

PESQUISA EXPERIMENTAL EM MARKETING

RESUMO

Considerando o crescimento do número de estudos científicos publicados na área de marketing e o consequente desenvolvimento de teorias próprias (Hunt, 2010), o uso de experimentos parece ser cada vez mais pertinente para elucidar o funcionamento dos fenômenos mercadológicos. Este artigo pretende discutir os principais elementos para a realização de um estudo experimental, além de estimular os pesquisadores a adotarem este método de pesquisa. Vários periódicos internacionais (p. ex. JCR, JCP, JMR, JR, JBR) têm publicado artigos com base em estudos experimentais que evidenciam não somente a relação entre dois eventos, mas também como tais eventos ocorrem, utilizando análises de mediação e moderação. Este artigo pretende servir como um guia na condução de experimentos para pesquisadores iniciantes e oferecer novas perspectivas da pesquisa experimental para pesquisadores mais experientes.

Palavras-chave: Método Experimental; Pesquisa em Marketing; Pesquisa em Comportamento do Consumidor; Relações Causais.

EXPERIMENTAL RESEARCH IN MARKETING

ABSTRACT

Considering the growing number of scientific studies published in the marketing field and the development of unique theories of the area (Hunt, 2010), using experimental designs seems increasingly appropriate to investigate marketing phenomena. This article aims to discuss the main elements in conducting experimental studies and also to stimulate researchers to adopt this research method. Several international journals (e.g., JCR, JCP, JMR, JR, JBR) have been publishing articles based on experiments that not only demonstrate a relationship between two events, but also elucidate how they occur by means of mediation and moderation analyses. This article intends to be a roadmap for novice researchers on how to conduct experiments and to offer new perspectives in experimental research for experienced researchers.

Key words: Experimental Design; Marketing Research; Consumer Behavior Research; Causal Relationships.

José Mauro da Costa Hernandez¹
Kenny Basso²
Marcelo Moll Brandão³

¹ Doutor em Administração de Empresas pela Fundação Getúlio Vargas – FGV. Professor da Universidade de São Paulo – USP e do Centro Universitário da FEI, Brasil. E-mail. jmhernandez@uol.com.br

² Doutor em Administração pela Universidade Federal do Rio Grande do Sul – UFRGS. Professor da Faculdade Meridional – IMED, Passo Fundo – RS, Brasil. E-mail. bassokenny@gmail.com

³ Doutor em Administração pela Fundação Getúlio Vargas – FGV. Professor de Marketing do PPGA - Programa de Pós-Graduação em Administração, Universidade Nove de Julho – UNINOVE, São Paulo, Brasil. E-mail: marcelo.brandão@uninove.br

1 INTRODUÇÃO

Considerando o notável crescimento do número de estudos científicos publicados na área de marketing e o conseqüente desenvolvimento de teorias próprias (Hunt, 2010), o uso de experimentos parece ser cada vez mais pertinente para elucidar o funcionamento dos fenômenos mercadológicos. Quando o campo teórico de um determinado assunto encontra-se bem desenvolvido, métodos de natureza qualitativa, de uma forma geral, podem ser ineficientes para a geração de conhecimento, uma vez que não conseguirão fornecer subsídios para a falsificação das proposições e hipóteses existentes (Bonoma, 1985).

Como a disciplina de marketing já tem teorias bem desenvolvidas, o ponto focal de pesquisa passa a ser a identificação e a mensuração de relações causais por meio de métodos como o experimento. Ao contrário de outros métodos de pesquisa, os experimentos são caracterizados por dois principais fatores: (1) a manipulação de uma ou mais variáveis independentes; (2) o controle sobre as variáveis externas por meio de diversas estratégias como a atribuição aleatória dos sujeitos às condições experimentais. A partir da manipulação de uma ou mais variáveis independentes, o pesquisador, respeitando os controles necessários, tem condições de inferir sobre as conseqüências observadas em uma ou mais variáveis dependentes. Por meio desse processo é possível compreender a relação de causa e efeito existente nos fenômenos mercadológicos.

