

WLADIMIR G. GRAMACHO

# INTRODUÇÃO À METODOLOGIA EXPERIMENTAL



**Blucher**

Wladimir G. Gramacho

INTRODUÇÃO À  
METODOLOGIA  
EXPERIMENTAL

*Introdução à metodologia experimental*

© 2023 Wladimir G. Gramacho

Editora Edgard Blücher Ltda.

*Publisher* Edgard Blücher

*Editor* Eduardo Blücher

*Coordenação editorial* Jonatas Eliakim

*Produção editorial* Lidiane Pedroso Gonçalves

*Preparação de texto* Ana Lúcia dos Santos

*Diagramação* Guilherme Henrique

*Revisão de texto* Maurício Katayama

*Capa* Leandro Cunha

*Imagem da capa* iStockphoto

## Blucher

---

Rua Pedroso Alvarenga, 1245, 4º andar

04531-934 – São Paulo – SP – Brasil

Tel.: 55 11 3078-5366

[contato@blucher.com.br](mailto:contato@blucher.com.br)

[www.blucher.com.br](http://www.blucher.com.br)

Segundo o Novo Acordo Ortográfico, conforme 6. ed. do *Vocabulário Ortográfico da Língua Portuguesa*, Academia Brasileira de Letras, julho de 2021.

É proibida a reprodução total ou parcial por quaisquer meios sem autorização escrita da editora.

---

Todos os direitos reservados pela Editora Edgard Blucher Ltda.

Dados Internacionais de Catalogação na Publicação (CIP)  
Angélica Ilacqua CRB-8/7057

---

Gramacho, Wladimir G

*Introdução à metodologia experimental* / Wladimir G. Gramacho. – São Paulo : Blucher, 2023.

216 p.

Bibliografia

ISBN 978-65-5506-430-8

1. Pesquisa – Metodologia 2. Metodologia científica  
I. Título

22-4518

CDD 001.42

---

Índice para catálogo sistemático:

1. Pesquisa – Metodologia

# Conteúdo

<b>INTRODUÇÃO</b>	<b>19</b>
Por que utilizar metodologia experimental?	19
<i>Ensaio normativo e estudos empíricos</i>	20
<i>Um dilema para enfrentar a pandemia</i>	21
<i>A revolução experimental</i>	23
<i>A organização do livro</i>	24
<i>Um convite à curiosidade e à igualdade</i>	25
<b>1. CONCEITOS BÁSICOS</b>	<b>27</b>
1.1 Um pouco de história	27
1.2 O que são experimentos?	30
1.3 Validade interna e validade externa	36
1.3.1 <i>Validade interna</i>	36
1.3.2 <i>Validade externa</i>	39

1.4 Na prática, para que servem os experimentos?	41
1.5 E para que servem os experimentos na ciência?	42
1.6 Vantagens e desvantagens do uso dos experimentos	45
1.7 Metodologias similares	47
1.7.1 Quase experimentos	47
1.7.2 Experimentos naturais	49
1.8 Conclusão	51
Exercícios	51
Olha essa... Trabalhador não precisa descansar? A história do efeito Hawthorne	54
Gabarito dos exercícios	54
<b>2. CONTEXTOS EXPERIMENTAIS</b>	<b>57</b>
2.1 Experimentos de laboratório	57
2.1.1 Vantagens de se realizarem experimentos em laboratório	61
2.1.2 Desvantagens de se realizarem experimentos em laboratório	62
2.1.3 Quando utilizar experimentos de laboratório	64
2.2 Experimentos em surveys	64
2.2.1 Vantagens de se realizarem experimentos em survey	68
2.2.2 Desvantagens de se realizarem experimentos em survey	71
2.2.3 Quando utilizar experimentos em survey	73
2.3 Experimentos de campo	73
2.3.1 Vantagens de se realizar um experimento de campo	78
2.3.2 Desvantagens de se realizar um experimento de campo	78
2.3.3 Quando utilizar experimentos de campo	80
2.3.4 Mídias sociais: uma nova fronteira dos experimentos de campo	81
2.4 Conclusão	83
Exercícios	84

Olha essa... Pesquisa com universitários em laboratório virou prêmio Nobel: a descoberta da teoria prospectiva	86
Gabarito dos exercícios	87
<b>3. DESENHOS DE PESQUISA EXPERIMENTAL</b>	<b>89</b>
3.1 Desenho entressujeitos	90
3.2 Desenho intrassujeito	92
3.3 Desenhos fatoriais	100
3.3.1 <i>Desenhos fatoriais completos</i>	101
3.3.2 <i>Desenhos fracionários</i>	103
3.3.3 <i>Desenhos fatoriais mistos</i>	104
3.4 Desenhos experimentais para o estudo de assuntos socialmente sensíveis	105
3.4.1 <i>Experimentos em lista</i>	106
3.4.2 <i>Experimentos de endosso</i>	107
3.4.3 <i>Experimentos de resposta randomizada</i>	108
3.5 Experimentos de escolha declarada	109
3.5.1 <i>Experimentos de escolha declarada pareados</i>	110
3.5.2 <i>Experimentos de escolha declarada simples</i>	111
3.5.3 <i>Vantagens e desvantagens de experimentos de escolha declarada</i>	113
3.6 Outros desenhos experimentais	114
3.6.1 <i>Desenho multinível</i>	114
3.6.2 <i>Desenho com blocagem</i>	115
3.6.3 <i>Desenho de pares combinados</i>	117
3.7 Conclusão	118
Exercícios	118
Olha essa... Meu partido é mais importante que a democracia?	
Um experimento de escolha perturbador	120
Gabarito dos exercícios	121

<b>4. VARIÁVEIS, MENSURAÇÃO E MODELOS DE ANÁLISE DE DADOS EXPERIMENTAIS</b>	<b>123</b>
4.1 Variáveis	123
4.1.1 Variáveis nominais	124
4.1.2 Variáveis ordinais	125
4.1.3 Variáveis discretas	125
4.1.4 Variáveis binárias	126
4.2 Mensuração de efeitos experimentais	127
4.2.1 Autodeclarações	127
4.2.2 Observação do comportamento	129
4.2.3 Reações fisiológicas	131
4.3 Modelos de análise	132
4.3.1 Variáveis dependentes	133
4.3.2 Variáveis independentes	133
4.3.3 Variável de tratamento	134
4.3.4 Covariáveis	136
4.3.5 Variáveis moderadoras	136
4.3.6 Variáveis mediadoras	137
4.3.7 Variáveis de confusão	138
4.3.8 Variáveis de controle	139
4.4 Conclusão	140
Exercícios	141
Olha essa... Homens heterossexuais fazem mais sexo que mulheres heterossexuais?	143
Gabarito dos exercícios	144
<b>5. A IMPLEMENTAÇÃO DOS EXPERIMENTOS</b>	<b>147</b>
5.1 Recrutamento de participantes	147
5.1.1 Técnicas de amostragem	148

5.1.2	<i>Tipos de participantes</i>	150
5.1.3	<i>Tamanho da amostra</i>	151
5.1.4	<i>Incentivos à participação</i>	152
5.2	Formação dos grupos	153
5.2.1	<i>Técnicas de designação aleatória</i>	153
5.2.2	<i>Número de condições experimentais</i>	154
5.2.3	<i>Número de participantes em cada condição experimental</i>	154
5.3	Sessão experimental	155
5.3.1	<i>Antes do tratamento e da mensuração</i>	155
5.3.2	<i>Durante o tratamento e a mensuração</i>	156
5.3.3	<i>Após o tratamento e a mensuração</i>	157
5.4	Conclusão	158
	Exercícios	158
	Olha essa... O maior experimento da história	161
	Gabarito dos exercícios	161

## **6. A QUALIDADE EM EXPERIMENTOS** **163**

6.1	Pré-testes e estudos piloto	163
6.2	Testes de balanceamento	165
6.3	Uso de placebo	166
6.4	Checagem de manipulação	168
6.5	Checagem de atenção	169
6.6	Análise de efeitos aura e chifre	170
6.7	A transparência no relato do experimento	171
6.8	A disponibilização dos dados e o incentivo à replicação	173
6.9	Pré-registro de experimentos	175
6.10	Conclusão	175
	Exercícios	176



Olha essa... Estudantes desmascaram fraude de professor	178
Gabarito dos exercícios	179
<b>7. QUESTÕES ÉTICAS SOBRE EXPERIMENTOS</b>	<b>181</b>
7.1 Os piores experimentos da história	181
7.2 O Relatório Belmont e suas limitações	183
7.3 O sistema brasileiro de ética em pesquisa	184
7.4 O Termo de Consentimento Livre e Esclarecido (TCLE)	185
7.5 O uso de engano	186
7.6 Conclusão	188
Exercícios	189
Olha essa... Obama pediu desculpas por experimento na Guatemala	191
Gabarito dos exercícios	191
<b>8. CONCLUSÕES</b>	<b>193</b>
8.1 E agora, o que mais?	195
<b>BIBLIOGRAFIA</b>	<b>197</b>

# Conceitos básicos

Este capítulo conta os episódios centrais na história da pesquisa experimental e introduz os conceitos essenciais para se compreender como essa metodologia funciona e por que a designação aleatória dos participantes faz com que os experimentos sejam considerados, hoje, uma estratégia poderosa nos estudos de relações causais. Não é por outra razão que experimentos são cada vez mais utilizados, em escalas cada vez maiores, e influenciam os demais métodos de pesquisa, na busca da inferência de causalidade. Este capítulo também introduz os temas da validade interna e da validade externa dos experimentos, discute suas principais utilidades práticas e científicas, aponta vantagens e desvantagens no uso dessa metodologia e distingue os experimentos verdadeiros de outras metodologias similares, como os quase experimentos e os experimentos naturais.

## 1.1 UM POUCO DE HISTÓRIA

Apontar o marco zero na trilha histórica de uma inovação científica não é nada fácil. Afinal, uma geração de intelectuais sempre avança de algum modo, marca o caminho com ideias preliminares e aponta destinos promissores que põem a geração seguinte mais próxima de uma descoberta. E, de tempos em tempos, alguém enxerga mais longe e registra um grande avanço nessa jornada, ao mapear com clareza

o que se aprendeu até aquele momento, ao explicar esse progresso de forma mais persuasiva e ao mostrar que relevância esse novo conhecimento tem sobre nossa capacidade coletiva de resolver problemas em qualquer área do conhecimento. No caso da metodologia experimental, esse papel parece caber com justiça ao britânico Ronald Fisher (Yates, 1964), que, ainda no início do século XX, formalizou argumentos matemáticos e estatísticos para lidar com dilemas sobre o desenvolvimento de sementes na agricultura, sobre a evolução de espécies na biologia e, surpreendentemente, sobre como saber se o paladar de uma pessoa pode descobrir corretamente qual líquido foi derramado primeiro em uma xícara de chá com leite!

