

Laboratório de Psicologia Experimental
Departamento de Psicologia – UFSJ
Disciplina: Método de Pesquisa Quantitativa

TEXTO 4: VALIDADE INTERNA E EXTERNA DE UMA PESQUISA

VIESES

Autora: Prof. Marina Bandeira, Ph.D.

Um experimento não confirma definitivamente uma hipótese. Ele apenas coloca à prova esta hipótese. Como resultado do experimento, nós podemos tomar duas decisões opostas: podemos decidir refutar, eliminar a hipótese, quando os dados foram contrários à nossa hipótese; ou então, podemos decidir não refutar, não eliminar aquela hipótese, quando nossos dados foram a favor de nossa hipótese. Neste último caso, podemos apenas dizer que por enquanto esta hipótese permanece e não será eliminada, nem a teoria de onde ela derivou será contestada. Neste caso, nossa hipótese estará provisoriamente sobrevivendo. Mas ela será colocada à prova novamente por um outro pesquisador, mais tarde. As hipóteses estão sempre à disposição para serem colocadas à prova. Portanto, fazer pesquisa é colocar à prova uma hipótese. Uma hipótese experimental consiste em um enunciado relacionando duas variáveis em uma relação causal. Ela é redigida em termos afirmativos, explicitando que a variável independente causará um determinado efeito em uma variável dependente.

Validade interna:

A validade interna de uma pesquisa se refere à certeza da relação entre a VI e a VD. Uma pesquisa será mais válida se pudermos ter mais certeza desta relação entre as duas variáveis, ou seja, se pudermos ter mais certeza de que nossa hipótese é a única explicação provável para os resultados obtidos e que outras possíveis explicações não são viáveis. Mas como fazer isto?

Para verificar uma hipótese através da realização de uma pesquisa e ter mais certeza do efeito da VI sobre a VD é necessário que a montagem de nossa pesquisa (o delineamento) seja capaz de evitar que variáveis espúrias interfiram com o efeito de nossa VI. Quando montamos a estratégia de nossa pesquisa, precisamos utilizar um delineamento adequado e, também, ter o cuidado de padronizar os procedimentos, na execução da pesquisa, de forma que os grupos de comparação, as instruções e os instrumentos de medida não sofram modificações. Estas duas atitudes visam evitar a interferência de variáveis espúrias, eliminando possíveis hipóteses rivais para explicar nossos dados. Cada variável espúria que pudesse interferir com o efeito da VI seria uma hipótese rival, ou seja, uma hipótese que estaria rivalizando com nossa hipótese experimental para explicar os resultados.

Uma pesquisa tem alta validade interna quando ela é capaz de diminuir as hipóteses rivais plausíveis, deixando apenas a nossa hipótese experimental como explicação plausível dos dados. A pesquisa só conseguirá diminuir hipóteses rivais plausíveis se sua montagem, ou seja, seu delineamento, puder controlar as variáveis espúrias que interferem com a interpretação dos dados. A validade interna, portanto, é garantida quando o delineamento escolhido permite se ter certeza de que as relações observadas empiricamente entre as variáveis VI e VD não podem ser

explicadas por outros fatores, outras variáveis além da VI, ou seja, quando se controla as explicações rivais plausíveis que poderiam explicar os resultados. Vieses são as variáveis espúrias que interferem com o efeito da VI. Quando controlamos os vieses de uma pesquisa, estamos garantindo sua alta validade interna. Não é possível controlar todos os vieses, mas procuramos diminuir o máximo possível. A montagem de pesquisa que é a melhor e mais respeitada como estratégia de garantia da validade interna é a distribuição aleatória dos sujeitos aos grupos.

Validade externa:

A validade externa de uma pesquisa se refere à capacidade de generalização dos resultados obtidos com a amostra da pesquisa para a população-alvo de onde ela foi retirada. Uma pesquisa possui validade externa quando ela permite ao pesquisador generalizar os resultados obtidos para a população-alvo de onde foi retirada a sua amostra. A validade externa de uma pesquisa vai depender de se poder mostrar que os resultados obtidos nesta pesquisa não são dependentes da amostra ou da situação particular desta pesquisa, mas que suas conclusões são verdadeiras também para população alvo de onde foi retirada a amostra. A melhor estratégia para se garantir a validade externa de uma pesquisa é compor a amostra que será estudada com sujeitos que sejam selecionados aleatoriamente da população-alvo, de modo que a amostra seja representativa da população.

VALIDADE INTERNA

Existe uma série de vieses ou fatores espúrios que interferem com a interpretação dos resultados quando o delineamento de uma pesquisa não permite eliminar a interferência destes fatores. Surgem, então, explicações rivais plausíveis dos resultados que são baseadas nestes vieses. Ao se elaborar um projeto de pesquisa, é necessário indicar quais são os procedimentos que o experimentador adotará para tentar controlar estes vieses ou, pelo menos, diminuir o seu efeito.

Há três procedimentos básicos que podem ajudar a controlar a maioria dos vieses. Dentre eles, entretanto, o principal procedimento e o mais respeitado para garantir a validade interna de uma pesquisa é a distribuição aleatória dos sujeitos da amostra aos grupos que serão comparados na pesquisa. Os três procedimentos básicos são os seguintes:

PROCEDIMENTOS BÁSICOS:

1. Ter um Grupo Controle com o qual se vai comparar os efeitos da VI. O grupo controle não é submetido a VI, apenas o grupo experimental é submetido à VI. Mas o grupo controle é importante, justamente para vermos o que acontece com os sujeitos quando eles não são submetidos à VI. Por exemplo, em uma pesquisa sobre o efeito de uma terapia, é importante saber se as pessoas podem melhorar após receber a terapia, mas é igualmente importante saber se um grupo que não foi submetido à terapia não melhora. Para que se conclua que a terapia foi eficaz é necessário observar que o grupo experimental melhorou após a terapia e que o grupo

controle, sem terapia, não melhorou. O grupo controle recebe o efeito das variáveis espúrias. Ao observarmos o que acontece com o grupo controle, vemos o efeito de outras variáveis que não são nossa VI. Por isto, podemos comparar o efeito da VI no nosso grupo experimental com o efeito dos vieses no grupo controle. Se o nosso grupo experimental obteve uma melhora igual ao do grupo controle, então concluímos que não foi a nossa terapia a causa da melhora, pois o nosso grupo controle também melhorou e, no entanto, não foi submetido a ela.

