um desses exemplos que intrigam os cientistas sociais, as variáveis independentes são simplesmente impossíveis de serem manipuladas experimentalmente. Cientistas sociais não podem "atribuir" a pessoas uma identificação partidária ou uma renda, "atribuir" a um país um nível de democratização ou de gasto militar, "atribuir" a especificidade de uma campanha com maior cobertura da mídia, ou "atribuir" diferentes ocupantes de cadeiras no parlamento. Essas variáveis simplesmente existem na natureza e não podemos controlar a exposição a elas e atribuir randomicamente diferentes valores para diferentes casos (isto é, pessoas ou países). E, no entanto, cientistas sociais se sentem compelidos a estudar esses fenômenos, o que significa que, nessas circunstâncias, devemos empregar um desenho não experimental de pesquisa.

Uma segunda desvantagem potencial do desenho de pesquisa experimental é que experimentos frequentemente sofrem por ter baixo grau de validade externa. Temos notado que uma das principais forças dos experimentos é que eles tipicamente possuem altos níveis de validade interna. Isto é, podemos estar bastante confiantes de que as conclusões sobre causalidade encontradas na análise não sofrem a influência de outras variáveis. Validade externa, neste sentido, é o outro lado da moeda; ela representa o grau em que podemos estar confiantes de que os resultados da nossa análise se aplicam não somente aos participantes do estudo, mas também a toda população.

Existem dois tipos de preocupações reais com respeito à validade externa. A primeira é a validade externa da amostra em si. Lembre-se que não existe nada na nossa definição de experimento que descreva como pesquisadores recrutam ou selecionam pessoas para participar de um experimento. Reiterando: *experimentos não requerem uma amostra que represente a população*. De fato, é extremamente raro que experimentos utilizem uma amostra da população. Em experimentos para teste de medicamentos, por exemplo, é comum o uso de anúncios em jornais ou em programas de rádio para convidar pessoas a participar, usualmente envolvendo alguma forma de compen-

sação financeira. Claramente, pessoas que veem e respondem a anúncios como esses não conformam uma amostra aleatória da população de interesse, que tipicamente é formada por todos os potenciais usuários do remédio. Similarmente, quando professores "recrutam" pessoas de suas salas de aula (ou de colegas), os participantes não são uma seleção aleatória de nenhuma população⁵. O grupo de participantes, nesse caso, representa o que poderíamos chamar de **amostra de conveniência**, o que significa que esse é mais ou menos o grupo de pessoas que poderíamos implorar, coagir, seduzir ou convencer a participar.

Com uma amostra de conveniência, é simplesmente nebuloso generalizar os resultados de um experimento para uma população mais ampla, se é que existe um modo de fazê-lo. Como aprenderemos no capítulo 6, essa é uma questão crítica nas ciências sociais. Muitos experimentos fazem uso de tais amostras de conveniência, como em qualquer experimento único, sendo difícil saber se os resultados da análise são de alguma forma semelhantes aos que acharíamos em uma amostra diferente. Com desenhos de pesquisa experimentais, então, cientistas aprendem sobre como seus resultados se aplicam a amostras mais amplas por meio do processo de **replicação**, no qual pesquisadores implementam os mesmos procedimentos repetidamente de forma idêntica para observar se as relações se sustentam de um modo consistente.

Existe um segundo problema de validade externa com experimentos que é mais discreto, mas talvez tão importante quanto. Ele se relaciona ao problema da validade externa do estímulo. Continuando nosso exemplo de se anúncios de campanha afetam as intenções de voto, se conduzirmos um experimento para responder a essa questão, o que faríamos? Primeiro, necessitaríamos de uma amostra de voluntários. Lembre-se que ela não precisa ser aleatória. Segundo, dividiríamos os participantes em dois grupos a partir de um processo randômico: o experimental (que seria nosso grupo de tratamento) e o de controle. Então perguntaríamos aos indivíduos de ambos os grupos sua intenção de voto e faríamos a comparação entre os dois grupos. Do mesmo modo que podemos ter preocupações sobre a validade da nossa amostra, porque ela pode não ser representativa da população de interesse, devemos também nos preocupar com a validade externa do nosso estímulo. O que queremos dizer? O estímulo é a variável X. Neste caso, o estímulo consiste no ato de fazer os indivíduos assistirem a (diferentes) mensagens de vídeo. Quão semelhante é esse estímulo ao que uma pessoa experimenta em sua casa – isto é, em um ambiente mais natural? Em alguns pontos, ele é bastante diferente. Em nosso experimento hipotético, os indivíduos não escolhem a que assistem. A exposição ao anúncio é forçada (uma vez que os indivíduos aceitam participar). Em casa, pessoas que não desejam ser expostas a propagandas políticas podem facilmente evitar se quiserem, simplesmente não assistindo a determinados programas ou canais, ou não assistindo à televisão, ou mudando de canal quando um anúncio político começa. Mas a comparação