Ainda que pareça improvável, uma tentativa de galanteio com a preparação de uma xícara de chá com leite está na origem da revolução experimentalista (Salsburg, 2001). Era o ano de 1920, e Ronald Fisher estava interessado em aproximar-se da pesquisadora Muriel Bristol, sua companheira de trabalho na prestigiosa estação experimental agrícola de Rothamsted, na Inglaterra. Em uma tarde, o estatístico decidiu oferecer uma xícara de chá com leite a Bristol. Preparou a bebida, típica entre ingleses, com esmero, entregou-lhe pessoalmente a xícara e esperou pelo elogio. A reação dela, entretanto, não poderia ter sido pior: disse que jamais tomaria aquilo. Afinal, explicou, o leite sempre deve ser servido na xícara depois do chá, nunca antes.

Desconfiado, Fisher duvidou da capacidade da colega em distinguir qual líquido havia sido depositado antes na xícara que lhe havia entregado. O engenhoso pretendente preparou, então, oito xícaras: em quatro, o leite foi colocado primeiro; em outras quatro, foi colocado depois. Submeteu a coleção à prova de Bristol. Ela aceitou o desafio e, sem titubear, identificou corretamente a ordem de composição de todas as oito xícaras. Fisher foi malsucedido na abordagem amorosa... Pouco tempo depois, viu a aproximação e o casamento de Bristol com outro colega de trabalho, William Roach. Mas a frustração e a insistente desconfiança com o teste das oito xícaras fizeram com que Fisher desenvolvesse nos anos seguintes ideias fundamentais para seu livro *Statistical Methods for Research Workers*, publicado em 1925 e que, desde então, orientou o desenvolvimento da metodologia experimental e transformou a capacidade explicativa do trabalho científico.

Ainda que vários outros especialistas também tenham contribuído com ideias relevantes sobre conceitos e desenhos experimentais,<sup>1</sup> o livro de Fisher teve uma influência ampla e duradoura sobre diversas áreas do conhecimento, tanto nas ciências da saúde como nas ciências exatas e nas ciências sociais (Yates, 1951; Henderson, 1986; Thompson, 1990; Efron, 1998; Levitt e List, 2009; Stanley, 1966). O legado de Fisher inclui, entre outros avanços, uma teoria sobre a designação aleatória de indivíduos em estudos experimentais, um teste exato para a análise de tabelas de contingência e o desenvolvimento de modelos estatísticos conhecidos como Análise de Variância (ANOVA).

---

1 Para saber mais sobre a história da metodologia experimental, veja Oakley (1998).

No campo dos estudos de comunicação, por exemplo, o primeiro livro a documentar um amplo esforço de utilização da metodologia experimental foi publicado décadas depois, em 1949. *Experiments on Mass Communication*, de Carl Hovland, Arthur Lumsdaine e Fred D. Sheffield, apresenta uma série de estudos que buscaram medir a efetividade de filmes, programas de rádio e revistas impressas em influenciar as atitudes e aumentar o conhecimento factual de recrutas americanos durante a Segunda Guerra Mundial. Para aumentar o desempenho de suas tropas, o exército americano criou uma Seção de Experimentos, subordinada ao Departamento de Pesquisas da Divisão de Educação e Informação, que deveria testar os efeitos de produtos de comunicação dirigidos aos soldados.

O estudo conduzido por Hovland e seus colegas procurou medir os efeitos do filme *The Battle of Britain*, de Frank Capra, sobre recrutas americanos. O filme é o quarto da série *Why We Fight*, produzido pelo governo dos Estados Unidos para justificar sua participação na Guerra. Os pesquisadores estavam especialmente interessados em analisar as atitudes dos soldados em relação à sua participação no conflito e em medir seu nível de conhecimento sobre fatos específicos associados à resistência britânica diante das tropas nazistas, durante o outono europeu de 1940.

A pesquisa adotou um desenho bastante simples. Primeiramente, diversas tropas foram pareadas com base na semelhança entre algumas características preponderantes dos seus soldados, como educação, idade e região do país da qual vinham. Cada par tinha as duas tropas mais parecidas entre si. Então, os pesquisadores lançavam uma moeda ao ar – literalmente – e, conforme o resultado (cara ou coroa), designavam a tropa ao grupo que seria tratado (ou seja, as tropas que assistiram ao filme e depois responderiam a um questionário), ou a unidades que serviriam como grupo de referência, ou controle (que não assistiria ao filme e simplesmente responderia a um questionário).

Conforme planejado, o grupo tratado assistiu ao filme e respondeu às questões, enquanto o grupo controle respondeu diretamente o questionário, sem ver o filme. Ao totalizarem as respostas e compararem os resultados entre os grupos, os pesquisadores encontraram evidências de que a exposição ao filme teria aumentado atitudes positivas em relação à atuação britânica contra os nazistas, melhorado seu conhecimento factual sobre a batalha e movido as opiniões das tropas em uma direção, em geral, mais favorável à participação americana na guerra.

Apesar da relevância histórica do livro, a pesquisa do trio não pôde assegurar, em termos lógicos, que assistir ao filme foi a *única causa* das diferenças entre as respostas dos grupos tratado e controle. Isso porque Hovland e seus colegas optaram por parear unidades de recrutas com base em características observadas (como idade e educação), em vez de designar aleatoriamente (ou seja, sortear) cada indivíduo para uma das duas condições experimentais. Ao não usarem o sorteio, os pesquisadores jamais poderiam saber se as opiniões mais favoráveis se deveram,

de fato, à exposição ao vídeo e não a outras características dos soldados que não foram medidas e que poderiam estar concentradas em um dos grupos experimentais. Por exemplo, se os filhos de britânicos que imigraram para os Estados Unidos fossem muito mais numerosos no grupo que viu o filme, as respostas ao questionário poderiam ter sido mais positivas não pela exposição ao vídeo, mas pelo efeito dos vínculos afetivos com o Reino Unido e a preocupação com os familiares que ficaram em seu país de origem.

Mas como assegurar que a designação aleatória dos participantes em um estudo experimental pode determinar uma relação de causa e efeito? Como saber se a exposição a um filme, a uma peça de marketing ou a uma vacina produz, de fato, algum efeito? A seção seguinte responde a essas questões e apresenta a definição de experimentos que será adotada ao longo deste livro. O exemplo que ilustrará a explicação, entretanto, não será o de um filme de guerra. Felizmente, a evolução da pesquisa experimental nos últimos anos tem nos ajudado a enfrentar outros inimigos, como a pobreza, a desigualdade e as enfermidades. Voltemos, portanto, ao dilema da ministra da Saúde, que precisa escolher o porta-voz de uma campanha de comunicação em meio a uma pandemia.

## 1.2 O QUE SÃO EXPERIMENTOS?

Suponha que seja possível duplicar perfeitamente qualquer grupo de pessoas que se queira estudar.<sup>2</sup> Imagine que, ao pesquisar porta-vozes mais efetivos para uma campanha contra uma pandemia, fosse possível ter dois grupos de idosos exatamente iguais em todas as suas características, pois cada um deles estaria duplicado. Poderíamos reunir em uma sala um dos grupos – digamos, grupo A – e distribuir a cada uma das pessoas um panfleto ilustrado por uma atriz aposentada, defendendo a importância do distanciamento social para enfrentar a pandemia. Ao mesmo tempo, poderíamos reunir em uma segunda sala, idêntica à primeira, o outro grupo – digamos, grupo B – e dar a cada um ali um panfleto com as mesmas características físicas e de conteúdo, salvo por um detalhe: ali estaria representada a imagem de um ex-jogador de futebol, e não a da atriz aposentada. Para manter o exemplo simples, suponha que a pesquisa pretende saber qual é a intenção de cada pessoa, após ler o panfleto, de vacinar-se anualmente contra a covid-19 e de usar máscaras durante o contato social, a fim de conter a pandemia. Para manifestar sua intenção em adotar esses comportamentos, cada pessoa só precisará indicar um número em uma escala de 1 a 7, em que 1 significa que ela não pretende “de

---

2 Exemplo adaptado de Gerber e Green (2012, p.23).

jeito nenhum” se vacinar anualmente e continuar usando máscaras em público, e 7 significa que ela pretende “com certeza” fazer ambas as coisas<sup>3</sup>.

A Tabela 1.1 representa um resultado possível desse teste, limitado a dez indivíduos apenas por uma questão de espaço, já que ela poderia ter qualquer outro número de linhas sem que isso afetasse a lógica do exemplo. O grupo A seria formado pelas versões “a” dos indivíduos um a dez, e o grupo B, pelas versões “b” dos mesmos indivíduos. A diferença de intenção de comportamento das observações 1a e 1b, portanto, só poderia ser explicada pela leitura de panfletos diferentes, já que essas observações 1a e 1b se referem ao mesmo indivíduo 1. No caso do exemplo, o indivíduo 1 registrou maior intenção em aderir à vacinação anual e usar máscaras ao ler o panfleto ilustrado pelo ex-jogador; porém, note que foi o único, já que, em geral, a associação da imagem da atriz aposentada à recomendação de vacinação e proteção facial parece ter sido mais efetiva. Nesse caso, o maior interesse do estudo não é descobrir qual dos porta-vozes tem maior influência sobre a intenção de comportamento do indivíduo 1, mas o efeito médio de cada porta-voz sobre o conjunto de idosos pesquisados. A última linha da tabela mostra que, em média, a intenção de distanciamento do grupo A após ler o panfleto foi de 5,5, não muito perto do ideal, em uma escala de 1 a 7. Ainda assim, melhor do que a intenção de distanciamento do grupo B depois da leitura, de 4,6. A diferença entre as duas médias sugere que o efeito médio do tratamento (EMT) foi de 0,9. Ou seja, a leitura do panfleto com a atriz teria aumentado em 20% ( $5,5/4,6$ ) a intenção de manter o distanciamento social em comparação com a exposição à peça ilustrada pelo ex-jogador. Mas esse é o resultado de um exemplo hipotético, em um mundo em que fosse possível duplicar pessoas.