2. Igualar os grupos por meio da distribuição aleatória dos sujeitos aos grupos: Esta é a melhor maneira de controlar os vieses que ameaçam a validade interna de uma pesquisa, aceita pela comunidade científica. Ao distribuímos aleatoriamente (por sorteio) os sujeitos aos grupos, nós escolhemos ao acaso o grupo para o qual cada sujeito será designado. Ou seja, para cada sujeito, tiramos a sorte para decidir se ele irá para o grupo experimental ou para o grupo controle. Com isto, estamos distribuindo aleatoriamente as variáveis individuais, de modo que estas variáveis fiquem igualmente distribuídas em um e outro grupo e, desta maneira, estamos tornando os grupos equivalentes. Por exemplo, após fazermos a distribuição aleatória dos sujeitos teremos um grupo experimental com sujeitos de várias características sócio-demográficas, clínicas, biológicas, etc. mas também o grupo controle será igualmente composto por sujeitos com variadas características. Ambos os grupos serão igualmente variados e heterogêneos, pois a distribuição aleatória garante que os grupos terão uma composição variada, mas semelhante de variáveis. Os grupos se tornam equivalentes. Portanto, ficaremos mais seguros em afirmar que as diferenças de resultados que encontrarmos na pesquisa não serão resultantes de diferenças iniciais entre os grupos, mas sim o resultado da nossa variável independente.

3. Tentativa de igualar os dois grupos sem distribuição aleatória: Esta é uma maneira que não é eficaz para igualar os grupos. O pesquisador procura igualar os grupos com relação a algumas variáveis que ele considera importantes, tais como sexo, idade, escolaridade. Porém, há um grande número de outras variáveis (sócio demográficas, clínicas, biológicas) que também podem afetar os resultados. Ele não consegue igualar os grupos em relação a todas as variáveis que possam ser relevantes para a pesquisa. Aliás, ele pode até desconhecer quais são elas. Além disso, ao igualar os grupos em relação a algumas variáveis ele pode estar deixando os grupos ficarem diferentes em relação a diversas outras variáveis. Deste modo, usando este procedimento, o pesquisador poderá apenas concluir que os resultados obtidos na sua pesquisa não estão sendo afetados pelas poucas variáveis que ele conseguiu igualar nos grupos, mas será afetado pelas outras variáveis que ele não igualou nos grupos.

4. Outros procedimentos: Para aumentar a validade interna de uma pesquisa, dois outros procedimentos são utilizados, especialmente em pesquisas de avaliação dos efeitos de um tratamento (ex. efeitos de um novo medicamento). São eles: a) Usar um “grupo placebo” para controlar o “efeito placebo” de um tratamento (ver descrição dentro do subtítulo de Viés Testagem) b) usar o procedimento “duplo cego” durante a liberação do novo tratamento (idem)

Observação: Os vieses se referem a variáveis que podem afetar o efeito da VI sobre a VD. A VI pode ser uma intervenção ou uma variável natural. Para efeito didático, vamos nos referir muitas vezes à VI como sendo uma intervenção. No entanto, as VI se referem a muitas outras variáveis além de uma intervenção.

VIESES QUE AMEACAM A VALIDADE INTERNA:

São oito os principais vieses que ameaçam a validade interna de uma pesquisa

1A. História:

É o viés provocado pela ocorrência de um evento externo simultaneamente à ocorrência da VI e que irá influenciar a variável dependente durante o desenrolar do experimento. As mudanças observadas na nossa VD poderão, neste caso, ter sido causadas tanto pela nossa VI, quanto pelo evento história. Se não utilizarmos nenhum procedimento de controle ou um bom delineamento, não saberemos se as mudanças observadas foram provocadas pela VI ou por este viés. Há vários tipos de eventos que podem provocar este viés, tais como:

Eventos sociais: Suponhamos que você esteja fazendo uma pesquisa sobre a atitude das pessoas a respeito dos negros. Você quer verificar se as pessoas mudam de atitude após ver um filme positivo sobre negros. Para testar esta hipótese, você faz um pré-teste (O1) medindo a atitude de seus sujeitos através de um teste de atitude. Em seguida, você os expõe ao filme, que é seu tratamento ou sua variável independente (que chamaremos de X). Depois, você mede outra vez a atitude dos sujeitos, aplicando-lhes o mesmo teste de atitude, em um pós-teste (O2) para ver se mudaram de atitude. Teremos então o seguinte esquema: O1-X-O2. Neste tipo de pesquisa, é fácil imaginar um evento social capaz de enviesar nossos resultados. Por exemplo, suponha que entre O1 e O2 ocorra uma calamidade na cidade e um indivíduo negro se destaca como herói que salvou pessoas, este evento tendo sido noticiado nos jornais e comentado pelas pessoas. Este evento constitui um viés história que, ocorrendo simultaneamente ao tratamento (X), pode afetar nossos resultados e não saberemos se foi ele a causa das mudanças positivas obtidas em O2. O contrário também poderia ocorrer, ou seja, um evento social negativo praticado por um negro, por exemplo, um crime, poderia afetar nossos resultados no sentido oposto, ou seja, no sentido de diminuir o efeito positivo de nosso filme. Também neste caso não saberíamos se nossa VI não teve efeito porque é ineficaz ou devido à interferência deste viés história.