⁵ Pense nisso por um momento. Experimentos com estudantes de graduação de psicologia ou ciência política não são uma amostra aleatória de indivíduos com idade entre 18 e 22 anos, ou mesmo uma amostra aleatória de estudantes de graduação, ou mesmo uma amostra aleatória de estudantes da sua faculdade ou universidade. Sua turma de psicologia é composta por indivíduos com mais interesse em ciências sociais do que em ciências físicas ou engenharia.

entre o nosso experimento hipotético é insensível a essa diferença fundamental entre o ambiente do experimento e o ambiente natural do indivíduo. Na medida em que um experimento cria um ambiente inteiramente artificial, podemos nos preocupar com os resultados que encontraríamos em um contexto mais próximo do real⁶.

Desenhos de pesquisa experimentais, às vezes, podem ser confrontados com uma terceira desvantagem: nominalmente, que eles trazem dilemas éticos ao pesquisador. Questões éticas sobre o tratamento de participantes humanos ocorrem frequentemente com experimentos médicos. Se desejamos estudar experimentalmente os efeitos de diferentes tipos de tratamento do câncer sobre as chances de sobrevivência, isto requeria obter uma amostra de pacientes com câncer e então atribuir randomicamente aos pacientes diferentes regimes de tratamento. Isso não é tipicamente considerado uma prática médica aceitável. Em tais situações médicas de alto risco, a maioria dos indivíduos prefere tomar essa decisão em uma consulta com o médico, e não relegariam importantes decisões sobre o tratamento a um gerador de número aleatório.

Em experimentos de ciências sociais, questões éticas são menos frequentes, e tipicamente menos dramáticas, mas elas surgem em algumas ocasiões. Durante a revolução behaviorista na psicologia, durante os anos 1960, muitos experimentos famosos conduzidos nas universidades produziram debates éticos vigorosos. O psicólogo Stanley Milgram (1974) conduziu experimentos sobre quão facilmente ele poderia fazer indivíduos obedecerem a uma autoridade. Nesse caso, a variável dependente era a disposição de um participante a levar o que ele ou ela acreditava ser um choque de outro participante que era, na verdade, um empregado de Milgram. (O artifício foi que Milgram disse aos participantes que estava testando como reforços negativos – choques elétricos – afetavam o “aprendizado” de “estudantes”.) A variável independente era em que grau Milgram transmitia seu *status* de autoridade. Em outras palavras, o *X* que Milgram manipulou foi em que grau ele se apresentou como autoridade que devia ser obedecida. Para alguns participantes, Milgram vestiu um jaleco branco de laboratório e informou que era um professor da Universidade de Yale. Para outros, ele se vestiu mais casualmente e não mencionou sua filiação institucional. A variável dependente, então, foi quão forte o (falso) choque seria antes de o indivíduo simplesmente se recusar a prosseguir. No extremo mais alto, o instrumento dava os “choques” a “450 volts”. Os resultados do experimento foram fascinantes porque, para sua surpresa, Milgram descobriu que uma grande maioria de seus participantes estava disposta a administrar até mesmo os choques mais altos nos “aprendizes”. Hoje em dia, os comitês de revistas científicas consideram tais experimentos antiéticos, porque esse tipo de experimento cria um grau elevado de estresse emocional entre os participantes reais.