No mundo real, entretanto, tudo o que se pode fazer é indicar o indivíduo 1 para o grupo A ou B, o mesmo para o indivíduo 2, e assim sucessivamente, até o indivíduo 10. Conhecemos essa condição como problema fundamental da inferência causal.<sup>4</sup> Podemos ter uma amostra aleatória das várias formas de sortear os indivíduos 1 a 10 entre os grupos A e B. Suponha que, como são apenas dez indivíduos, queremos ter cinco em cada grupo. Há 252 formas possíveis de sortear cinco indivíduos para cada grupo ( $10!/5!(10-5)! = 252$ ). Uma delas está apresentada na Tabela 1.2. Aqui, só podemos saber a resposta do indivíduo 1 submetido à condição do grupo A (atriz). Não sabemos como ele reagiria ao estímulo do grupo B (jogador), muito menos o efeito da mudança de porta-vozes para esse indivíduo. Mas, ainda assim, conseguimos calcular o efeito médio da atriz para os indivíduos do grupo A (5,6) e o efeito médio do jogador para os indivíduos do grupo B (4,4). Para essa amostra aleatória de todas as 252 possíveis, portanto, o EMT foi de 1,2.

---

3 Saiba mais sobre escalas no Box 1.1.

No mundo real, entretanto, tudo o que se pode fazer é indicar o indivíduo 1 para o grupo A ou B, o mesmo para o indivíduo 2, e assim sucessivamente, até o indivíduo 10. Conhecemos essa condição como problema fundamental da inferência causal.<sup>4</sup> Podemos ter uma amostra aleatória das várias formas de sortear os indivíduos 1 a 10 entre os grupos A e B. Suponha que, como são apenas dez indivíduos, queremos ter cinco em cada grupo. Há 252 formas possíveis de sortear cinco indivíduos para cada grupo ( $10!/5!(10-5)! = 252$ ). Uma delas está apresentada na Tabela 1.2. Aqui, só podemos saber a resposta do indivíduo 1 submetido à condição do grupo A (atriz). Não sabemos como ele reagiria ao estímulo do grupo B (jogador), muito menos o efeito da mudança de porta-vozes para esse indivíduo. Mas, ainda assim, conseguimos calcular o efeito médio da atriz para os indivíduos do grupo A (5,6) e o efeito médio do jogador para os indivíduos do grupo B (4,4). Para essa amostra aleatória de todas as 252 possíveis, portanto, o EMT foi de 1,2.

**Tabela 1.1** Exemplo de um teste hipotético de porta-vozes para dois grupos idênticos (*continua*)

Indivíduo	Grupo A (atriz aposentada)	Grupo B (ex-jogador de futebol)	Efeito do tratamento (A-B)
Indivíduo 1	4	5	-1
Indivíduo 2	6	3	3
Indivíduo 3	5	5	0
Indivíduo 4	7	6	1
Indivíduo 5	4	4	0
Indivíduo 6	5	3	2
Indivíduo 7	6	5	1
Indivíduo 8	7	7	0
Indivíduo 9	4	3	1
Indivíduo 10	7	5	2
Média	5,5	4,6	0,9*

(Fonte: Gerber & Green, 2012, p.23. Adaptado.)

\* Este exemplo hipotético considera que seria possível observar os comportamentos individuais sob ambos os tratamentos e que isso produziria uma estimativa precisa do efeito experimental. Na prática, entretanto, os dados que conseguimos obter estão normalmente sujeitos a variações e isso com frequência produz algum grau de incerteza sobre os efeitos observados.

4 Para conhecer uma discussão mais aprofundada sobre o “problema fundamental da inferência causal”, veja, por exemplo, Holland (1986).

Qual é a diferença entre os dois resultados? A Tabela 1.1 relata o efeito médio esperado do tratamento nesse teste hipotético, ou seja, a média dos efeitos para todas as amostras possíveis desse pequeno número de dez indivíduos. Isso quer dizer que, apesar do efeito relatado na Tabela 1.2 ter sido de 1,2, o valor verdadeiro desse efeito é um pouco menor, de 0,9. Mesmo assim, testar uma amostra aleatória dentre as 252 possíveis é o mais perto que se pode chegar do resultado obtido em um mundo hipotético, em que é possível duplicar as pessoas. Podemos adquirir um estimador (parâmetro estatístico) confiável e não enviesado para que o resultado obtido na Tabela 1.2 tenha sido o melhor resultado observável (veremos mais sobre estimadores confiáveis e não enviesados no Box 1.3).

Ao designar aleatoriamente cinco indivíduos para cada um dos dois grupos, pode-se reduzir o problema causado pela nossa condição do problema fundamental da inferência causal gerado, como dito anteriormente, pela impossibilidade de se observar o comportamento de uma mesma pessoa em dois mundos diferentes, ou seja, a inviabilidade de se conhecer a reação do indivíduo 1 se for submetido à condição A e a reação do mesmo indivíduo se submetido à condição B.<sup>5</sup> Como explicam Gerber & Green (2012, p.31), “a designação aleatória dos participantes resolve o problema desses ‘dados ausentes’ ao criar dois grupos de observações cuja expectativa de comportamento antes

Box 1.1 – ESCALAS LIKERT: O uso de escalas é um recurso valioso em estudos experimentais (assim como nos observacionais), pois permite obter mais informações que as oferecidas por questões com respostas binárias, que oferecem escolhas simplificadas de tipo “sim” ou “não” para assuntos que às vezes são mais complexos. A escala mais famosa é a Likert, cuja pronúncia correta, ‘*liquerti*’, é com frequência confundida com ‘*laiquerti*’). A versão tradicional dessa escala oferece aos respondentes de um questionário opções simétricas e bipolares de concordância com uma série de itens. Por exemplo, uma escala com 5 pontos teria as opções “discordo totalmente”, “discordo”, “indiferente”, “concordo” e “concordo totalmente”. Há, no entanto, outras versões, com diferentes números de pontos, e as unipolares, que medem a ausência total ou a presença total de determinado atributo.

Box 1.2 – O CICLO CIENTÍFICO: O conhecimento científico avança sempre que somos capazes de avaliar as teorias existentes que procuram explicar como o mundo funciona, desenvolvemos hipóteses que exploram ideias novas ou alternativas, testamos essas hipóteses, observamos os resultados e voltamos a avaliar a teoria de partida para validá-la, aprimorá-la ou descartá-la.

5 O desenho experimental intrassujeito (*within-subject*) representa uma exceção a essa regra e traz vantagens e desvantagens, que serão discutidas no Capítulo 3.



da aplicação dos tratamentos é idêntica”. A designação aleatória dos participantes em experimentos, portanto, é o elemento fundamental que permite assumir que os efeitos medidos ao final de um estudo devem-se, única e exclusivamente, à diferença de tratamentos a que cada grupo foi submetido. A utilização desse mecanismo tende a evitar, por motivos probabilísticos, que características não observadas dos indivíduos – sejam elas físicas, sociais, emocionais, cognitivas, políticas, econômicas, religiosas ou de outro tipo – estejam distribuídas de modo atipicamente desigual entre os grupos.

**Box 1.3 – UMA CONFUSÃO NADA ALEATÓRIA:** Atenção para não confundir *designação aleatória* com *amostra aleatória*. A *designação aleatória* é determinada por sorteio ou qualquer outro procedimento que assegure que cada indivíduo tem a mesma probabilidade de ser designado a cada um dos grupos do experimento. Portanto, a composição dos grupos não será correlacionada a nenhum fator individual. Já a *amostra aleatória* é uma técnica de seleção de indivíduos que busca extrair um grupo representativo de uma população maior, em que cada indivíduo tem a mesma probabilidade de ser escolhido para participar do estudo. Assim, garantimos que os parâmetros na formação da amostra não serão enviesados, ou seja, não representarão de modo desproporcional nenhuma característica da população da qual foram extraídos esses indivíduos.

**Tabela 1.2** Exemplo de um teste de porta-vozes para dois grupos

Indivíduo	Grupo A (atriz aposentada)	Grupo B (ex-jogador de futebol)	Efeito do tratamento (A-B)
Indivíduo 1	4	?	?
Indivíduo 2	6	?	?
Indivíduo 3	5	?	?
Indivíduo 4	7	?	?
Indivíduo 5	?	4	?
Indivíduo 6	?	3	?
Indivíduo 7	6	?	?
Indivíduo 8	?	7	?
Indivíduo 9	?	3	?
Indivíduo 10	?	5	?
Média	5,6	4,4	1,2*

(Fonte: Gerber & Green, 2012, p.31. Adaptado.)

\* Neste exemplo, a estimativa também traz um grau de incerteza pois não é possível observar o comportamento de todos os indivíduos em ambas as condições experimentais, assim como não é possível saber se outros indivíduos dariam respostas semelhantes às dadas nesta amostra hipotética. Qualquer estimativa, portanto, sempre terá associada a ela um grau de incerteza.

Mas como podemos definir um experimento? Ao longo deste livro, os experimentos – também conhecidos como *testes aleatorizados controlados* (em inglês, *randomized controlled trials*, ou RCT) – serão entendidos como estudos que combinam quatro características básicas:

- i) em primeiro lugar, devem ser testes científicos, no sentido de que precisam estar baseados em um método e apresentar uma hipótese que possa ser testada e refutada empiricamente;
- ii) em segundo lugar, esses testes devem conter uma proposição causal, ou seja, precisam explicitamente descrever uma relação de causa e efeito;
- iii) em terceiro lugar, o controle sobre as condições experimentais deve assegurar que mudanças em uma certa variável dependente  $y$  só pode ter respondido a diferentes valores de uma variável independente  $x$ ;
- iv) finalmente, em quarto lugar, esses testes precisam adotar o mecanismo de designação aleatória dos indivíduos às diferentes condições experimentais;

**Box 1.4 – PARÂMETROS CONFIÁVEIS E NÃO ENVIESADOS:** Um parâmetro estatístico é um número que sintetiza um conjunto de dados. Esses dados podem representar uma característica da população, como o percentual de pessoas que se definem como mulheres, homens ou transgênero no Brasil. Duas características importantes e intimamente conectadas dos parâmetros estatísticos é que eles devem ser confiáveis e não enviesados.

Parâmetros confiáveis são aqueles que, durante sucessivas medidas, retornam valores dentro de um intervalo de confiança esperado. Um exemplo simplificado é que, se uma estratégia para se obterem amostras da população brasileira encontra sucessivamente valores próximos a 52% de mulheres – como 50%, 53% ou 54% –, esse é um parâmetro confiável – já que, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), 51,8% das pessoas no país se identificam com o gênero feminino. Todavia, se uma amostra nacional se compõe de 90% de mulheres, algo errado ocorreu, e esse parâmetro deixa de ser confiável.

Para que se obtenham parâmetros confiáveis, é importante que a estratégia de coleção de dados não produza nenhum viés sobre essas informações, ou seja, que não aumente nem diminua as chances de ocorrência de um determinado evento. Mas, por exemplo, se uma estratégia de obtenção de uma amostra da população brasileira só selecionar indivíduos em horário comercial em endereços residenciais, provavelmente vai encontrar mais mulheres que homens – dada a divisão de trabalho ainda existente em parte da sociedade – e mais idosos que jovens. Esse procedimento, portanto, terá produzido um viés na formação dessa amostra.