Efeito da estação do ano: O início de uma estação pode provocar mudanças no humor das pessoas. Se este evento acontecer simultaneamente à ocorrência de nossa VI não saberemos se foi este viés a causa das mudanças que observamos ou se a causa foi nossa VI.

Efeito de eventos institucionais: eventos tais como exames finais ou periódicos, proximidade das férias, greves, eleições, etc. afetam as pessoas e se ocorrerem simultaneamente à introdução de nossa VI, poderão confundir nossos resultados.

Ciclos da semana: a chegada do fim de semana pode afetar as pessoas, provocando mudanças em seu humor, aumentando ou diminuindo a frequência do comportamento observado (VD). Se deixarmos que nossa VI seja introduzida no fim de semana ou no início da semana, então as mudanças obtidas na VD poderão ter sido causadas por este viés história.

Uma maneira de controlar o viés história é utilizar um grupo controle (GC), porque se um destes eventos espúrios ocorrerem ao mesmo tempo em que nossa VI, este evento afetará não só nosso grupo experimental (GE), mas ele afetará também nosso grupo controle. Então veremos que o grupo controle também apresenta mudanças na mesma direção do outro grupo. Neste caso,

saberemos que houve um viés história e saberemos também a intensidade de seu efeito. Se o GE apresentar mudanças mais fortes do que o GC no pós-teste, isto nos indicaria que o GE sofreu efeito tanto do viés história quanto da nossa VI. A diferença entre o efeito obtido no GC e o efeito obtido no GE nos dará a dimensão do efeito da VI. Se, ao contrário, o grupo experimental apresentar o mesmo grau de mudança em O2 que o grupo controle então saberemos que o efeito observado é um viés e não o efeito do tratamento.

1B. História interna:

Entretanto, há um viés parecido com este que não estará controlado com a utilização do grupo controle. Trata-se do viés “história interna”. Dizemos “interna” para nos referirmos a eventos que ocorrem diferentemente no GE e no GC. Por exemplo, ao fazermos nossa medida da VD nos nossos dois grupos, após o tratamento (VI), é possível que ocorra um evento durante esta sessão do GE que não ocorre no GC, como por exemplo, uma piada engraçada do experimentador que não foi dita no outro grupo ou instruções mais detalhadas para o GE do que para o GC, ou o horário da sessão que pode favorecer o GE e desfavorecer o GC. Se um destes eventos tem o poder de afetar o comportamento dos sujeitos e se ele ocorre somente no grupo experimental, favorecendo-o, então estará acontecendo o viés história interna e o fato de termos um grupo-controle não serviu para evitar este viés.

A melhor solução para controlar este viés é testar cada sujeito individualmente, separadamente. Não se deve fazer a sessão para todos os sujeitos de um grupo juntos ao mesmo tempo e depois para todos os sujeitos do outro grupo. Mesmo ao fazermos a medida da VD individualmente, não devemos avaliar os sujeitos de um grupo e depois os sujeitos do outro grupo, mas sim distribuídos ao acaso. A melhor maneira de controlar este viés é designar sujeitos e os períodos ou horários aleatoriamente a duas condições experimentais.

2. Maturação:

Este é um viés ocasionado pela influência de variáveis ligadas a processos internos aos sujeitos, ou seja, processos biológicos ou psicológicos que ocorrem entre O1 e O2, independentemente da variável independente. Estes processos podem provocar mudanças nos sujeitos que podem interferir com o efeito do tratamento, se elas ocorrerem simultaneamente à VI.

Por exemplo, se estivermos investigando o efeito de uma intervenção terapêutica no grau de depressão de um grupo de sujeitos e se ocorrer remissão espontânea de uma parte do grupo, o resultado médio do grupo no pós-teste pode ser bem melhor do que no pré-teste, como resultado desta remissão espontânea e não como efeito de nossa intervenção.

Em outras palavras, mesmo se não fosse introduzida nenhuma intervenção terapêutica neste grupo de sujeitos, poderia ocorrer uma melhora no seu grau de depressão por um processo natural de evolução dos sujeitos em suas vidas. Este efeito de melhora natural (remissão espontânea) é bem conhecido, podendo ocorrer em pacientes, por exemplo, que se encontram em lista de espera para receber tratamento psicoterápico.

Outros processos internos podem ocorrer durante uma pesquisa, simultaneamente à introdução da variável independente, de modo a interferir com a interpretação de nossos dados. As principais fontes de influência que podem acarretar no viés maturação são:

1. O desenvolvimento motor ou cognitivo dos sujeitos que ocorre com a idade, quando os sujeitos da pesquisa são crianças e estamos pesquisando o efeito de intervenções que envolvem estes processos;
2. O processo de envelhecimento dos sujeitos, quando os sujeitos são idosos e o efeito do envelhecimento pode mascarar o efeito da variável independente.
3. Processos de fadiga, sono ou fome durante um experimento que duram algumas horas, quando o efeito da VI pode ser confundido com os efeitos destes processos.
4. Mudanças no nível de motivação dos sujeitos durante a introdução da VI, quando uma baixa de motivação pode interferir com o efeito da VI.
5. A própria experiência adquirida pelo sujeito em sua vida, quando este efeito tem probabilidade de interferir com o efeito da VI.

Para controlarmos o efeito do viés maturação, é necessário utilizar um delineamento que inclui um grupo de controle. Neste caso, o viés maturação ocorrerá nos dois grupos igualmente, e assim poderemos interpretar os resultados obtidos, sem haver o perigo de nos confundirmos com a interferência deste viés. Mudanças ocorridas no grupo experimental que forem significativamente mais acentuadas do que as possíveis mudanças ocorridas no grupo de controle poderão ser interpretadas como sendo resultado da VI e não do viés maturação. Neste caso, o grau da mudança obtida no grupo experimental será muito mais acentuado do que o grau de mudança ocorrida no grupo de controle, pois esta última seria provocada apenas pelos fatores ligados à maturação.