Uma quarta desvantagem potencial dos desenhos de pesquisa experimentais é que, quando interpretamos os resultados de um experimento, algumas vezes cometemos

⁶ Para uma discussão sobre a validade externa de experimentos utilizando *surveys* nacionais, ver Barabas e Jerit (2010). Para uma aplicação substantiva em que as questões de validade externa do estímulo são essenciais para determinar os resultados dos experimentos, ver Arceneaux e Johnson (2011). Ver também Morton e Williams (2010, p. 264), que se referem a esse problema como “validade ecológica”.

erros de ênfase. Se um experimento produz um achado de que algum X de fato causa Y , não significa que esse particular X é a causa mais proeminente de Y . Como temos enfatizado repetidamente, uma variedade de variáveis independentes pode ser causalmente relacionada com quaisquer variáveis dependentes interessantes nas ciências sociais. Desenhos de pesquisa experimentais frequentemente não nos ajudam a descobrir qual causa da variável dependente tem o maior efeito e quais outras possuem efeitos menores.

4.3 ESTUDOS OBSERVACIONAIS (EM DOIS SABORES)

Tomadas conjuntamente, as desvantagens dos experimentos significam que, para qualquer situação de pesquisa em ciência política, a implementação de um experimento frequentemente se prova impraticável e, algumas vezes, impossível. Como resultado, o método experimental não é o desenho de pesquisa utilizado mais comumente por cientistas políticos. Em alguns subcampos, como o da psicologia política – que, como o nome indica, estuda os fatores cognitivos e emocionais subjacentes à tomada de uma decisão política –, experimentos são bastante comuns. Eles têm se tornado mais comuns também no estudo da opinião pública e da competição eleitoral. Mas experimentos, para muitos pesquisadores, e por uma variedade de razões, se mantêm como uma ferramenta que não é aplicável a muitos dos fenômenos que buscamos estudar.

Isso significa que pesquisadores devem dar de ombros e abandonar a busca por conexões causais antes mesmo de começar? De modo algum. Mas quais são as opções que estudiosos têm quando não podem controlar a exposição de diferentes valores das variáveis independentes? Em tais casos, a única escolha é recolher dados do mundo como eles são e realizar comparações entre unidades individuais – como pessoas, partidos políticos ou países – ou entre uma quantidade **agregada** que varia ao longo do tempo. Esses dois exemplos representam duas das variantes mais comuns dos chamados estudos observacionais. Estes não são experimentos, mas buscam emulá-los. Eles são conhecidos como estudos observacionais porque, à diferença da natureza controlada e de certo modo artificial da maioria dos experimentos, nesses modelos de pesquisa, os pesquisadores simplesmente tomam a realidade como ela é e a “observam”, na tentativa de descobrir conexões causais sem o benefício da atribuição randômica dos participantes ao grupo de tratamento. Em vez disso, diferentes valores da variável independente já existem no mundo, e o que cientistas fazem é observá-las e, então, avaliar suas afirmações teóricas por meio dos mesmos quatro obstáculos causais para descobrir se X causa Y .

Isso leva à definição de um estudo observacional: um estudo observacional é um desenho de pesquisa no qual o pesquisador *não* tem controle dos valores da variável independente, que ocorrem naturalmente. Todavia, é necessário que exista algum grau de variabilidade na variável independente entre os casos, assim como variação na variável dependente.

Como não existe atribuição randômica ao grupo de tratamento, como nos experimentos, alguns estudiosos afirmam que é impossível falar em causalidade em estudos

observacionais e, portanto, se referem a eles algumas vezes como **estudos correlacionais**. Juntamente com a maioria dos cientistas políticos, não compartilhamos dessa visão. Certamente experimentos produzem um grau de confiança mais alto sobre relações causais do que estudos observacionais. Todavia, em estudos observacionais, se atenção suficiente é dada à análise de todas as outras possíveis causas da variável dependente sugeridas pela literatura corrente, então podemos fazer avaliações informadas com base na confiança que temos em que a variável independente causa a variável dependente.

Estudos observacionais, como essa discussão indica, enfrentam os mesmos quatro obstáculos causais que os experimentos. Lembre-se de que os obstáculos estão presentes em qualquer desenho de pesquisa. Então como, em estudos observacionais, superamos esses obstáculos? O primeiro obstáculo causal – existe um mecanismo crível que conecta X a Y ? – é idêntico em experimentos e em estudos observacionais.