- v) Em síntese, experimentos são testes científicos de proposições causais que pretendem medir a influência de uma variável independente  $x$  sobre uma variável dependente  $y$  por meio da comparação de respostas entre dois ou mais grupos formados pela designação aleatória dos participantes.<sup>6</sup>

### 1.3 VALIDADE INTERNA E VALIDADE EXTERNA

Nosso conhecimento sobre qualquer assunto pode ser ampliado pela nossa experiência direta com um objeto – seja ela episódica ou cotidiana –, ou pelo aprendizado indireto, por meio de relatos de outras pessoas, entre esses os reportados em pesquisas como as que usam metodologia experimental. Mas esses estudos só constituem, de fato, uma contribuição para ampliar nossa compreensão do mundo se eles demonstrarem coerência lógica e persuasiva em duas dimensões conhecidas como a validade interna e a validade externa dos experimentos. Sobre esses conceitos, dois alertas são importantes. O primeiro é que ambos são de definição bastante simples, porém, em geral, de execução complexa. O segundo é que, se for preciso escolher entre assegurar a validade interna ou a validade externa de um estudo, deve-se, certamente, optar pelo primeiro e abandonar o segundo. Ao final desta seção, você saberá o porquê.

#### 1.3.1 Validade interna

A validade interna dos experimentos trata de sua capacidade em demonstrar que um tratamento experimental foi, de fato, a única causa de uma mudança observada. Ou seja, um experimento é internamente válido se for convincente quando demonstra que a manipulação de uma certa variável independente  $x$  é a única razão para que se observem variações na variável dependente  $y$ . Nas ciências sociais, a tradicional pergunta que resume a importância da discussão da validade interna dos experimentos foi formulada por Donald Campbell, ainda em 1957, e subsiste até hoje: “O estímulo experimental fez, de fato, alguma diferença no caso específico?”. Se a resposta a essa questão for um convincente “sim”, então, pode-se dizer que o experimento é internamente válido.

Responder positivamente a essa pergunta, entretanto, não é nada fácil. Durante a realização de um experimento, muitos outros fatores – além do tratamento – podem levantar suspeitas de terem provocado uma mudança na variável dependente que representa o interesse central de uma pesquisa. À primeira vista, alguns desses fatores podem parecer inclusive irrelevantes. Por exemplo, no estudo hipotético de porta-vozes para a campanha pró-vacinação, o grupo que leu o panfleto estampado

---

<sup>6</sup> Para outras definições, ver McDermott (2002, p.32), Druckman *et al.* (2011, p.4) e Banerjee & Duflo (2017, p.78).

pelo ex-jogador de futebol poderia ter a presença de um participante que escolheu ir à sessão trajando a camisa de um clube de futebol com igual número de fãs e opositores. Esse evento, aparentemente corriqueiro, poderia, nesse estudo específico, estimular a aversão de alguns dos participantes, ao associarem o futebol, que às vezes divide opiniões, ao combate a uma pandemia, que precisa unir esforços. Se isso houvesse ocorrido, o melhor desempenho da atriz aposentada poderia não ter sido devido a uma comparação direta e livre de interferências entre ela e o ex-jogador, mas ao fato de que um gatilho tenha disparado uma atitude de resistência ou aversão a símbolos associados ao futebol. Isto é, o estimador do efeito causal médio se torna enviesado.

Não é possível fazer uma lista exaustiva de eventos que podem reduzir a validade interna dos experimentos, ou enviesar o cálculo do EMT entre os grupos experimentais. Em boa parte dos casos, cada desenho experimental tem desafios de validade interna específicos, derivados do problema de pesquisa que procura responder e do contexto experimental que decidiu adotar. Entretanto, alguns problemas de validade interna são mais comuns e sérios, como os cinco a seguir.

- 1) **Mortalidade (*mortality*):** ocorre quando não é possível medir os resultados do experimento em alguns indivíduos, seja porque estes se recusam a responder ao questionário, seja porque se perde o contato com eles. Se esses indivíduos estão distribuídos aleatoriamente entre as diferentes condições experimentais, isso não tem muita importância. Entretanto, se isso ocorre especialmente em um dos grupos, a validade do experimento pode estar ameaçada. Por exemplo, quando, ao receber um tratamento, parte dos indivíduos desiste de participar do experimento. Algum fator não previsto pode ter afastado esses sujeitos da pesquisa, e deixa de fazer sentido comparar os resultados entre os grupos. Essa ameaça à validade interna dos experimentos também é chamada de atrito (*attrition*);
- 2) **Descumprimento (*non-compliance*):** ocorre quando sujeitos designados para um grupo tratado não recebem o tratamento que deveriam receber, e/ou quando sujeitos designados para o grupo controle recebem o tratamento que não deveriam receber.<sup>7</sup> Nesse caso, não é possível calcular o EMT, uma vez que nem todo o grupo tratado recebeu o estímulo e, portanto, pode manifestar uma variação menos intensa da que deveria; e, por outro lado, parte do grupo controle recebeu o estímulo e, desse modo, não oferece um ponto de referência neutro, senão parcialmente alterado em direção ao tratamento proposto. Um controle de qualidade importante para avaliar esse problema é a checagem de manipulação, que será discutida no Capítulo 8;

---

<sup>7</sup> Para uma discussão avançada das soluções ao problema de descumprimento, ver os Capítulos 5 e 6 de Gerber & Green (2012).

- 3) *Vazamento (spillover)*: ocorre quando participantes de um experimento que receberam o tratamento transmitem informações contidas nesse estímulo a participantes do grupo controle (Sinclair *et al.*, 2012). De modo semelhante ao que pode ocorrer no caso de descumprimento, deixam de ser válidas as medidas obtidas do grupo controle, uma vez que alguns indivíduos desse grupo foram tratados indiretamente;
- 4) *Desvio comportamental*: ocorre quando os indivíduos aceitam participar do estudo, mas, por algum motivo, atuam de modo artificial, de modo propositalmente aleatório ou boicotam a pesquisa. Isso pode ocorrer por diversas razões, como por tentarem “ajudar a pesquisa a encontrar um resultado”;<sup>8</sup> por cansaço diante de um estudo longo demais; por aborrecimento diante de um estudo desinteressante; por desconforto em um local muito quente, frio ou ruidoso; por já estarem acostumados a participar de pesquisas experimentais; pela intervenção indevida de um fator imprevisto (como a presença de um participante cuja roupa pode ter efeitos sobre a resposta do grupo); ou pelo fato de manifestarem um comportamento atípico apenas por estarem sendo observados (conhecido como efeito *Hawthorne* – leia mais na Seção “Olha essa...”), entre outras razões. O desvio comportamental não é necessariamente um problema se estiver distribuído aleatoriamente entre as condições experimentais e puder ser controlado posteriormente. Mas ele pode enviesar os resultados e inviabilizar o estudo se estiver concentrado em uma das condições experimentais. Controles de qualidade importantes para analisar um desvio comportamental são a checagem de atenção e a análise dos efeitos *halo* e *horn*, que serão discutidas no Capítulo 6;
- 5) *Confusão (confounds)*: ocorre uma determinada variável pode influenciar tanto o tratamento quanto o resultado que se pretende medir, o que traz dúvidas sobre o processo causal. Por exemplo, se o panfleto estampado pela atriz tivesse um texto mais curto e a versão do ex-jogador tivesse um texto mais longo, não seria possível saber exatamente se a diferença de opinião entre os grupos seria devida à mudança de porta-vozes ou à variação no tamanho e conteúdo dos textos.<sup>9</sup> Além disso, a passagem do tempo entre a exposição a um estímulo e a mensuração da resposta pode permitir que os indivíduos se exponham também a outros eventos que logo criarão dúvidas sobre a exigência de uma relação causal. Nesse sentido, se a demora em mensurar os resultados de um experimento não for necessária para um estudo específico, quanto mais rápido ocorrer essa etapa, melhor.

---

8 Isso pode ocorrer devido a efeitos de expectativa (quando a pesquisadora comunica de modo sutil e/ou inadvertido a intenção do estudo) ou por efeito de demanda (quando os participantes se sentem motivados a cooperar para a obtenção de algum resultado específico).

9 Para uma discussão sobre esse problema, veja Thorson *et al.* (2012).

A validade interna dos experimentos, portanto, diz respeito à capacidade persuasiva do desenho experimental e de sua implementação em descartar todas as demais explicações plausíveis para a observação de uma mudança na variável dependente de uma pesquisa. Assim, um estudo pode reivindicar que está apontando uma relação causal, uma vez que o tratamento adotado parece ser a única causa de uma mudança observada, e a estimação do efeito causal médio não está enviesada.

### 1.3.2 Validade externa

A validade externa de um experimento diz respeito, essencialmente, à *possibilidade de generalizar suas conclusões, ou seja, de acreditar que os resultados observados para uma amostra valem também para toda a população da qual essa amostra foi extraída*. Se os resultados de um estudo sobre porta-vozes para uma campanha contra uma pandemia só valem para os participantes *daquele* estudo, que leram *aquele* panfleto, *naquela* sala de universidade, e medidos com *aquelas* perguntas de questionário, esse estudo não tem nenhuma validade externa. Ou seja, ainda que a pesquisa mostre um efeito importante na escolha de um porta-voz, os resultados não poderiam ser extrapolados para outros indivíduos, em outros contextos e nos quais a mensagem tenha outro formato. Isto é, não saberíamos como seriam as reações de outros idosos diante do panfleto; também não saberíamos quais seriam os efeitos desse panfleto se fosse enviado para as casas das pessoas ou para lares de idosos; e, finalmente, ignoraríamos os resultados da exposição a um anúncio de TV, em vez da leitura do panfleto. Finalmente, desconheceríamos o comportamento efetivo de cada indivíduo, para além das respostas dadas no questionário.

A validade externa de qualquer estudo experimental pode ser analisada a partir de seis fatores:<sup>10</sup>

- 1) se os *participantes* do estudo representam adequadamente os indivíduos que receberiam o tratamento em um contexto natural;
- 2) se o *contexto* espacial e temporal em que ocorre o estudo se assemelha ao contexto natural no qual os indivíduos receberiam o tratamento (também conhecido como “validade ecológica”);
- 3) se o *tratamento* utilizado no estudo representa corretamente o estímulo que os indivíduos receberiam naturalmente;
- 4) se a *mensuração* das respostas dos participantes ao tratamento do estudo registra corretamente o resultado que se pretende investigar;

---

10 Parágrafo adaptado de Shadish, Cook & Campbell (2002, p.83). Ver também Druckman, Green, Kuklinski e Lupia (2011, p. 19).