3. Testagem:

Este viés ocorre quando os sujeitos aprendem a responder melhor ou pior a um questionário aplicado no pré-teste e depois no pós-teste, para avaliar os efeitos da VI na VD. Este viés tem maior probabilidade de ocorrer quando há muitas aplicações de um mesmo questionário ou quando o intervalo entre as aplicações é muito pequeno.

Por exemplo, a aplicação repetida de um teste de inteligência pode levar a uma melhora nos resultados devido à aprendizagem dos sujeitos com este teste. Os testes de personalidade, quando repetidos, provocam geralmente um melhor resultado, mas ocasionalmente podem provocar um resultado pior. Já os questionários de atitude podem apresentar, em sua repetição, uma leve evidência de aumento no preconceito (Campbell e Stanley, 1987).

Um efeito semelhante ao citado acima ocorre no caso das medidas reativas, por exemplo, quando fazemos uma auto-observação para medir o número de cigarros que fumamos por dia ou o número de itens extra de alimentos que ingerimos por dia. Outro exemplo de medida reativa é a presença de um observador para avaliar, por exemplo, o comportamento dos sujeitos em uma sala de aula. A utilização de microfones ou câmeras para registrar o comportamento dos sujeitos também pode produzir um efeito de interferência sobre os resultados.

O controle do viés testagem pode ser feito utilizando-se um grupo de controle, pois havendo dois grupos de sujeitos na pesquisa o viés testagem ocorrerá em ambos os grupos igualmente. Neste caso, podemos interpretar os resultados mais acentuados do grupo experimental ao efeito da VI e não ao efeito deste viés.

Um efeito semelhante ao do viés testagem, geralmente considerado dentro deste item, consiste no efeito placebo de uma intervenção. O efeito placebo se refere à influência positiva exercida por uma intervenção, pelo simples fato dos sujeitos estarem sendo submetidos a um tratamento, mesmo quando este tratamento não possui nenhum ingrediente ativo. Por exemplo, se estivermos testando o efeito de um novo medicamento, o simples fato dos sujeitos estar tomando um remédio pode provocar efeitos positivos, mesmo se este remédio for ineficaz. Para controlar este viés específico, não basta utilizar um grupo de controle comum, ou seja, aquele grupo que não é submetido a nenhum tratamento.

Para controlar o efeito placebo é necessário utilizar, além disso, um grupo de controle especial, chamado de “grupo placebo”, ou seja, aquele em que os sujeitos recebem uma intervenção placebo. Uma intervenção placebo é aquela intervenção que não possui nenhum ingrediente ativo. Por exemplo, no caso da pesquisa que testa um novo medicamento, deveríamos ter um grupo de controle placebo que recebesse um remédio com as mesmas características de cor, odor e sabor do remédio novo dado ao grupo experimental, mas sem nenhum ingrediente ativo (simplesmente água colorida). Evidentemente, é necessário que os sujeitos deste grupo de controle placebo não percebam a diferença entre os dois medicamentos, o verdadeiro e o placebo. Deste modo, qualquer efeito placebo do novo medicamento ocorrerá igualmente nos dois grupos, no grupo experimental e neste grupo de controle placebo. Neste caso, se o grupo experimental apresentar um resultado significativamente mais acentuado do que o resultado deste grupo placebo podemos ter certeza de que é devido ao efeito do ingrediente ativo do novo medicamento.

Além disso, é importante que as pessoas que vão administrar os dois tipos de medicamento também não percebam a diferença entre eles, a fim de não comunicarem indevidamente esta informação aos sujeitos. É o que se chama usualmente de procedimento do duplo cego, ou seja, tanto os sujeitos quanto os aplicadores estão cegos quanto ao tipo de intervenção aplicada.

4. Instrumentação:

Chamamos de instrumentação o viés que ocorre quando nosso instrumento de medida sofre mudanças entre o pré-teste e o pós-teste, de modo a interferir com os resultados de nossa variável independente. Não saberemos então se a mudança obtida em nossa VD se deve ao efeito da intervenção ou se deve à interferência destas modificações do instrumento de medida.

Este viés pode ser provocado, por exemplo, por mudanças ocorridas na calibragem de um instrumento de medida entre o pré-teste e o pós-teste, tais como a distensão de uma balança. Quando utilizamos observadores como instrumentos de medida para registrar, por exemplo, a frequência de um comportamento dos sujeitos, estes observadores podem sofrer mudanças também, tais como uma maior familiaridade ou habituação com o processo de medida ou mudanças de expectativa entre o pré-teste e o pós-teste, que consideraremos como mudanças no instrumento de medida.

Para controlar este viés, não basta utilizarmos um grupo de controle, porque não podemos ter certeza de que este viés ocorrerá igualmente nos dois grupos. Por exemplo, se o observador

sabe quais sujeitos pertencem ao grupo experimental e se ele conhece a intervenção a ser utilizada, sua expectativa de que esta intervenção irá aumentar a frequência do comportamento observado pode afetar sua forma de observar e medir. Sua expectativa pode induzi-lo inconscientemente a prestar mais atenção e a observar uma maior frequência de comportamentos após a intervenção, só no grupo experimental. Seria necessário, portanto, garantir que o observador não saiba se está observando o grupo de controle ou o grupo experimental, nem se o grupo observado está em período de pré ou pós-teste.

O mesmo problema teríamos se nossa VD fosse uma prova de conhecimentos a ser aplicada antes e após a intervenção. O indivíduo que irá corrigir estas provas também poderá sofrer influência de suas próprias expectativas a respeito da intervenção, resultando em uma correção diferente no pré e no pós-teste. Neste caso, deveríamos misturar todas as provas, de modo que a pessoa que as corrige não saiba se trata de uma prova do grupo experimental ou do grupo de controle e também não saiba se se trata de uma prova aplicada antes ou após a intervenção.