Quando se considera o volume de publicações em marketing com base em estudos experimentais, fica explícita a relevância deste tipo de pesquisa. A preferência dos periódicos por estudos experimentais, mas principalmente aqueles focados em comportamento do consumidor, pode ser explicada por algumas características da evolução do conhecimento na área. Vários periódicos (p. ex. JCR, JCP, JMR) têm publicado artigos com base em estudos experimentais que evidenciam não somente a relação entre dois eventos (condição necessária para a inferência de causa e efeito), mas também como tais eventos ocorrem, utilizando análises complexas de mediação e moderação. Por meio dessas análises, os fenômenos em marketing ganham uma explicação mais aprofundada sobre os mecanismos pelos quais eles ocorrem e geram os efeitos observados.

Mesmo periódicos considerados mais equilibrados entre uma visão acadêmica e gerencial, que não apresentam uma preferência clara por estudos em comportamento do consumidor (p. ex. JM, JR e JBR), vêm publicando cada vez mais estudos que se propõem a testar relações causais com a utilização de experimentos. Sendo assim, o experimento se apresenta como o desenho de pesquisa causal mais adequado para proporcionar o avanço do conhecimento sobre fenômenos já investigados por meio de estudos correlacionais.

Apesar deste cenário internacional bastante favorável aos estudos experimentais, Mazzon e Hernandez (2013) verificaram que estudos experimentais representaram menos de 5% dos artigos publicados no Brasil na área de marketing entre 2000 e 2009. Embora o objetivo central deste artigo seja discutir os principais elementos de um estudo experimental, espera-se também que ele estimule mais pesquisadores a adotarem este método de pesquisa. Para cumprir estes objetivos, inicialmente é apresentado o conceito de causalidade, seguido pela descrição dos elementos de um experimento. Na seqüência, as classificações dos estudos experimentais são apresentadas. Análises de mediação e moderação são discutidas em seguida, assim como os fatores que podem ameaçar uma boa execução e análise dos resultados dos experimentos. As tendências e desafios para os experimentos em marketing são apresentados juntamente com a conclusão do trabalho.

1.1 O Conceito de Causalidade

A pesquisa causal é indicada quando se deseja verificar se existe uma relação de causa e efeito entre duas ou mais variáveis. Entretanto, a noção de causa e efeito que as pessoas costumam usar no dia-a-dia geralmente não é a mesma dos pesquisadores acadêmicos. Por exemplo, quando alguém diz que o aumento do volume de vendas do chocolate X foi provocado pelo aumento do investimento em propaganda, parece uma implicação óbvia que a variação do investimento em propaganda (a causa) tenha sido responsável pela variação no volume de vendas (o efeito). Embora a maioria das pessoas sintasse confortável com essa noção de causa e efeito, o pesquisador acadêmico provavelmente diria que não se pode estar seguro de que o aumento no volume de vendas se deve ao aumento do investimento em propaganda, caso não se possam eliminar outras possíveis explicações. Por exemplo, as vendas do chocolate X poderiam ter aumentado devido à menor intensidade de propaganda dos concorrentes, a um aumento na distribuição do chocolate X ou até mesmo devido a uma mudança na preferência dos consumidores. De fato, não se pode estar seguro nem mesmo de que a pretensa causa (o aumento de investimento publicitário) tenha ocorrido antes do efeito (o aumento do volume de vendas): caso o investimento publicitário tenha sido fixado como uma porcentagem fixa das vendas, é possível que o aumento do volume de vendas tenha sido responsável pelo aumento do investimento publicitário, e não o inverso.

O significado de causalidade já suscitou, e ainda suscita, um ardoroso debate em torno de seu significado para a ciência. No exemplo anterior, o fato de que vendas e investimentos em propaganda variaram ao mesmo tempo não implica causalidade, mas apenas a existência de uma variação concomitante. Logo, o que distingue a afirmação “X causa Y” da afirmação “X

varia concomitantemente com Y”? De acordo com Hunt (2010), quatro condições são necessárias e suficientes para que se possa inferir uma relação de causalidade: sequência temporal, variação concomitante, associação não espúria e suporte teórico. Por conseguinte, as relações entre variáveis que não cumprirem todos os quatro critérios não podem ser consideradas causais.