- 5) se os efeitos estimados utilizaram a *análise estatística* adequada, isolando corretamente o efeito do tratamento experimental de outras variáveis que podem estar associadas ao resultado de interesse de um estudo; e
- 6) se os resultados têm *reprodutibilidade*, isto é, se os achados podem ser encontrados também por outras pesquisadoras e em outros contextos.

Assegurar a completa validade externa de um experimento, portanto, é um desafio grande. Afinal, não só é difícil atender adequadamente aos cinco fatores apontados acima, como, com frequência, os experimentos analisam apenas um evento sobre os participantes, e pouco se pode antecipar a respeito de seu efeito continuado ou de seu efeito interativo com outros estímulos que ocorrem de modo imprevisível e caótico no mundo. Esses efeitos permanecem no tempo? Aumentam? Diminuem? Em algumas áreas de conhecimento, como as ciências sociais, pouco se sabe ainda sobre isso, mas, em outras, como na área de saúde, os experimentos em larga escala, como os utilizados para a análise da eficácia e segurança de vacinas, em geral, atendem satisfatoriamente aos requisitos de validade externa.

Quando esses requisitos não são atendidos, portanto, qualquer resultado de um estudo experimental sempre poderá estar sub ou superestimado em relação ao seu efeito real. Ainda que os resultados de um experimento estejam baseados em evidências empíricas, sua aplicação teórica e prática inevitavelmente acarretará algum grau de extrapolação lógica e dedutiva, cuja precisão é incerta.

Mesmo assim, diante do caráter essencialmente especulativo da metodologia observacional, a pesquisa experimental tem sido considerada como o “padrão ouro” dentre as narrativas científicas para construirmos nosso conhecimento sobre o mundo. Essa confiança tem sido particularmente ampliada com o uso de técnicas que aumentaram significativamente a validade externa dos tradicionais estudos de laboratório, com sua transposição para os experimentos de campo (Gerber *et al.*, 2014), que serão discutidos no Capítulo 2.

A esta altura, você provavelmente já sabe responder à pergunta do início desta seção sobre “a escolha de Sofia” entre a validade interna e a validade externa dos experimentos. Esse dilema não deve existir. Ainda que seja desejável assegurar ambas as validades de um estudo, isso muitas vezes é impossível. Mesmo assim, se um estudo tiver validade interna e nenhuma validade externa, podemos dizer que aprendemos algo sobre um fenômeno, ainda que ele tenha sido observado em um contexto muito específico e não possa ser generalizado. Entretanto, que sentido teria generalizar os resultados de um estudo sem validade interna, em que não se pode ter alguma certeza de que mudanças observadas foram causadas por uma certa variável manipulada, e não por outros fatores, sejam eles mais ou menos importantes? Pois é. Não faria sentido algum.

## 1.4 NA PRÁTICA, PARA QUE SERVEM OS EXPERIMENTOS?

Como podemos ter certeza de que um porta-voz é mais efetivo que outro para incentivar a vacinação periódica contra um vírus ou o uso de máscaras em locais públicos? Como alguém pode assegurar que uma nova campanha de comunicação recém-criada pelo governo para algum desses objetivos é mais eficaz do que simplesmente repetir os anúncios veiculados dois anos antes? Como optar entre uma narrativa emotiva e outra não emotiva ao desenhar uma mensagem efetiva para lidar com outro problema, como a proliferação do mosquito *Aedes aegypti*, que causa doenças graves, como a dengue, a zika ou a *chikungunya*?

Não deve ser surpresa imaginar que os experimentos, e não outras técnicas de pesquisa observacional, são a estratégia metodológica mais indicada para se chegar a respostas ao menos satisfatórias a essas perguntas de caráter aplicado. Entretanto, há uma razão ainda mais relevante para usar experimentos ao buscar soluções para questões como essas, por uma razão simples: sem testes experimentais, não é possível saber se um porta-voz escolhido piora, em vez de melhorar, a intenção das pessoas em vacinarem-se ou em usarem máscaras em tempos de pandemia, ou diminui, em vez de aumentar, a atenção das pessoas no enfrentamento a um mosquito que causa doenças que podem matar.

Conhecidos como *efeitos bumerangue*, esses resultados contrários aos pretendidos não são, infelizmente, tão raros. Ainda que de boa-fé, algumas decisões profissionais são tomadas sem a devida investigação sobre seus efeitos. De volta ao exemplo inicial, mesmo equipes competentes de comunicação veiculam campanhas que são mal recebidas, soam como desrespeitosas e, às vezes, podem pôr em risco a vida ou a saúde de quem pretendiam proteger. É o caso, por exemplo, de uma campanha de 2015 do Ministério da Justiça, que foi percebida como machista (e depois retirada do ar) ao usar o slogan “Bebeu, perdeu”, responsabilizando a vítima pelo que poderia ser feito com ela durante o carnaval. Já no final de 2016, o Ministério dos Transportes divulgou uma série de anúncios polêmicos. A campanha da pasta argumentava que mesmo pessoas que praticam boas ações podem matar se forem descuidadas no trânsito. Algumas das frases utilizadas foram: “Quem faz a alegria das crianças pode matar” e “Quem resgata animais na rua pode matar”. Outra campanha duramente criticada foi promovida pela revista *Vogue*, em 2016, que amputou – via Photoshop – partes do corpo dos artistas Cléo Pires e Paulinho Vilhena, para divulgar os Jogos Paralímpicos de 2016. Logo, a primeira utilidade prática de um experimento no campo da comunicação é assegurar que seus efeitos – se existirem – não serão o contrário do desejado. Se produzem os efeitos desejados, cabe, então, considerar qual dentre as alternativas é a mais efetiva.

Mas, além dessa utilidade, os experimentos são importantes por três outras razões. Em primeiro lugar porque permitem entender se um estímulo, como a escolha



de uma atriz aposentada e não de um ex-jogador de futebol, tem algum efeito significativo no sentido esperado, como incentivar a vacinação e o uso de máscaras em meio a uma pandemia depois de ler um panfleto sobre isso. Em segundo lugar, um experimento consegue mensurar o EMT, o que é uma medida objetiva, confiável e muito importante quando uma decisão não se dá apenas entre duas opções, mas entre três ou mais. Saber qual alternativa produz maior efeito tende a aumentar a eficiência das decisões. Finalmente, em terceiro lugar, a realização de sucessivos experimentos sobre um mesmo tema tende a aumentar nosso conhecimento sobre relações causais entre estímulos em qualquer área do conhecimento e a resposta dos indivíduos, fazendo com que as intervenções em momentos de crise – como em uma pandemia, em uma crise política ou em um colapso econômico – sejam mais efetivas e certas, e não um exercício de tentativa e erro.

## **1.5 E PARA QUE SERVEM OS EXPERIMENTOS NA CIÊNCIA?**

Além de terem uma relevância aplicada, os experimentos são considerados o “padrão-ouro” no processo de avanço científico. Isso porque eles oferecem condições ideais para a identificação e estimação de relações de causa e efeito entre duas variáveis. Essa é a principal utilidade da metodologia experimental em comparação com a metodologia observacional, que, em geral, só pode especular sobre relações entre variáveis que parecem estar associadas. Todavia, segundo Rose McDermott (2002), além desse interesse, os experimentos também são úteis na ciência por outras quatro razões.

Em primeiro lugar, quando os métodos de investigação existentes produzirem resultados inconsistentes ou contraditórios, os experimentos poderão ajudar a arbitrar entre esses achados. Um bom exemplo dessa vantagem na utilização dos experimentos vem dos estudos sobre a existência de estereótipos de gênero no processo de decisão do voto. As pesquisas sobre esse tema, inauguradas nos Estados Unidos, traziam resultados discordantes nos anos 1970. Enquanto alguns trabalhos sugeriam que o gênero dos candidatos e das (poucas) candidatas não fazia muita diferença para a decisão do voto (Darcy & Schramm, 1977), ou inclusive mostravam que o eleitorado estava mais propenso a votar em mulheres (Ferree, 1974), outros estudos registravam a resistência de alguns segmentos sociais em eleger mulheres (Bernstein & Polly, 1975).

Em 1981, entretanto, Virginia Sapiro realizou um experimento com estudantes universitários para tentar identificar e explicar a existência de estereótipos que prejudicassem eleitoralmente as mulheres e limitassem sua participação na política. Alunos e alunas receberam aleatoriamente uma de duas versões de um texto, feitas para apresentar uma candidatura à Câmara dos Deputados americana (*House of*

*Representatives*). Em um dos textos, quem se apresentava era John Leeds, enquanto na outra versão aparecia o nome de Joan Leeds – o que no Brasil seriam João e Joana, respectivamente. Os tratamentos eram idênticos, salvo na referência ao gênero do(a) candidato(a). Analisadas as reações dos estudantes, Sapiro não identificou uma diferença na intenção de votar em John ou em Joan, mas descobriu que tanto alunos como alunas acreditavam que as chances de que Joan ganhasse a disputa eram menores. Esse resultado não só explicou achados discordantes de pesquisas observacionais anteriores como inaugurou uma agenda de pesquisa ainda muito atual e relevante para a promoção da igualdade de gêneros na política.<sup>11</sup>

Em segundo lugar, uma vantagem em se usarem experimentos, segundo Rose McDermott, reside na possibilidade de se validarem empiricamente modelos formais e ilustrarem-se melhor conceitos e processos mais abstratos. Estudos de modelos formais baseiam-se em análises matemáticas e lógicas para se chegar a conclusões generalizáveis, sem qualquer coleta ou análise de dados, uma fragilidade que pode ser corrigida com o uso de experimentos. Em 1985, Randall Calvert sugeriu, em um modelo formal, baseado unicamente na teoria da maximização de utilidade, que as pessoas economizariam tempo para tomar decisões se usassem como referência as recomendações e as opiniões de fontes não neutras, ainda que isso pareça contraintuitivo. E Calvert foi ainda mais longe: sugeriu que essas fontes não neutras, claramente identificadas com um dos lados de uma questão, deveriam ser as primeiras a serem consultadas. Isso porque, se uma fonte que normalmente defende a posição A opta por – em um caso concreto – defender a posição B, isso deve significar que – naquele caso concreto – B é melhor ou mais verdadeiro que A.