Para evitar que o cansaço dos observadores afete sua maneira de observar e medir os comportamentos, e que isto favoreça um ou outro grupo de sujeitos, devemos distribuir aleatoriamente os observadores para os diversos horários de observação e para os diversos sujeitos a serem observados. Assim, os fatores cansaço, mudança de motivação ou mudança na familiaridade dos observadores estarão distribuídos igualmente para os diversos sujeitos dos dois grupos. Deste modo, os fatores que induzem falhas destes instrumentos de medida estarão controlados.

5. Mortalidade Experimental:

Este viés estará presente quando a taxa de abandono ou evasão dos sujeitos alterar a composição do(s) grupo(s), de forma que o resultado obtido no pós-teste seja decorrente desta mudança e não do efeito da VI. No caso de termos um delineamento com dois grupos de sujeitos, pode ocorrer, por exemplo, uma evasão diferencial dos sujeitos nos grupos GE e GC, de forma a favorecer o GE nos resultados finais. Ou seja, pode ser que os melhores sujeitos do GE permaneçam na pesquisa e os piores sujeitos desistam e o contrário ocorrer no GC. Neste caso, os resultados superiores do GE no pós-teste seriam decorrentes deste viés e não do efeito da VI. No caso de termos um delineamento de um só grupo na nossa pesquisa, pode ocorrer que a evasão dos sujeitos durante o experimento faça com que os piores sujeitos do grupo desistam, ficando apenas os melhores. Neste caso, o pós-teste apresentará resultados melhores não pelo efeito da VI, mas simplesmente por causa deste viés.

Para diminuir o efeito deste viés, podemos, por exemplo, solicitar os sujeitos que abandonaram o tratamento para participarem assim mesmo do pós-teste (02). Deste modo teremos resultados conservadores no pós-teste, o que significa que o efeito real da VI seria ainda mais forte do que parece. No caso de não termos mais nenhum acesso aos sujeitos que abandonaram a pesquisa, o que devemos fazer é comparar os dados de pré-teste dos sujeitos que abandonaram a pesquisa com os dados de pré-teste dos sujeitos que permaneceram na pesquisa, para verificarmos se há ou não diferença significativa entre estes dois grupos de sujeitos, com relação a variáveis pertinentes para os resultados do efeito da VI, tais como as variáveis individuais e clínicas.

Por exemplo, ao verificarmos se um grupo de 40 doentes mentais desinstitucionalizados, que foram devidamente preparados para tal, apresentam resultados melhores na sua reinserção social na comunidade do que um grupo de doentes mentais que não foram preparados para a

reinserção social, observamos a evasão de 5 sujeitos do grupo experimental, uns por morte natural, outros por não quererem participar. Neste caso, nós devemos observar se estes sujeitos diferem dos demais, com relação a variáveis clínicas, tais como tempo de hospitalização, duração da doença, diagnóstico, número de rehospitalizações e outras variáveis que podem afetar os resultados do tratamento. Para isto, devemos computar os dados e aplicar um teste estatístico de significância. A diferença entre os sujeitos que ficaram e os que não ficaram deverá então ser relatada, e os resultados do tratamento devem ser discutidos em função deste fato. Se os que saíram não apresentam diferenças significativas em relação aos que ficaram, então a análise dos resultados se faz normalmente. Caso contrário, devemos ter cautela ao concluirmos a respeito do efeito de nossa VI, nos reportando sempre ao viés mortalidade experimental.

Um problema semelhante ocorre quando recrutamos sujeitos voluntários para um tratamento e apenas alguns aceitam participar. Por exemplo, suponhamos que 100 sujeitos são escolhidos em um ambulatório para participar de uma pesquisa sobre o efeito de um novo tratamento para a depressão, enquanto outros 100 servirão de grupo controle avaliados normalmente pela rotina de avaliação do ambulatório. Entretanto, podemos constatar que dos 100 que foram convidados para participar do GE, somente 30 sujeitos aceitam participar da pesquisa. Como faremos para selecionar 30 sujeitos do grupo controle dentre os 100 que foram selecionados para este grupo GC? Quais sujeitos do grupo controle deverão compor o grupo dos 30? Também neste caso, devemos comparar as características dos sujeitos que aceitaram com as dos sujeitos que não aceitaram. Para evitar que este problema ocorra, Stanley e Campbell (1966) sugerem que, neste caso, devemos proceder de maneira diferente do que foi exposto acima. Devemos primeiro convidar a totalidade dos 200 sujeitos que farão parte tanto do grupo controle quanto do grupo experimental. . Após verificarmos quantos aceitam participar, só então devemos dividir os grupos em dois, GE e GC.

6. Regressão estatística em direção à média:

Este viés ocorre quando escolhemos sujeitos com escores extremos no pré-teste para fazer parte de nossa pesquisa, principalmente quando temos apenas um delineamento de um só grupo. Por exemplo, suponhamos que escolhemos os piores alunos para participar de uma pesquisa sobre o efeito de um novo método de ensino. Estes foram escolhidos com base nas piores notas obtidas em um pré-teste que fizemos. Acontece que os escores extremos de um teste, os piores ou os melhores, ou seja, aqueles valores que diferem demais da distribuição dos demais valores da amostra, são considerados como erros aleatórios positivos ou negativos.

Se repetirmos a aplicação deste mesmo teste na nossa amostra algum tempo mais tarde, observaremos que aqueles sujeitos que tiveram os valores positivos mais extremos do grupo tenderão a obter valores menores da segunda vez. Por outro lado, aqueles sujeitos que tinham alcançado os valores negativos mais extremos do grupo, tenderão a obter notas maiores na segunda vez. Ou seja, existe uma tendência a uma regressão dos valores extremos em direção à média do grupo. Este fenômeno, segundo os estatísticos, é provocado pela ocorrência de uma correlação imperfeita teste-reteste do instrumento de medida usado. É impossível termos um instrumento de medida que possua correlação teste-reteste igual a 1.0, que seria a correlação perfeita. Quanto menor a correlação teste-reteste de um instrumento de medida, maior a probabilidade de ocorrer este viés de regressão em direção à média. Este fenômeno da regressão à média ocorre somente quando nós selecionamos deliberadamente nossos sujeitos com base nos piores ou nos melhores escores do pré-teste.