Por sequência temporal deve-se entender que a variável responsável pela causa, também chamada de variável independente ou explicativa, deve anteceder a variável responsável pelo efeito, também chamada de variável dependente ou critério. Entretanto, em muitas situações não é possível distinguir claramente o que ocorreu antes do que ocorreu depois. Como visto no exemplo do chocolate X, por acreditar que o aumento das vendas sempre decorre do aumento do investimento publicitário, um observador menos atento poderia concluir que, de fato, foi isso que ocorreu. Entretanto, é possível que o contrário tenha ocorrido em determinada situação e, portanto, não se pode concluir definitivamente sobre a relação de causalidade do fenômeno. Embora se tenha presumido por vários anos que a mudança da atitude precede a mudança de comportamento, pesquisas mostram que as mudanças de comportamento geralmente precedem as mudanças de atitude (Fishbein & Ajzen 1972; Ray 1973). Em ambientes complexos é praticamente impossível distinguir a sequência temporal dos fatos. Por exemplo, o aumento da taxa de desemprego é causa ou consequência do aumento da taxa de juros? A desvalorização cambial de uma moeda é causa ou consequência da aceleração da inflação? Provavelmente há argumentos favoráveis para ambos os lados.

A segunda condição para se inferir causalidade - a variação concomitante - sugere que mudanças no nível ou na presença da variável-causa devam estar sistematicamente associadas a mudanças no nível ou na presença da variável-efeito (Hunt, 2010). Quando existe correlação entre duas variáveis (uma medida estatística da associação entre essas variáveis), há evidência em favor da causalidade; a ausência de correlação, por sua vez, em geral é suficiente para excluir a hipótese de causalidade.

Para se assegurar que a relação entre causa e efeito não seja espúria - a terceira condição para se inferir causalidade - é necessário que não haja uma variável qualquer que, ao ser introduzida como variável explicativa, faça desaparecer a associação sistemática entre as variáveis causa e efeito. Por exemplo, suponha-se que um pesquisador esteja analisando a relação entre duas variáveis: X (variável relativa à causa) e Y (variável relativa ao efeito). Para verificar se existe uma associação entre X e Y, o pesquisador executa uma regressão linear ou um método equivalente usando a variável X como explicativa e a variável Y como critério. Suponha-se que inicialmente se tenha verificado que o coeficiente relativo à variável explicativa seja significativo. Suponha-se ainda que o

pesquisador agora insira na regressão outra variável Z, uma explicação alternativa a X para o efeito sobre Y. Pode-se dizer que a relação entre X e Y é espúria, caso o coeficiente associado à variável Z seja significativo e o coeficiente associado à variável X deixe de ser significativo. Se, por outro lado, o coeficiente associado à variável X permanecer significativo, pode-se afirmar que a relação entre X e Y não é espúria e a relação de causalidade não pode ser descartada com base neste argumento.

A necessidade de mostrar que uma relação não é espúria levou muitos filósofos da ciência a questionarem se é possível, de fato, inferir uma relação de causalidade. O argumento é que não é possível testar todas as explicações alternativas possíveis e, portanto, nunca seria possível eliminar completamente o argumento de relação espúria. Portanto, nenhuma hipótese poderia ser refutada, levando mesmo alguns teóricos de marketing a afirmarem que a “ciência é relativa” (Peter & Olson 1983, p. 120-21).

Entretanto, para os defensores do pragmatismo científico (Hunt, 2010), o suporte teórico, a quarta condição para que se possa inferir uma relação de causalidade, resolveria muitos dos problemas levantados pelos defensores do relativismo científico. Por exemplo, existe suficiente teoria em marketing para explicar porque o aumento do investimento em propaganda pode provocar um aumento do volume de vendas de um produto qualquer, e, portanto, poderia ser usada para se inferir a relação de causalidade proposta.