Em um estudo observacional, no entanto, superar o segundo obstáculo causal – podemos eliminar a possibilidade de que Y causa X ? – pode algumas vezes ser problemático. Por exemplo, países com alto grau de desenvolvimento econômico (X) têm, como consequência, regimes democráticos mais estáveis (Y)? Cruzar o segundo obstáculo causal, nesse caso, é uma questão mais complicada. É totalmente plausível que ter um governo democrático mais estável faz com que a prosperidade econômica seja mais provável, o que é um cenário de causalidade reversa. Afinal, é mais provável que investidores se sintam mais confortáveis em arriscar em regimes democráticos do que nos autocráticos. Esses riscos, por sua vez, são mais prováveis de elevar a prosperidade econômica. É possível, claro, que X e Y se reforcem mutuamente – isto é, que X cause Y e Y cause X .

O terceiro obstáculo – existe covariação entre X e Y ? – não é, como mencionamos, mais difícil em um estudo observacional do que em um experimento. (As técnicas para examinar se a relação entre duas variáveis são diretas, e você as conhecerá nos capítulos 7 e 8.) Mas, ao contrário de um desenho experimental, se falharmos em achar uma covariação entre X e Y em um desenho observacional, devemos proceder ao quarto obstáculo, porque a possibilidade de acharmos uma covariação entre X e Y se mantém uma vez que controlamos por uma variável Z .

A comparação mais frequente entre experimentos e estudos observacionais, porém, ocorre com respeito ao quarto obstáculo causal. A quase mágica que acontece em experimentos por causa da atribuição aleatória ao grupo de tratamento – que possibilita aos pesquisadores saberem que nenhum outro fator interfere na relação entre X e Y – não está presente em um estudo observacional. Então, em um estudo observacional, a comparação entre grupos com diferentes valores para a variável independente pode muito bem estar poluída por outros fatores, interferindo em nossa habilidade de fazer afirmações conclusivas sobre se X causa Y .

Dentro dos estudos observacionais existem dois tipos puros – **os estudos observacionais de corte transversal**, com foco na variação entre **unidades espaciais** em uma única **unidade temporal**, e **os estudos observacionais de séries temporais**, com foco

na variação de uma única unidade espacial em múltiplas unidades de tempo. Existem, adicionalmente, desenhos híbridos, mas por uma questão de simplicidade focaremos os modelos puros⁷. Antes de entrarmos nos dois tipos de estudos observacionais, precisamos apresentar uma breve introdução aos dados observacionais.

4.3.1 DATUM, DATA E BANCO DE DADOS

A palavra *data* (dados) é uma das palavras gramaticalmente mais mal utilizadas da língua inglesa. Por quê? Porque a maioria das pessoas utiliza essa palavra como se estivesse no singular, quando ela está, na verdade, no plural. Todas as vezes que ler *the data is* (“os dados é”), você terá encontrado um erro gramatical. Quando descrevemos dados, a frase deve ser *the data are* (“os dados são”). Acostume-se a fazer isso: você agora é um dos soldados na cruzada pela utilização adequada dessa palavra, e essa será uma longa e difícil batalha.

A forma singular da palavra *data* (dados) é *datum* (dado). Juntos, uma coleção de *datum* (dado) produz *data* (dados) ou um *dataset* (conjunto de dados ou banco de dados). Definimos um conjunto de dados observacionais pelas variáveis que eles contêm e as unidades espaciais e temporais pelas quais eles são mensurados. Cientistas políticos utilizam dados mensurados por uma variedade de unidades espaciais. Por exemplo, em pesquisas de *survey*, a unidade espacial é o indivíduo que respondeu o *survey*. Em estudos comparados dos governos estaduais americanos, a unidade espacial são os estados dos Estados Unidos. Nas relações internacionais, a unidade espacial mais comum é o país. Unidades temporais mais comuns são meses, trimestres e anos. Também é comum se referir às unidades espacial e temporal que definem os conjuntos de dados como **dimensões do conjunto de dados**.

Os dois tipos mais comuns de conjunto de dados correspondem diretamente aos dois tipos de estudos observacionais que acabamos de introduzir. Por exemplo, a Tabela 4.1 apresenta um conjunto de dados no qual a unidade de tempo é o ano de 1972 e a unidade espacial são países. Esses dados podem ser usados para testar uma teoria que sustenta que o percentual de desemprego (X) \rightarrow a dívida do governo medida como percentual do produto interno bruto (Y).