No caso de usarmos um delineamento de dois grupos em nossa pesquisa, o fenômeno da regressão à média estaria controlado, na medida em que a regressão iria ocorrer igualmente no GE e no GC. Porém, nós só poderemos ter certeza que a regressão ocorrerá igualmente nos dois grupos, se eles forem retirados aleatoriamente da população. Porque somente neste caso é que os dois grupos serão representativos da população, ou seja, terão a mesma distribuição de valores da população e, portanto, terão uma distribuição equivalente de valores, se comparado um com o outro. Sendo equivalentes, a regressão ocorrerá da mesma forma em um e em outro grupo.

Campbell e Stanley (1966) advertem contra uma prática errônea dos pesquisadores quando analisam subgrupos dentro de um grupo experimental que teve valores extremos. Por exemplo, não se pode tomar um subgrupo de sujeitos do GE para verificar se aqueles que foram os piores no pré-teste foram os que melhoraram mais no pós-teste. Esta análise não é válida, segundo estes autores. O que se pode fazer, no máximo, é apenas comparar este subgrupo do GE com o subgrupo equivalente do GC.

7. Seleção:

Este é um viés ocasionado por diferenças que podem existir desde o início entre o grupo experimental e o grupo controle, quando se usa um delineamento de dois grupos não equivalentes. Esta diferença inicial poderia ser então a causa dos resultados no pós-teste e não o efeito da nossa VI. Em outras palavras, mesmo se não tivéssemos aplicado nenhum tratamento, os grupos dariam resultados diferentes no pós-teste, devido à suas diferenças iniciais.

Por exemplo, suponhamos que queremos avaliar dois grupos de pacientes alcoolistas que já estão sendo tratados por dois tratamentos diferentes, um tradicional e outro com atividades programadas em um esquema de hospital-dia. Os sujeitos haviam se oferecido para participar do novo tratamento alternativo, compondo assim o GE e os demais haviam ficado compondo o grupo controle. Portanto, os dois grupos não eram equivalentes. Neste caso, é possível que os sujeitos que se ofereceram para participar do grupo experimental sejam, em sua maioria, sujeitos mais motivados e com menor grau de severidade do alcoolismo do que o outro grupo. Neste caso, eles seriam diferentes desde o início e esta diferença favoreceria o GE. Ao aplicarmos um pós-teste para comparar os dois grupos, não poderíamos saber se nosso GE apresenta melhores resultados devido ao tratamento alternativo ou devido ao fato de ser um grupo diferente, mais propício a uma melhora.

Este viés só pode ser controlado se, no início do experimento, nós distribuímos os sujeitos aleatoriamente aos GE e GC, não deixando assim a composição dos grupos ao desejo dos sujeitos. Esta distribuição aleatória serve para nos assegurar de que os dois grupos serão equivalentes um ao outro com relação à distribuição das principais variáveis relevantes à avaliação do tratamento. Este é o procedimento mais frequentemente utilizado pelos pesquisadores, porém não é considerado pelos estatísticos como o recurso mais perfeito para controlar o viés seleção, porque o acaso também pode permitir, embora raramente, uma distribuição desigual de variáveis importantes. O procedimento mais recomendado pelos estatísticos para garantir a equivalência dos grupos é primeiro igualarmos todos os sujeitos da pesquisa com relação às variáveis mais importantes, através do emparelhamento dos casos. Depois, então, distribuímos aleatoriamente estes sujeitos aos GE e GC, de modo a garantir a distribuição equitativa de todas as variáveis entre os dois grupos.

8. Interação seleção-maturação ou outras:

Pode haver também interação entre os vieses expostos acima. Por exemplo, o viés seleção pode interagir com o viés maturação, provocando um viés particular desta interação. Neste caso, a maturação ocorrerá diferentemente no GE comparativamente ao que ocorrerá no GC. Por exemplo, suponhamos que queremos avaliar o efeito de um novo tratamento para a depressão. Os sujeitos se ofereceram voluntariamente para participar do tratamento e serão comparados aos sujeitos-controle, formando assim dois grupos não-equivalentes. É provável que os sujeitos que se ofereceram sejam diferentes dos demais desde o início podendo, por exemplo, ser mais motivados e apresentarem um grau de depressão menos severa. Além disso, é provável também que a evolução da depressão em um e outro grupo ocorra de forma diferente. Neste caso, podemos supor facilmente que estes sujeitos possam apresentar uma remissão espontânea maior do que o outro grupo. Ou seja, pode se tratar de um grupo de sujeitos que teriam de qualquer forma evoluído melhor do que o outro grupo. Assim os dois vieses estariam combinados em um efeito conjunto de interação, que favoreceria ainda mais o GE.

Uma maneira de controlar este viés é utilizar o delineamento de séries temporais múltiplas, no qual todos os dois grupos são avaliados várias vezes antes e depois do tratamento, como se fossem duas séries temporais (duas séries A-B). Por exemplo, todos os sujeitos seriam avaliados todas as semanas, durante 6 semanas. Depois do tratamento, o mesmo procedimento seria feito. Neste caso, poderíamos observar a evolução de cada grupo separadamente. Se houver alguma tendência a uma maturação diferente no GE do que no GC, devido ao viés interação, poderemos notar isto desde o início, nas 6 primeiras medidas anteriores ao tratamento.