Esses quatro critérios, tomados em conjunto, reforçam o valor científico do desenho de pesquisa experimental, especialmente se comparado com outros métodos de pesquisa não experimental como o levantamento e a pesquisa exploratória. Ao longo do tempo, os conceitos de experimento e pesquisa causal se tornaram praticamente sinônimos, a ponto de muitos se referirem à pesquisa causal como pesquisa experimental, uma vez que o experimento é o único método de pesquisa aceito que permite inferências sobre uma relação de causalidade. A próxima seção analisa o conceito e os elementos de um desenho de pesquisa experimental.

2 ELEMENTOS DE UM EXPERIMENTO

Em um experimento, o pesquisador manipula os níveis das variáveis independentes e observa o resultado produzido sobre a variável dependente, enquanto controla o efeito de outras variáveis que podem oferecer explicações alternativas. Ao se examinar atentamente cada um dos elementos de um experimento (Fig. 1), fica mais fácil entender porque o experimento é o único método que garante as condições necessárias e suficientes para se inferir uma relação de causalidade.

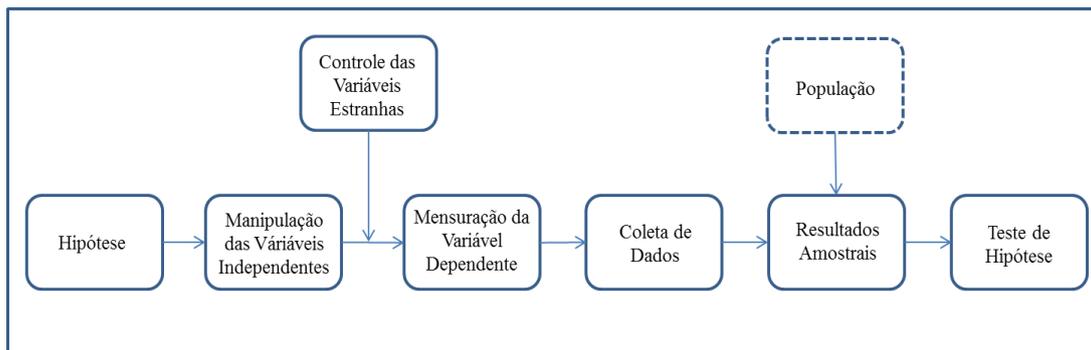


Figura 1 - Elementos de um experimento

Uma das condições para se inferir causalidade é o suporte teórico para a suposta relação em análise. O suporte teórico permite a formulação de hipóteses que serão testadas após a coleta de dados amostrais. Embora para os desenhos de pesquisa exploratório e descritivo a formulação de hipóteses não seja necessariamente um pré-requisito, em estudos experimentais científicos a existência de uma hipótese teórica é fundamental.

O elemento central do experimento, que o distingue dos demais tipos de desenhos de pesquisa, é a manipulação das variáveis independentes. Ao manipular a variável independente, o pesquisador garante a sequência temporal (primeira condição para se inferir causalidade), pois ele controla quando o estímulo correspondente à variável independente será apresentado ao sujeito da pesquisa, bem como em que momento será mensurada a variável dependente. Desta forma, o pesquisador pode garantir que a variável dependente seja sempre mensurada após o estímulo correspondente à variável independente ter sido apresentado.

Manipulando a variável independente, o pesquisador também tem controle preciso sobre a natureza de cada nível do tratamento. Sua expectativa é demonstrar que determinados níveis da variável independente provocam respostas diferentes na variável dependente. Em geral, uma variável independente é manipulada em dois níveis, embora possam ser usados tantos níveis quantos forem necessários, de acordo com a necessidade. O objetivo é garantir que os níveis escolhidos sejam suficientes para produzir efeitos diferentes sobre a variável dependente.