Estudos observacionais de séries temporais contêm medidas de X e Y ao longo do tempo para uma única unidade espacial. Por exemplo, a Tabela 4.2 expõe um conjunto de dados de série temporal na qual a unidade espacial é os Estados Unidos e a unidade de tempo são meses. Poderíamos utilizar esses dados para testar uma teoria em que a inflação (X) \rightarrow a aprovação presidencial (Y). Com um conjunto de dados, pesquisadores analisam apenas aqueles dados que contêm valores mensurados tanto para a variável independente (X) quanto para a variável dependente (Y) a fim de determinar se o terceiro obstáculo causal é superado.

⁷ As afirmações clássicas dos estudos observacionais apareceram em 1963 no trabalho seminal de Donald Campbell e Julian Stanley, *Experimental and Quasi-experimental Designs for Research*.

Tabela 4.1 – Exemplo de dados transversais.

País	Dívida do governo como percentual do PIB	Taxa de desemprego
Finlândia	6,6	2,6
Dinamarca	5,7	1,6
Estados Unidos	27,5	5,6
Espanha	13,9	3,2
Suécia	15,9	2,7
Bélgica	45,0	2,4
Japão	11,2	1,4
Nova Zelândia	44,6	0,5
Irlanda	63,8	5,9
Itália	42,5	4,7
Portugal	6,6	2,1
Noruega	28,1	1,7
Holanda	23,6	2,1
Alemanha	6,7	0,9
Canadá	26,9	6,3
Grécia	18,4	2,1
França	8,7	2,8
Suíça	8,2	0,0
Reino Unido	53,6	3,1
Austrália	23,8	2,6

Tabela 4.2 – Exemplo de dados de série temporal.

Mês	Aprovação presidencial	Inflação
2002.1	83,7	1,14
2002.2	82,0	1,14
2002.3	79,8	1,48
2002.4	76,2	1,64
2002.5	76,3	1,18
2002.6	73,4	1,07
2002.7	71,6	1,46
2002.8	66,5	1,80
2002.9	67,2	1,51
2002.10	65,3	2,03
2002.11	65,5	2,20
2002.12	62,8	2,38

4.3.2 ESTUDOS OBSERVACIONAIS TRANSVERSAIS

Como o nome indica, um estudo observacional transversal examina a realidade social transversalmente, focando a variação entre unidades espaciais individuais – novamente, como cidadãos, políticos eleitos, distritos eleitorais ou países – e a explicação da variação da variável dependente entre elas.

Por exemplo, qual é a conexão, se existir, entre as preferências dos eleitores de um distrito (X) e o comportamento do parlamentar em votações nominais (Y)? Em um estudo observacional transversal, a estratégia que um pesquisador perseguiria para responder a essa pergunta envolveria a comparação da preferência agregada dos eleitores de vários distritos (X) com o histórico de votações nominais dos parlamentares (Y). Uma análise como essa, claro, teria que ser observacional, em vez de experimental, porque esse X não é passível de ser manipulado experimentalmente. Uma análise pode ocorrer confinada a uma única legislatura, por um número variado de razões práticas (um fator complicador óbvio seria a inexistência de mudança de parlamentares).

Tenha em mente, claro, que estudos observacionais têm que superar os mesmos quatro obstáculos causais que os experimentos. E, como temos notado, estudos observacionais, diferentemente de experimentos com atribuição aleatória do grupo de tratamento, frequentemente têm dificuldades de superar o quarto obstáculo causal. Esse pode ser o caso em nosso exemplo. Assumindo que os outros três obstáculos foram superados, considere a possibilidade de que existem outras variáveis colineares que causam Y e que também são correlacionadas com X , o que faz a conexão $X - Y$ espúria. (Você consegue pensar em algum destes fatores?) Como estudos observacionais transversais lidam com esse problema crítico? A resposta é que, na maioria dos casos, esse obstáculo pode ser superado por meio de uma série de controles estatísticos. Em particular, no começo do capítulo 9, você conhecerá a ferramenta mais comum nas ciências sociais para “controlar por” outras possíveis causas de Y , nominalmente o modelo de regressão múltiplo. O que você aprenderá é que a regressão múltipla pode possibilitar ao pesquisador observar como controlar, quando necessário, por outra variável (como Z) afeta a relação entre X e Y .