Três décadas depois, em 2017, Adam Berinsky publicou um artigo sobre como desmentir com mais eficiência falsos rumores sobre uma mudança regulatória, mais especificamente, sobre a reforma do sistema de saúde americano proposta pelo democrata Barack Obama quando foi presidente dos Estados Unidos (2009-2017). Alguns boatos sugeriam que a proposta considerava que os idosos deveriam se apresentar a um “painel da morte”, em que funcionários avaliariam quem continuaria vivo e quem seria submetido à eutanásia. Ainda que pareça absurda a informação, políticos republicanos, como a ex-candidata a vice-presidente Sarah Palin (2008), ajudaram a difundir o rumor em suas páginas de Facebook. Por meio de um experimento, Berinsky mostrou que a correção mais eficiente desse boato ocorria quando um político também republicano como Palin – e não um democrata – desmentia o rumor. O resultado é uma comprovação empírica e aplicada do modelo formal proposto por Calvert.

Voltando à lista de McDermott, em terceiro lugar, uma vantagem na utilização de experimentos é que eles funcionam muito bem em *triangulações metodológicas*,

---

11 Veja, por exemplo, Dolan & Kira (2011) e, sobre o caso brasileiro, Biroli (2018).

em que uma pesquisa procura reforçar ou explicar alguns achados utilizando de forma combinada metodologias observacionais e experimentais. Nesse caso, um dos melhores exemplos vem da aplicação de experimentos de lista, que serão vistos com mais detalhe no Capítulo 3. Os experimentos de lista (*list experiments*) são uma técnica que combina, em um só instrumento de pesquisa (o questionário), a metodologia experimental e observacional, e que oferece um contexto de privacidade para que os entrevistados possam informar opiniões e comportamentos sinceros a respeito de temas polêmicos, sobre os quais é mais difícil falar abertamente, como questões morais, sexuais ou de preconceito.

Um bom exemplo vem da experiência da Universidade de Brasília, que, em 2003, passou a adotar um sistema de cotas raciais, com reserva de 20% das vagas para estudantes negros. Apesar de seus benefícios, a decisão gerou polêmica e dividiu opiniões por muito tempo, criando um contexto desconfortável para que os próprios alunos da raça negra declarassem seu apoio às cotas. Esse foi um dos principais achados de uma pesquisa publicada em 2014 por Mathieu Turgeon, Bruno Chaves e William Wives. Eles descobriram, em um estudo on-line com 1.190 estudantes, que, usando uma pergunta tradicional de *survey*, apenas 29% de estudantes negros declaravam abertamente seu apoio às cotas raciais, mas, quando se adotava um experimento de lista, o resultado era que o apoio sincero à nova regra reunia 68% dos estudantes negros. A combinação de abordagens metodológicas mescladas no experimento de lista permitiu não só estimar o apoio aberto à medida (29%), mas também medir seu apoio verdadeiro (68%) e a deseabilidade social, que fez com que expressivos 39% dos entrevistados da raça negra (68%-29%) não emitissem abertamente sua opinião sincera. Já o percentual de apoio às cotas entre os estudantes da raça branca, de 27%, não variou entre a pergunta aberta ou a versão em lista.

Finalmente, em quarto lugar, os experimentos podem ajudar a *encontrar evidências mais robustas para apoiar afirmações contundentes* sobre relações causais. Por exemplo, se há uma diferença de gênero no rigor com que eleitores penalizam homens ou mulheres que se envolvem em corrupção; mais especificamente, se as mulheres são punidas de forma mais severa na política quando associadas à malversação de recursos. Segundo Andrew Eggers, Nick Vivyan e Markus Wagner (2018), os estereótipos de gênero tendem a fazer com que esperemos dos homens mais assertividade e egocentrismo, e, das mulheres, mais compaixão e altruísmo. Por essa razão, diante de um caso de corrupção igualmente grave cometido por um homem e uma mulher, o eleitorado tenderia a castigar mais a mulher do que o homem, pois ela teria se desviado de uma expectativa mais exigente do que a aplicada aos homens. Isso de fato ocorre?

Os três pesquisadores investigaram essa questão designando aleatoriamente quase 2.000 eleitores do Reino Unido para diferentes simulações possíveis da eleição para uma vaga no Parlamento britânico. Essa vaga hipotética era disputada

por um(a) opositor(a) contra um(a) detentor(a) do cargo (que podia ser homem ou mulher, corrupto ou não corrupto). Cada simulação de votação apresentava só uma versão dentre todas as possíveis. Os resultados do estudo mostraram que as mulheres eram, sim, menos escolhidas que os homens caso estivessem envolvidas em um caso de corrupção (o desvio de 10.000 libras para gastos pessoais). Entretanto, só as eleitoras eram mais rigorosas com as mulheres. Os dados – ao menos para o caso britânico – não mostraram que os homens tendem a castigar mais as mulheres por corrupção, nem que castigam mais aos homens.

## 1.6 VANTAGENS E DESVANTAGENS DO USO DOS EXPERIMENTOS

A principal vantagem do uso dos experimentos, em relação às técnicas observacionais, é que eles dão ao(à) pesquisador(a) maior controle sobre todos os aspectos da condução de um estudo científico. Esse controle tende a aumentar a validade interna das pesquisas e permitir que os estudos façam afirmações sobre relações de causa e efeito. Para Rose McDermott, as vantagens de se usarem experimentos podem ser resumidas em cinco aspectos (McDermott, 2022, pp. 38-39):

- 1) a possibilidade de se identificarem relações causais;
- 2) o controle que o(a) pesquisador(a) tem sobre o recrutamento de participantes, o tratamento e a mensuração de respostas;
- 3) a capacidade de se estabelecerem mensurações objetivas e precisas;
- 4) a oportunidade de se estudarem diferentes etapas do processo causal; e
- 5) o custo relativamente baixo de se realizarem experimentos, especialmente os tradicionais estudos de laboratório com estudantes de universidades.

Os experimentos, entretanto, não são uma panaceia; eles têm suas desvantagens e não são a solução para todos os problemas de pesquisa. Em primeiro lugar, algumas perguntas de pesquisa só podem ser respondidas com estudos observacionais, mais apropriados para descrever um objeto de estudo ou um contexto. Por exemplo, quando se pretende identificar os hábitos de alimentação mais comuns na população, medir a intenção de voto do eleitorado em uma disputa presidencial ou apontar a prevalência de certas doenças em uma região. Nesses casos, estudos como os tradicionais *surveys* com amostras representativas são a melhor técnica para se obterem respostas válidas, uma vez que o objetivo é apenas descrever uma dada população, e não testar hipóteses de relações causais.

No entanto, mesmo as questões logicamente dirigidas a um estudo com desenho experimental encontram limitações importantes ou desvantagens quando o interesse pelos aspectos mais abstratos e conceituais de um estudo se transpõem

para uma realidade específica, concreta e complexa do desenho da pesquisa, da coleta de dados e da interpretação dos resultados. Essas desvantagens e limitações podem ser resumidas essencialmente a três aspectos.

O primeiro deles diz respeito a um *desenho experimental artificial*, que ocorre em um contexto físico muito diferente do real, que estuda pessoas muito diferentes das quais se pretende conhecer e que usa medidas experimentais muito diferentes do fenômeno que realmente se quer observar. Tudo isso reduz, como vimos antes, a validade externa de um estudo experimental. Desenhos artificiais de um experimento, todavia, têm sido menos comuns com a crescente adesão a estudos não laboratoriais (Shanto, 2002) e, especialmente, com o uso mais frequente de experimentos de campo, que serão discutidos no Capítulo 2. Contudo, como lembra McDermott (2002, p. 39), mesmo estudos artificiais têm suas vantagens quando pode ser antiético ou perigoso estudar um fenômeno fora do laboratório (por exemplo, o efeito da exposição a videogames violentos sobre atitudes no trânsito), ou quando é muito difícil reproduzir as condições naturais em que um evento ocorre (como em estudos sobre eleições, crises econômicas ou pandemias).

O segundo aspecto trata de olhar para o outro lado da moeda, representada por uma das vantagens de se usarem os experimentos citados. Ao manter mais controle sobre várias etapas de um estudo, o(a) pesquisador(a) pode aumentar a validade interna de seu estudo, mas também pode enviesar o desenho, a implementação e a análise de seu experimento, de modo consciente ou não. Logo, é especialmente relevante que os autores de cada estudo façam uma análise da presença de três potenciais problemas:

- 1) viés do(a) pesquisador(a), que ocorre quando crenças, opiniões e atitudes das pessoas que conduzem o estudo influenciam no desenho de um experimento e aumentam as chances de obtenção de um certo resultado desejado;
- 2) efeitos de expectativa, que ocorre quando o(a) pesquisador(a) indica aos participantes, ainda que de modo sutil, qual resposta espera receber; e
- 3) efeitos de demanda, que ocorre quando são os participantes que conseguem identificar o interesse da pesquisa, não suficientemente mantido em sigilo, e atuam de forma a ajudar a se obterem os resultados desejados.

Finalmente, o terceiro aspecto diz respeito aos impedimentos éticos que tornam muitos experimentos impraticáveis. Por exemplo, uma pesquisadora pode ter interesse em saber quais os efeitos psicológicos da exposição a *fake news* sobre vacinas. Para testar suas hipóteses, ela poderia designar aleatoriamente a sua variável de interesse (a exposição a essas *fake news*) a uma série de participantes, fazendo com que alguns experimentassem ler notícias falsas por um certo período, enquanto outros leriam apenas notícias corretas e verificadas

sobre o mesmo assunto. Mas seria essa uma estratégia válida do ponto de vista ético? É moralmente correto fazer com que seres humanos sejam expostos à desinformação apenas para aumentar nosso conhecimento científico sobre um determinado tema? Os comitês de ética na pesquisa lidam com questões como essa e procuram, em suas avaliações, ponderar os riscos e os benefícios de pesquisas experimentais. Nesse sentido, a ética é um fator fundamental a ser considerado no desenho de experimentos e pode, não raro, estar em conflito com as pretensões de cientistas, como veremos em mais detalhes no Capítulo 7. Em suma, ao optarmos pelo método experimental, estamos, assim como em outros tipos de método, limitando as questões que podemos investigar.

## 1.7 METODOLOGIAS SIMILARES

Esta seção trata de dois métodos de pesquisa que compartilham algumas características similares às dos estudos experimentais, mas que não cumprem os requisitos indicados anteriormente para serem considerados experimentos verdadeiros (*true experiments*). Um deles são os quase experimentos, como são denominados os estudos que não contemplaram a designação aleatória dos participantes às diferentes condições experimentais, mas que comparam efeitos de uma variável manipulada pelas pesquisadoras sobre um grupo tratado em relação a um grupo controle. O outro caso são os experimentos naturais, que, apesar do nome, não são experimentos, mas, sim, estudos baseados em dados tipicamente observacionais que adotam uma estratégia de análise experimental.