VALIDADE EXTERNA

A validade externa de uma pesquisa se refere, como dissemos no início deste texto, à capacidade de podermos generalizar os resultados, obtidos com a nossa amostra, à população-alvo de onde esta amostra foi retirada. O que queremos, ao realizar uma pesquisa, é descobrir relações causais que possam ser verdadeiras, não só para a nossa amostra, mas para as pessoas da população-alvo de onde retiramos a nossa amostra, senão nossas conclusões seriam muito limitadas e, portanto, de pouco interesse.

Alguns vieses, quando não controlados, provocam esta limitação da validade externa da pesquisa, ou seja, conduzem a uma grande especificidade dos resultados obtidos àquele tipo de amostra estudada ou ao contexto específico da pesquisa ou a algum conjunto indesejável de condições específicas daquele experimento.

Existe um procedimento básico para se aumentar a validade externa de uma pesquisa, que é selecionar os sujeitos que vão compor a nossa amostra aleatoriamente da população-alvo, para que a nossa amostra seja representativa desta população (ver diagrama 2). Somente com este procedimento, evitaremos ter uma amostra particular cujos resultados fiquem restritos a ela.

VIESES QUE AMEAÇAM A VALIDADE EXTERNA:

São quatro os principais vieses que ameaçam a validade externa de uma pesquisa:

1. Interação entre o pré-teste e a VI (testagem-X):

Este viés ocorre quando o pré-teste produz um efeito sobre os sujeitos, no sentido de torná-los mais sensíveis ou menos sensíveis à variável experimental. O fato de se submeterem ao pré-teste pode aumentar ou diminuir a sensibilidade ou capacidade dos sujeitos à intervenção utilizada. Neste caso, o pré-teste terá modificado os sujeitos e, portanto, eles não serão mais representativos da população. Os resultados que obtivermos não serão, portanto, generalizáveis à população-alvo, pois esta é composta de sujeitos que não foram pré-testados. Em outras palavras, o efeito de nossa VI fica limitado a sujeitos que tenham sido submetidos antes a um pré-teste. Neste caso, não podemos ter certeza que nossa intervenção poderia ter algum efeito em sujeitos que não se submeterem a um pré-teste.

Por exemplo, consideremos o exemplo citado anteriormente referente à pesquisa sobre atitudes em relação aos negros. A pesquisa visa testar o efeito de um filme a favor dos negros sobre o grau de racismo de uma amostra de sujeitos. Para isto, aplicamos um questionário de atitude sobre os negros em nossa amostra (pré-teste) para avaliarmos seu grau de racismo inicial. Em seguida, submetemos nosso GE a assistir um filme de romance que mostra uma imagem muito positiva dos negros. Após esta intervenção, aplicamos de novo o mesmo questionário de atitude (pós-teste) para verificarmos se a atitude da amostra em relação aos negros se tornou mais positiva em consequência daquele filme.

Suponha, entretanto, que o fato de ter respondido ao pré-teste, fez com que nossos sujeitos despertassem para a questão do racismo, o que não teria acontecido se simplesmente tivéssemos utilizado o filme, sem nenhum pré-teste. Neste caso, despertando para esta questão, os sujeitos irão assistir ao filme com uma outra perspectiva, refletindo claramente sobre a questão do racismo. Depois, ao fazermos o pós-teste, os resultados podem mostrar um efeito do filme que talvez não tivesse ocorrido se os sujeitos não tivessem sido despertados explicitamente para a questão do racismo.

Neste caso, ficaremos com uma dúvida muito importante: será que nossos resultados seriam os mesmos se nossos sujeitos não tivessem feito o pré-teste? Talvez não. E se esta dúvida persiste, então não poderemos concluir genericamente, que o nosso filme terá um efeito benéfico para pessoas que não forem pré-testadas. Temos então um viés que afeta a validade externa de nossa pesquisa.

Uma solução para este problema consiste na utilização do delineamento experimental de Solomon (Campell), no qual temos 4 grupos, dois grupos que são submetidos a um pré-teste (GE e GC) e dois grupos que não são submetidos a um pré-teste (GE e GC). Neste caso, teremos dois tipos de resultados em nossa pesquisa: um resultado do efeito da VI em grupo pré-testado e um resultado da VI em um grupo que não foi pré-testado. Saberemos então se nossa intervenção é eficaz não só para grupos pré-testados, mas se ela é eficaz também para grupos que não foram pré-testados.

Este viés ocorre mais frequentemente em situações de pesquisa em que o pré-teste é de grande impacto, seja por incluir testagens insólitas, seja porque o processo de testagem envolve stress para os sujeitos ou os faz refletir e realizar uma reestruturação perceptual ou cognitiva.

2. Interação entre seleção e a intervenção (seleção-X):

Quando uma amostra é particular ou muito especial, os resultados de nossa pesquisa ficarão restritos àquela população específica de pessoas, da qual a nossa amostra foi retirada. Por exemplo, se utilizarmos sujeitos voluntários para realizarmos nossa pesquisa, as nossas conclusões a respeito do efeito de nossa VI só se aplicarão às pessoas voluntárias e não à população em geral. Os voluntários são pessoas de certo modo especiais, pois são pelo menos mais motivadas do que as demais, ou são pessoas até mais temerárias do que a maioria, no caso, por exemplo, do experimento envolver situações inusitadas.

Um outro exemplo consiste nas pesquisas avaliativas em geral, cujo objetivo é avaliar a eficácia de um procedimento pedagógico utilizado em uma escola, ou de um procedimento terapêutico utilizado em uma instituição de saúde ou de saúde mental. Se a maioria das instituições evita participar devido ao desconforto de estarem em uma situação em que se sentirão avaliados, as poucas instituições ou equipes que aceitarem participar da pesquisa pertencerão provavelmente a uma classe particular de pessoas altamente motivadas, que não são representativas das pessoas dos outros serviços.