4.3.3 ESTUDOS OBSERVACIONAIS DE SÉRIES TEMPORAIS

A outra grande variante dos estudos observacionais é o estudo observacional de séries temporais, que tem, em seu núcleo, a comparação ao longo do tempo de uma única unidade espacial. Diferentemente da variação transversal, na qual a relação examinada é entre unidades individuais tipicamente em um único ponto do tempo, em estudos observacionais de séries temporais cientistas políticos tipicamente examinam a variação dentro de uma unidade espacial ao longo do tempo⁸.

Por exemplo, como a mudança na cobertura da mídia sobre a economia (Y) afeta (se é que afeta) a preocupação pública sobre a economia (Y)?⁹ Sendo um pouco mais

⁸ As unidades espaciais analisadas em estudos de séries temporais são usualmente medidas de forma agregada.

⁹ Ver Iyengar e Kinder (2010).

específico, quando a mídia passa mais tempo falando sobre problemas potenciais de inflação, o público se mostra mais preocupado com a inflação? E, quando a mídia gasta menos tempo falando de inflação, a preocupação do público com inflação se esvai? Podemos mensurar essas variáveis de modo agregado aproveitando sua variação ao longo do tempo. Por exemplo, quantas matérias sobre inflação foram apresentadas no noticiário noturno em um dado mês? É quase certeza que a quantidade não será a mesma em todos os meses. E quanto o público que assiste ao noticiário se preocupa (podemos captar essa informação por meio de pesquisas de opinião, por exemplo) com a inflação em um dado mês? Novamente, o percentual de pessoas que identificam inflação como um problema urgente, podemos afirmar com quase certeza, variará de um mês para outro.

Claro, assim como nos estudos transversais, os estudos observacionais de séries temporais requerem que nos foquemos muito no quarto obstáculo causal. Controlamos por todas as variáveis colineares (Z) que estão relacionadas à variação no volume da cobertura da imprensa sobre inflação (X) e sobre a preocupação do público com a inflação (Y)? (O terceiro exercício no final deste capítulo pedirá para você pensar sobre esse tópico.) Se pudermos identificar qualquer outra possível causa do porquê de, algumas vezes, o público estar mais preocupado com a inflação e, outras vezes, estar menos preocupado, então necessitaremos controlar por esses fatores na nossa análise.

4.3.4 A MAIOR DIFICULDADE COM ESTUDOS OBSERVACIONAIS

Apontamos que os desenhos de pesquisa experimentais possuem algumas desvantagens. Assim também ocorre com os estudos observacionais. Focaremos apenas em uma dessas desvantagens, mas é uma grande desvantagem. Como os exemplos anteriores demonstraram, quando precisamos controlar outras possíveis causas de Y para superar o quarto obstáculo causal, precisamos controlar por *todas elas*, não por apenas uma¹⁰. Mas como sabemos se controlamos por todas as possíveis causas de Y ? Em muitos casos, não sabemos com certeza. Mas, claro, precisamos tentar controlar estatisticamente por todas as outras possíveis causas que pudermos. O que envolve considerar cuidadosamente pesquisas prévias sobre o assunto e coletar o máximo possível de dados sobre essas outras causas. Mas, em muitos casos, simplesmente não seremos capazes de fazer isso perfeitamente.

O que isso tudo significa é que, em nossa opinião, a análise observacional deve ser um pouco mais cautelosa em seus pronunciamentos sobre causalidade. De fato, se tivermos feito o melhor que pudermos para controlar por tantas causas de Y quanto possível, então a conclusão mais sensível a que podemos chegar, em muitos casos, é de que X causa Y . Mas, na prática, nossas conclusões raramente são definitivas e pesquisas subsequentes podem modificá-las. Sabemos que isso pode ser frustrante para a

¹⁰ Como poderemos observar no capítulo 9, tecnicamente, precisamos controlar apenas pelos fatores que podem afetar Y e que também são relacionados a X . Na prática, porém, essa é uma distinção muito difícil de ser feita.

maioria dos estudantes que estão se familiarizando com a pesquisa em ciência política – e isso também pode ser frustrante para pesquisadores. Mas o fato de que respostas conclusivas são difíceis de obter deve apenas nos fazer trabalhar duro para identificar outras causas de Y . Uma importante parte de ser um cientista é que raramente podemos chegar a conclusões sobre causalidade; devemos manter aberta a possibilidade de que alguma variável previamente desconsiderada (Z) emergirá e tornará espúria a relação que previamente encontramos.