### 1.7.1 Quase experimentos

Em estudos em que não é possível realizar a designação aleatória dos participantes aos grupos tratado e controle, os experimentos perdem parte de seu poder de identificação de causalidade e – nesses casos – devem ser chamados de quase experimentos (Campbell & Stanley, 1963). Ainda que tenha menor força explicativa, um quase experimento não deve ser visto como um estudo sem nenhum valor. Em alguns casos, esses estudos podem constituir uma extensão do repertório da pesquisa experimental, quando aplicarem a lógica dessa metodologia a contextos em que é factível se obter alguma evidência empírica; mas, por motivos éticos, procedimentais ou de acesso, não é possível designar aleatoriamente os participantes às diferentes condições experimentais (Campbell, 1968). Em outras palavras, os quase experimentos são – de fato – *quase* experimentos, no sentido de que, não fosse a ausência ou impossibilidade de designação aleatória dos participantes às diferentes condições experimentais, eles seriam considerados experimentos verdadeiros.

Um exemplo sobre campanhas de prevenção ao suicídio em universidades ajuda a entender por que, às vezes, pode ser impossível implementar a etapa de

designação aleatória dos participantes em um estudo. Em 2015, pesquisadores de cinco universidades americanas estudaram o efeito de diferentes versões de campanhas de enfrentamento ao suicídio entre estudantes de graduação. Com base na teoria das normas sociais, eles elaboraram pôsteres, peças para painéis digitais e e-mails com mensagens que incentivavam estudantes a recomendarem aos colegas sob risco que buscassem os serviços de auxílio da universidade.

A pesquisa pretendia formar três condições experimentais: uma em que os estudantes vissem a campanha protagonizada por celebridades do time de basquete da universidade; outra em que a campanha fosse apresentada por estudantes comuns; e a última formada pelo grupo controle, que não seria exposto a qualquer mensagem sobre o tema. Mas como fazer com que alguns estudantes no mesmo *campus* só vissem a versão do time de basquete, outros só vissem a versão dos estudantes comuns e os demais não vissem nenhuma delas? Isso seria impraticável. Afinal, os pesquisadores não poderiam designar aleatoriamente cada estudante para uma condição experimental diferente se as duas campanhas estivessem em andamento no mesmo *campus*.

A solução encontrada foi distanciar fisicamente os tratamentos. Para conseguir esse resultado, os pesquisadores implementaram cada condição experimental em diferentes unidades da universidade, localizadas em bairros distintos. Portanto, em uma unidade, os estudantes foram expostos à campanha com os atletas; em outra, às peças protagonizadas por universitários comuns; e na terceira, que funcionou como grupo controle, não foi exibida nenhuma comunicação sobre o tema. Ao final do estudo, os resultados mostraram que os estudantes que frequentavam a unidade que recebeu a campanha protagonizada por colegas comuns manifestaram maior inclinação que os demais grupos a falar sobre os serviços de cuidado com saúde mental da universidade com companheiros que considerassem estar sob risco. Já os estudantes da unidade que recebeu a campanha estrelada pelo time de basquete não manifestaram maior disposição a fazer o mesmo quando comparados aos universitários do grupo controle.

A implicação prática mais importante do estudo é a sugestão de que nem sempre as celebridades terão mais influência sobre o público em campanhas de comunicação em saúde, especialmente quando a ação recomendada depende do contato interpessoal. Nesses casos, campanhas protagonizadas por pessoas comuns podem ter até maior efeito que as estreladas por celebridades.

Mas veja que, nesse caso, é mais adequado dizer que os resultados “sugerem” uma determinada conclusão do que dizer que eles “permitem concluir”. Afinal, essa conclusão deve ser vista com cautela, pois o estudo tem uma limitação metodológica importante. Como não foi possível designar aleatoriamente os estudantes para cada uma das condições do experimento, os pesquisadores não podem assegurar que os grupos eram semelhantes ao início do estudo. Ou seja, outras variáveis, além da exposição à campanha, não foram observadas nem controladas pelos pesquisadores

durante o estudo. Por exemplo, suponha que o grupo exposto à campanha protagonizada pelo time de basquete da universidade tivesse, antes da pesquisa, um interesse muito menor por esportes, em comparação aos demais grupos. Digamos que, nessa unidade da universidade, a principal atração e vocação dos estudantes fosse a música ou a dança. Nesse caso, não se pode dizer que a campanha com atletas não produziu um efeito significativo se comparada ao grupo controle, porque os estudantes que a viram tinham um interesse atipicamente menor em esportes que os demais. Talvez, se os grupos tivessem sido formados por designação aleatória, o interesse por basquete estaria igualmente distribuído entre as três condições experimentais, e essa não seria uma variável de confusão (*confounding variable*) que dificultaria o isolamento da relação causal entre a campanha com o time de basquete e a intenção de recomendar os serviços de prevenção ao suicídio na universidade. O Capítulo 4 vai explicar em detalhes o que é e por que é importante identificar variáveis de confusão antes de tirar conclusões num estudo.

Portanto, é impossível assegurar que o grupo que foi exposto à campanha protagonizada por estudantes comuns teve maior inclinação a sugerir os serviços de ajuda universitária por causa da campanha e não por outro fator não observado ou controlado durante a pesquisa. Isso pode ter ocorrido porque esse grupo teve maior contato pessoal com casos de suicídio, ou porque estava mais acostumado a participar de trabalhos sociais ou comunitários. Em circunstâncias como essa, ainda que haja uma contribuição ao debate científico e ao acúmulo de evidências a favor de alguma teoria ou hipótese, não se pode assegurar causalidade. Por esse motivo, ainda que compartilhem grande similaridade metodológica e procedimental com os experimentos, os quase experimentos não têm a mesma potência explicativa e científica, ainda que sejam em alguns casos a melhor solução possível.

## 1.7.2 Experimentos naturais

Já nos experimentos naturais, as pesquisadoras não apenas não designam os indivíduos aleatoriamente a diferentes condições experimentais como tampouco elaboram um tratamento. A única característica comum entre os experimentos naturais e os experimentos verdadeiros é a comparação entre diferentes grupos, dada uma designação “natural” de um certo estímulo. Isso porque os experimentos naturais são, na verdade, estudos observacionais cujos dados são analisados ao modo de um experimento, e o estímulo não é conduzido pelo pesquisador, mas por forças “naturais”.

Um exemplo pode ajudar a entender melhor esse conceito. Imagine que você está preocupado com o fim da circulação de jornais importantes em algumas cidades e desconfia que a ausência de veículos de imprensa com maior credibilidade pode reduzir o engajamento cívico das pessoas, além de eliminar uma fonte de



referência, em tempos de crescente circulação de *fake news*. Afinal, os jornais impressos têm sido vistos tradicionalmente como “cães de guarda” da sociedade, e sua circulação pode muito bem estimular seus leitores diretos e, indiretamente, os demais cidadãos<sup>12</sup> a participarem mais da vida comunitária e exigirem mais de seus governos. Portanto, você pode formular uma hipótese de que, quando um jornal deixa de circular em uma cidade, seus habitantes manifestarão menor engajamento cívico, ou seja, terão menos incentivos para fazerem coisas como contatar servidores públicos para expressarem suas opiniões, boicotar produtos ou serviços por motivos políticos e sociais, ou participar de reuniões comunitárias ou organizações sociais. Em uma situação como essa, você não poderia designar aleatoriamente cidadãos de uma cidade para dois mundos diferentes: um em que o jornal local mais importante deixou de existir e outro no qual esse jornal continua existindo. Também não poderia fechar jornais para implementar o *tratamento* necessário em um desenho verdadeiramente experimental. Nesse caso, a melhor solução possível é identificar cidades que perderam seus jornais impressos nos últimos tempos e observar mudanças em seus comportamentos, usando medidas tomadas antes e depois do fechamento desses veículos.

Foi o que fez o pesquisador Lee Shaker (2014), da Portland State University. Por problemas financeiros agravados pela crise econômica americana de 2008, os jornais *Rocky Mountain News*, de Denver, e *Seattle Post-Intelligencer*, de Seattle, encerraram sua circulação impressa no início de 2009. Ao analisar dados da pesquisa de opinião *Current Population Survey*, coletados em novembro de 2008 e novembro de 2009, Shaker observou mudanças importantes de comportamento dos habitantes dessas cidades. Mais que a média da variação ocorrida em outras localidades que não tinham perdido seus jornais entre um ano e outro, os habitantes de Denver e Seattle manifestaram menor engajamento cívico, pois os entrevistados em 2009 indicaram menor participação e atuação junto a organizações sociais e aos governos do que os entrevistados em 2008. É possível afirmar que o fechamento dos jornais foi *a causa* da diminuição do engajamento cívico nessas cidades? Certamente não. Mas o estudo, ainda assim, contribui com evidências empíricas que apoiam a expectativa de que os jornais impressos cumprem um papel importante na vida democrática e comunitária.

Eventos extraordinários, como pandemias, golpes de Estado, crises econômicas, decisões judiciais históricas, atentados terroristas e guerras, têm enorme potencial de transformação social e de influência sobre o comportamento de indivíduos, de empresas e dos governos em qualquer país. Entretanto, nenhuma pesquisadora tem condições de manipular *tratamentos* como esses, podendo apenas esperar que ocorram para analisar dados coletados antes e depois do evento e compará-los ao modo experimental.

---

12 Saiba mais sobre o processo de influência indireta da comunicação (*two-step flow*) em Katz (1957) e Druckman *et al.* (2017).

O fator distintivo dos experimentos naturais é a argumentação de que os grupos tratado e controle foram formados *como se* tivessem sido objeto de designação aleatória, ainda que não tenham sido sorteados. Como explica Thad Dunning (2012), muitos temas de interesse científico simplesmente não podem ser estudados com experimentos verdadeiros. Nesses casos, os experimentos naturais são a melhor solução.

## 1.8 CONCLUSÃO

Ainda que não saibamos, interagimos com experimentos praticamente todos os dias, atuando como objetos de pesquisa, não como pesquisadores. Quando usamos Google, Youtube, Instagram, Facebook ou LinkedIn, estamos, muitas vezes, participando – sem saber – de testes experimentais simples, conhecidos no mundo digital como testes A/B. Esses testes, vale dizer, não respeitam padrões éticos de pesquisa experimental pois não contam com o consentimento dos participantes, como veremos no Capítulo 7. Ainda assim, sua existência revela que experimentos podem servir para responder a um amplo conjunto de questões, desde as mais práticas e procedimentais, dos desafios profissionais e aplicados, às interrogações mais abstratas e teóricas, que se interessam por processos cognitivos, formação de atitudes e intenções de comportamento e comportamentos nas mais diversas áreas do conhecimento.