Este viés tem maior probabilidade de ocorrer em pesquisas que envolvem situações particulares, por exemplo, a utilização de estímulos aversivos durante a experimentação, tais como choque elétrico ou térmico, isolamento social, privação de sono ou de outros elementos, utilização de drogas, etc. Situações de desconforto psicológico, tais como uma avaliação das pessoas ou dos métodos empregados por elas. Além disso, nos casos em que o recrutamento dos sujeitos para a pesquisa envolveu muita dificuldade de conseguir a aceitação dos sujeitos, temos grande chance de encontramos este viés. A dificuldade em conseguir a participação de sujeitos ou de instituições na nossa pesquisa constitui um indício de que estamos utilizando algum elemento inusitado, difícil ou peculiar que dificulta a aceitação das pessoas em participar. Isto constitui um indício de que nosso contexto de pesquisa é peculiar e nossa validade externa estará, portanto ameaçada.

No caso de não podermos evitar a peculiaridade de nossa pesquisa, devemos pelo menos tecer considerações a respeito deste viés no nosso projeto de pesquisa, no relatório da pesquisa ou na publicação de um artigo sobre os dados da pesquisa. Devemos deixar claro que este viés pode estar presente. Neste caso, devemos especificar quantos e quais sujeitos foram recrutados, quantos e quais aceitaram e quantos recusaram. Isto servirá para fornecer uma idéia do grau de peculiaridade da situação da pesquisa e da dimensão deste viés de validade externa. Quanto ao controle real deste viés, o único procedimento eficaz consiste na utilização de sujeitos não voluntários, mas sim sujeitos selecionados aleatoriamente da população-alvo.

3. Efeitos reativos da situação experimental:

Este viés ocorre quando a situação de realização da pesquisa apresenta um certo grau de artificialidade que difere da situação natural. Neste caso, não podemos generalizar os resultados da nossa intervenção a pessoas que seriam expostas a ela no ambiente natural. Os resultados obtidos por nós na pesquisa ficariam circunscritos a situações artificiais, tais como a utilizada na pesquisa. Ou seja, o efeito da nossa VI só deverá ocorrer, ou pelo menos naquela intensidade observada, no caso de se utilizar a intervenção na mesma situação artificial.

Por exemplo, quando utilizamos uma determinada intervenção no contexto de um laboratório, podemos colocar em dúvida se aquela intervenção também será eficaz no ambiente natural. Talvez não. Um outro exemplo é quando utilizamos uma nova intervenção terapêutica em uma unidade de um hospital de saúde mental. Antes da implantação do novo procedimento em uma das unidades, em caráter experimental, cria-se uma expectativa em torno do novo tratamento, tanto no meio dos pacientes, quanto no meio dos terapeutas, assistentes e toda a equipe que compõe aquela unidade. Geralmente esta expectativa se estende até para as demais unidades e o fato é comentado em larga escala na instituição. Isto cria um clima de excepcionalidade e entusiasmo a respeito do novo tratamento que o torna artificial. Talvez se não houvesse este clima particular, se estivéssemos avaliando o tratamento em questão em uma situação mais natural, o novo tratamento não teria o mesmo efeito ou então não teria efeito nenhum. Neste caso, a validade externa da pesquisa estaria, portanto ameaçada e não poderíamos generalizar nossos resultados e nossas conclusões às situações mais naturais de utilização daquela intervenção.

Para controlar este viés torna-se necessário evitar que as pessoas tenham conhecimento da introdução da nova intervenção, camuflando-a de alguma maneira. Na área educacional isto se torna mais fácil, pois a intervenção pode ser introduzida como uma rotina e as avaliações de pré-teste e pós-teste podem ser as próprias notas mensais ou semestrais, o que já ocorre normalmente.

No caso de pesquisas avaliando o efeito de tratamentos medicamentosos, todos os envolvidos devem estar ignorantes do fato de que um novo tratamento está sendo utilizado. Para isto, é preciso utilizar os envoltórios usuais dos medicamentos, porém com o conteúdo novo, sem que as pessoas envolvidas (sujeitos e aplicadores) percebam a mudança. No caso do tratamento terapêutico global em uma unidade de saúde mental, a questão se torna mais difícil de ser contornada, mas todo cuidado possível para inserir o novo tratamento como uma rotina a mais do hospital servirá para diminuir a influência deste viés de validade externa.

4. Interferência de tratamentos múltiplos:

Quando a pesquisa possui mais de um tratamento ou variável independente, é possível que haja interferência do efeito de um tratamento sobre o outro e, neste caso, temos um problema sério, pois não poderemos conhecer o efeito isolado de um tratamento sobre a nossa variável dependente. Por exemplo, se queremos testar um novo medicamento para saber se seu efeito é melhor do que o de um outro medicamento sobre o grau de depressão dos sujeitos, ao submetermos os mesmos sujeitos aos dois medicamentos, consecutivamente, para compararmos seus efeitos, poderá haver uma contaminação de um sobre o outro.

Neste caso, os efeitos do segundo medicamento serão alterados pela ingestão anterior do primeiro medicamento. Os resultados desta pesquisa não poderão ser generalizados para situações em que cada medicamento em questão será dado isoladamente, sem ser precedido pelo outro.

Uma maneira de controlar esta interferência seria ter dois grupos e em cada um deles introduzir uma ordem diferente de ingestão, de modo que um grupo receba o medicamento 1 e depois o medicamento 2 e o outro grupo receba a combinação inversa. O efeito da ordem de apresentação ficaria sob controle, se este procedimento puder ser aplicado.

Referências Bibliográficas:

Contandriopoulos, Ap., Champagne, F., Potvin, L., Denis, J.L. E Bouyle, P. (1990) **Savoir préparer une recherche**. Montréal: Les Presses de l'Université de Montréal.

Campbell, D.T., Stanley, J.C. (1979) **Delineamentos experimentais e quase-experimentais de pesquisa**. Editora da Universidade de São Paulo, SP.

Selltiz, C., Wrightsman, L.S., Cook, S.W. (1987) **Métodos de pesquisa nas relações sociais**. Editora da Universidade de São Paulo, SP.