4.4 RESUMO

Para quase todo fenômeno de interesse de cientistas políticos existe mais de um desenho de pesquisa que eles podem implementar para responder a perguntas sobre relações causais. Antes de começar um projeto, pesquisadores precisam decidir entre utilizar um método experimental ou observacional; e, como é comum acontecer, se optarem pelo último eles devem decidir qual tipo de estudo observacional utilizar. E, algumas vezes, pesquisadores escolhem mais de um tipo de desenho de pesquisa.

Diferentes desenhos de pesquisa ajudam a iluminar diferentes perguntas. Quando queremos focar uma questão simples como a preferência do público por uma política governamental liberal ou conservadora, estudos transversais e de séries temporais são ambos úteis. Eles simplesmente respondem a diferentes tipos de questões substantivas. A abordagem transversal busca observar por que alguns indivíduos preferem políticas governamentais mais liberais e outros preferem políticas mais conservadoras. Essa é uma tarefa perfeitamente válida para cientistas políticos: o que faz com que algumas pessoas sejam liberais e outras sejam conservadoras? Considere agora a abordagem de um desenho de pesquisa com uma série temporal, que foca porquê de o público (como um todo, isto é, de maneira agregada) preferir uma política governamental mais liberal ou mais conservadora em diferentes pontos no tempo. Essa é uma questão simples. Nenhuma das abordagens é inerentemente melhor ou pior do que a outra. Ambas jogam luz em diferentes aspectos da realidade social. Qual desenho de pesquisa os pesquisadores devem escolher depende do tipo de pergunta que eles pretendem fazer e responder.

CONCEITOS INTRODUZIDOS NESTE CAPÍTULO

- Agregado – uma quantidade que é criada pela combinação de valores de muitos casos individuais.
- Amostra aleatória – um método de seleção de casos individuais para um estudo no qual cada membro de população de interesse tem a mesma probabilidade de ser selecionado.
- Amostra de conveniência – uma amostra de casos da população de interesse em que o mecanismo de seleção não é aleatório.

- Atribuição randômica – quando os participantes de um experimento são atribuídos randomicamente a um dos diferentes valores possíveis de X , a variável independente.
- *Dataset* (conjunto de dados) – sinônimo de *data* (dados). Uma coleção de valores de variável para ao menos duas observações.
- *Datum* (dado) – forma singular da palavra *data* (dados).
- Desenho de pesquisa – estratégias que um pesquisador emprega para fazer comparações com o objetivo de avaliar afirmações causais.
- Dimensões do conjunto de dados – unidades espaciais e temporais que definem um conjunto de dados.
- Estudo correlacional – sinônimo de “estudo observacional”.
- Estudo observacional – desenho de pesquisa no qual o pesquisador não tem controle dos valores da variável independente, que ocorrem naturalmente; é necessário que exista um grau de variabilidade entre os casos na variável independente e na variável dependente.
- Estudo observacional de corte transversal – um desenho de pesquisa que foca a variação entre unidades espaciais em uma única unidade de tempo.
- Estudo observacional de série temporal – um desenho de pesquisa que foca a variação de uma única unidade espacial em múltiplas unidades de tempo.
- Experimento – desenho de pesquisa no qual o pesquisador controla e atribui randomicamente valores da variável independente aos participantes.
- Experimento de campo – um estudo experimental que ocorre em uma configuração natural na qual os indivíduos vivem suas vidas.
- Experimento de *survey* – uma técnica da pesquisa de *survey* em que o processo de entrevista inclui alguma randomização experimental no estímulo do *survey* (no que é perguntado).
- Experimento natural – situação na natureza que não é propriamente definida como experimento, mas em que os valores da variável independente surgem naturalmente de maneira a dar a impressão de que ocorreu uma verdadeira atribuição aleatória pelo pesquisador.
- Grupo de controle – em um experimento, o subgrupo de casos que não é exposto ao principal estímulo causal em investigação.
- Grupo de tratamento – em um experimento, o subconjunto de casos que é exposto ao principal estímulo causal em investigação.
- Placebo – em um experimento, um estímulo inócuo dado ao grupo de controle.
- População – o conjunto inteiro de casos para o qual nossa teoria se aplica.
- Replicação – um processo científico no qual pesquisadores implementam os mesmos procedimentos repetidamente de forma idêntica para observar se as relações se mantêm de modo consistente.