Este capítulo apresentou conceitos básicos e centrais da metodologia experimental para que você possa participar dos experimentos não só como objeto de pesquisa quando navega na internet, usa as redes sociais ou toma decisões de consumo, mas também como protagonista, desenhando, implementando e analisando experimentos. Além do breve resumo sobre os eventos históricos que deram origem à revolução experimental nas ciências sociais e na comunicação, o capítulo introduziu os principais conceitos para entendermos como funcionam os experimentos e para que servem. Expressões como designação aleatória, EMT, validade interna, quase experimentos ou experimentos naturais devem estar claras, a fim de que seja possível acompanhar o conteúdo dos próximos capítulos. O seguinte trata da mudança no contexto de realização dos experimentos, que cada vez mais têm saído dos laboratórios e ganhado as ruas, em busca de maior realismo e validade externa, que agora você já sabe o que é.

## EXERCÍCIOS

- 1) Os estudos experimentais se prestam, em geral, a:
  - a) indicar correlação entre duas ou mais variáveis
  - b) conhecer a opinião de uma amostra representativa da população, por meio de *surveys*

- c) conhecer a influência de um tratamento sobre a intenção de comportamento de um indivíduo
  - d) conhecer o efeito médio de um tratamento sobre o conjunto de indivíduos pesquisados
- 2) A principal característica dos experimentos que lhes permite estabelecer relações de causa e efeito é:
- a) a designação aleatória dos participantes do estudo para cada condição experimental
  - b) o controle amplo da condução do experimento em um laboratório;
  - c) o uso de técnicas estatísticas avançadas, como a análise multivariada e a regressão
  - d) a designação dos participantes do estudo para cada tratamento a partir de critérios preestabelecidos
- 3) Entre os atributos dos experimentos, não se encontra:
- a) a presença de método científico, com hipótese que possa ser testada e refutada
  - b) a obrigatoriedade de sorteio para a seleção de uma amostra representativa de participantes
  - c) a existência de proposição que descreva uma relação de causa e efeito
  - d) o controle que garanta que mudanças em uma variável dependente y advêm de mudanças em uma variável independente x
- 4) A validade interna de um experimento se refere:
- a) à capacidade de demonstrar que um tratamento experimental foi a única causa de uma mudança observada na variável dependente
  - b) à capacidade de generalizar as conclusões de um experimento;
  - c) à garantia de que os efeitos estimados utilizaram a análise estatística adequada
  - d) à garantia de que os participantes do estudo representam adequadamente os indivíduos que receberão o tratamento
- 5) Um exemplo de limitação à validade externa de um experimento é:
- a) o efeito *Hawthorne*
  - b) a mortalidade dos participantes
  - c) os problemas de validade ecológica
  - d) o efeito *halo*

- 6) Um exemplo de pergunta que pode ser respondida por meio de um desenho experimental é:
  - a) quem é o melhor porta-voz para uma campanha publicitária
  - b) qual é a intenção de voto para presidente dos eleitores de um país
  - c) quais são os hábitos de alimentação mais comuns em uma população
  - d) qual é o efeito do fechamento de um jornal impresso em dado município
- 7) Segundo McDermott, não se encontra entre as vantagens dos experimentos:
  - a) a possibilidade de arbitrar entre resultados prévios inconsistentes ou contraditórios
  - b) a facilidade de contornar dilemas éticos
  - c) a possibilidade de validar empiricamente modelos formais
  - d) a existência de triangulações metodológicas
- 8) A diferença entre os experimentos e os estudos quase experimentais é:
  - a) a ausência de designação aleatória de participantes nos estudos quase experimentais
  - b) a ausência de uso de Análise de Variância (ANOVA) nos estudos quase experimentais
  - c) a ausência de um grupo controle nos estudos quase experimentais
  - d) a ausência de controle em laboratório nos estudos quase experimentais
- 9) Os experimentos e os experimentos naturais possuem a seguinte característica em comum:
  - a) ambos designam aleatoriamente participantes a um tratamento
  - b) ambos comparam uma condição experimental a uma condição controle
  - c) ambos comparam diferentes grupos
  - d) ambos estabelecem relações causais com a mesma validade interna
- 10) O problema de descumprimento se refere às situações em que:
  - a) os participantes do estudo atuam de modo artificial
  - b) os participantes da condição experimental (tratamento) não transmitem informações aos participantes do grupo controle
  - c) os participantes não recebem o tratamento do grupo ao qual pertencem
  - d) os participantes não estão inseridos em um contexto espacial e temporal semelhante ao contexto real no qual os indivíduos receberão o tratamento

## **OLHA ESSA... TRABALHADOR NÃO PRECISA DESCANSAR? A HISTÓRIA DO EFEITO HAWTHORNE**

*Por Victor Gomes*

Durante os anos 1920 e 1930, uma série de experimentos foi realizada no complexo fabril Hawthorne Works, da Western Electric Company, em Illinois, nos Estados Unidos. O objetivo era descobrir os fatores que aumentam ou diminuem a produtividade dos trabalhadores. Primeiro, os pesquisadores investigaram o efeito da iluminação sobre o desempenho dos funcionários. Enquanto o grupo tratado foi submetido a diferentes intensidades de luz durante o expediente, o grupo controle trabalhou com uma iluminação constante, sem variações. Para a surpresa dos investigadores, o resultado desse estudo indicou um aumento generalizado na produtividade dos dois grupos de empregados, o que foi atribuído ao fato de que os funcionários teriam se comportado de uma maneira diferente simplesmente porque sabiam que estavam participando de um experimento (Brannigan & Zwerman, 2001). Outro experimento mostrou que a introdução de períodos de descanso não aumentava a produtividade dos empregados, ao contrário do que se supunha. Mais uma vez, o efeito foi atribuído à situação artificial em que a fábrica se encontrava, já que o contexto experimental era do conhecimento de todos. Essa interferência no processo experimental foi apelidada de “efeito *Hawthorne*”, sendo disseminada na literatura acadêmica a partir de 1953 (Wickström & Bendix, 2000). Apesar de críticas e contestações à conclusão de que resultados obtidos na Western Electric Company poderiam ser explicados pelo “efeito *Hawthorne*” (Jones, 1992), essa hipótese ganhou proeminência em campos como a psicologia social (O’Sullivan *et al.*, 2004).

### **GABARITO DOS EXERCÍCIOS**

- 1) Os estudos experimentais se prestam, em geral, a:
  - d) conhecer o efeito médio de um tratamento sobre o conjunto de indivíduos pesquisados.
- 2) A principal característica dos experimentos que lhes permite estabelecer relações de causa e efeito é:
  - a) a designação aleatória dos participantes do estudo para cada condição experimental.
- 3) Entre os atributos dos experimentos, não se encontra:
  - b) a obrigatoriedade de sorteio para a seleção de uma amostra representativa de participantes.

- 4) A validade interna de um experimento se refere:
  - a) à capacidade de demonstrar que um tratamento experimental foi a única causa de uma mudança observada na variável dependente
- 5) Um exemplo de limitação à validade externa de um experimento é:
  - c) os problemas de validade ecológica
- 6) Um exemplo de pergunta que pode ser respondida por meio de um desenho experimental é:
  - a) quem é o melhor porta-voz para uma campanha publicitária?
- 7) Segundo McDermott, não se encontra entre as vantagens dos experimentos:
  - b) a facilidade de contornar dilemas éticos
- 8) A diferença entre os experimentos e os estudos quase experimentais é:
  - a) a ausência de designação aleatória de participantes nos estudos quase experimentais
- 9) Os experimentos e os experimentos naturais possuem a seguinte característica em comum:
  - c) ambos comparam diferentes grupos
- 10) O problema de descumprimento se refere às situações em que:
  - c) os participantes não recebem o tratamento do grupo ao qual pertencem

**Este livro preenche uma importante lacuna no Brasil ao oferecer uma introdução à metodologia experimental, um método cada vez mais empregado para responder a importantes questões em diversas áreas do conhecimento, como as ciências sociais, de saúde e exatas, incluindo saúde pública, medicina, nutrição, comunicação, economia, pedagogia, ciência política, políticas públicas e relações internacionais.**

Por se esforçar em facilitar a compreensão de conceitos e técnicas complexas, trazendo exemplos de várias pesquisas aplicadas, este livro tem o potencial de ser um excelente texto introdutório para o ensino e para quem pensa em utilizar essas técnicas em suas pesquisas.

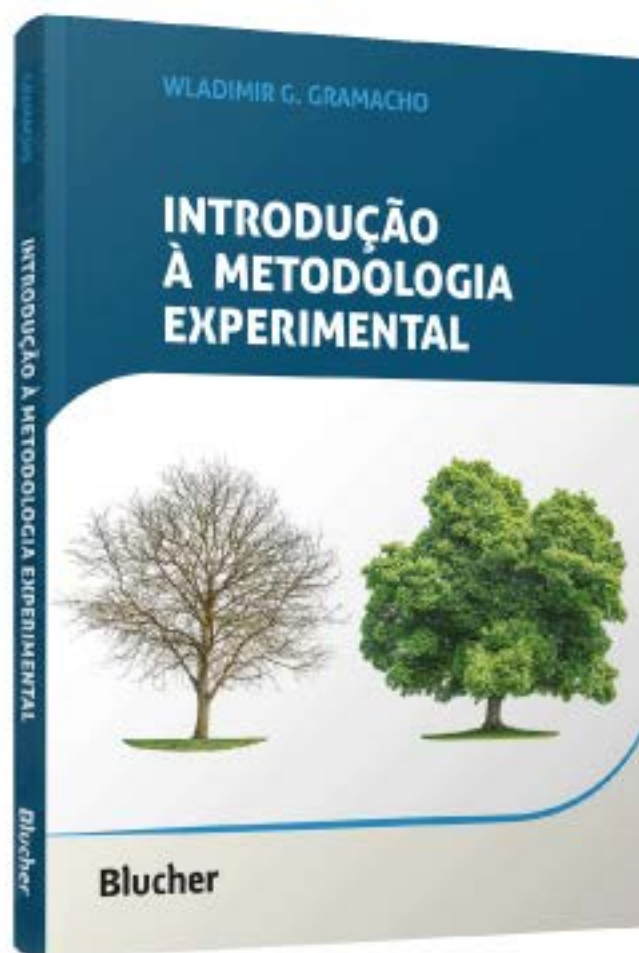
**Lorena G. Barberia**

*Professora livre-docente em Ciência Política da Universidade de São Paulo (USP)*



[www.blucher.com.br](http://www.blucher.com.br)

**Blucher**



Clique aqui e:

[VEJA NA LOJA](#)

## **Introdução à metodologia experimental**

Wladimir Ganzelevitch Gramacho

---

**Ana Cláudia Santos Meira**

ISBN: 9786555064308

Páginas: 216

Formato: 17 x 24 cm

Ano de Publicação: 2022

---