- Unidade espacial – a unidade geográfica que forma a base para a observação. (unidade geográfica na qual os casos são mensurados)
- Unidade temporal – a unidade de tempo que forma a base para a observação (unidade temporal na qual os casos são mensurados).
- Validade externa – o grau em que podemos confiar que os resultados da nossa análise se aplicam não somente aos participantes e as circunstâncias do estudo, mas também à população amplamente dita.
- Validade interna – o grau de confiança do nosso estudo em que podemos afirmar que a variável independente causa a variável dependente.

EXERCÍCIOS

1. Considere as seguintes relações causais entre uma variável independente e uma variável dependente. Em cada um dos casos, seria realista um pesquisador conduzir um experimento para testar a teoria? Se sim, descreva brevemente como seria a atribuição randômica no experimento; se não, explique brevemente as razões.
 - a) O nível de religiosidade de um indivíduo (X) e sua preferência por diferentes candidatos políticos (Y).
 - b) Exposição a notícias políticas negativas (X) e apatia política (Y).
 - c) Serviço militar (X) e atitudes em relação à política externa (Y).
 - d) As características pessoais de um palestrante (X) e o poder de persuasão (Y).
2. Considere a relação entre o nível educacional (X) e o comparecimento eleitoral (Y). Como o desenho de pesquisa observacional de corte transversal difere do de série temporal?
3. Na seção sobre estudos observacionais de séries temporais, introduzimos a ideia de como a variação no nível de cobertura sobre a inflação (X) pode causar variação nas preocupações do público sobre a inflação (Y). Você consegue pensar em alguma variável Z relevante pela qual, estatisticamente, precisamos controlar em nossa análise para termos confiança de que a relação entre X e Y é causal?
4. Nos capítulo anterior (especificamente, na seção intitulada “Por que o estudo da causalidade é tão importante? Três exemplos da ciência política”), demos exemplos de problemas de pesquisa. Para cada um desses exemplos, identifique a(s) unidade(s) temporal(is) e espacial(is). Também relate se o estudo era um experimento, um estudo observacional de corte transversal ou um estudo observacional de série temporal.
5. A Tabela 4.1 apresenta dados para o teste de uma teoria utilizando um estudo observacional de corte transversal. Se a mesma teoria fosse testada com um estudo observacional de série temporal, como a tabela de dados se pareceria?
6. Compare os dois desenhos no teste da teoria anterior. Quais são as variáveis Z pelas quais você precisa controlar em cada uma das duas formas de estudos observacionais?

7. A Tabela 4.2 apresenta os dados para o teste de uma teoria utilizando um estudo observacional de série temporal. Se a mesma teoria fosse testada com um estudo observacional de corte transversal, como a tabela de dados seria?
8. Compare os dois desenhos no teste da teoria anterior. Quais são as variáveis Z pelas quais você precisa controlar em cada uma das duas formas de estudos observacionais?
9. Utilize a biblioteca da sua universidade ou o Google Acadêmico para acessar cada um dos artigos a seguir e determinar se o desenho de pesquisa utilizado em cada um deles é um experimento, um estudo observacional de corte transversal ou um estudo observacional de série temporal. (Observação: para ter acesso a esses artigos, talvez você necessite estar conectado à rede de computadores da sua universidade.)
 - a) CLARKE, Harold D.; MISHLER, William; WHITELEY, Paul. Recapturing the Falklands: Models of Conservative Popularity, 1979–83. *British Journal of Political Science*, v. 20, n. 1, p. 63-81, 1990.
 - b) GIBSON, James L.; CALDEIRA, Gregory A.; BAIRD, Vanessa A. On the Legitimacy of National High Courts. *American Political Science Review*, v. 92, n. 2, p. 343-358, 1998.
 - c) DRUCKMAN, James N. The Implications of Framing Effects for Citizen Competence. *Political Behavior*, v. 23, n. 3, 2001.