Por exemplo, para verificar se o ceticismo em relação à propaganda modera o efeito da similaridade percebida de uma extensão de marca sobre a atitude em relação à extensão de marca, Hernandez e Marinelli (2013) conduziram um experimento em que era necessário manipular a similaridade percebida da extensão de marca. Uma das principais preocupações era que a similaridade percebida das extensões variasse suficientemente, a ponto de produzirem respostas diferentes, mas não a ponto de serem consideradas

absurdas. Isso poderia diminuir a validade externa do estudo (grau em que um estudo experimental representa a realidade). Por exemplo, no experimento em que o estímulo principal era o fabricante de computadores Dell, foram escolhidas como extensões de marca uma impressora multifuncional (muito similar), um óculos de sol (moderadamente similar) e uma máquina de café (muito dissimilar). A escolha apropriada de diferentes níveis para a similaridade percebida tornou mais fácil verificar o efeito pretendido sobre a atitude em relação às extensões de marca propostas.

Num mesmo desenho experimental, pode haver tanto variáveis independentes manipuladas quanto variáveis independentes mensuradas. As variáveis manipuladas são sempre discretas mas as variáveis mensuradas podem ser discretas ou intervalares. Por exemplo, suponha um experimento em que se deseja verificar o efeito do ceticismo em relação à propaganda sobre a persuasão de uma mensagem publicitária com a presença de uma celebridade. Uma das opções para este experimento é usar uma escala para mensurar o ceticismo e manipular a mensagem publicitária (com e sem celebridade). O maior problema de apenas se mensurar uma variável independente em um experimento é que ela pode estar fortemente correlacionada com outra variável e, portanto, não se pode eliminar a explicação alternativa de que o efeito foi provocado pela outra variável que não foi mensurada. Quando a variável é diretamente manipulada, a explicação alternativa é menos plausível. Por exemplo, para verificar se a necessidade de conclusão cognitiva (do inglês *need for cognitive closure*) influencia o estilo de busca de informações de um indivíduo (busca baseada em atributos vs. busca baseada em alternativas), Choi, Koo, Choic e Auh (2008) conduziram dois experimentos. No primeiro, a necessidade de conclusão cognitiva foi mensurada, enquanto no segundo ela foi manipulada. Esta estratégia (medir e manipular) é usada principalmente quando a variável independente do experimento é uma característica individual como, por exemplo, ceticismo,

autoestima, envolvimento, humor, autoconfiança ou o estilo de processamento de informações.

Assim, é importante lembrar que um desenho experimental requer a presença de pelo menos uma variável independente manipulada. Em determinadas situações, os níveis da variável independente são determinados a posteriori com base em determinadas características da população (por exemplo, níveis de renda mensal). Quando um estudo inclui apenas variáveis mensuradas, o estudo é dito observacional ou não experimental.

Em experimentos de ciências sociais, garantir que a manipulação da variável independente tenha sido realizada com sucesso nem sempre é uma tarefa simples. Como esta é uma condição necessária para que se possa caracterizar um experimento, costuma-se empregar variáveis de conferência de manipulação, cujo único objetivo é assegurar que a manipulação tenha sido feita como o planejado. Por exemplo, em estudos envolvendo diferentes extensões de marcas como estímulo, é necessário pedir aos sujeitos indicarem quanto as extensões apresentadas são similares à marca, e ao final do experimento conferir se a manipulação foi feita como havia sido previsto. Quando uma manipulação falha, é recomendável reiniciar o experimento até que se assegure que a variável independente tenha sido manipulada corretamente.

As variáveis independentes podem ser manipuladas de forma direta, quando observáveis, ou indiretamente, quando não observáveis. Em geral, variáveis do contexto, tais como preço, tipo de promoção, cores utilizadas em um ambiente de loja ou em uma peça de propaganda impressa, número de vendedores em uma loja, tipo de mensagem de uma propaganda ou densidade humana dentro de uma loja, podem ser manipuladas de forma direta. Esse tipo de variável é relativamente fácil de ser testada em relação à confiabilidade de sua manipulação, uma vez que basta perguntar ao sujeito se ele entendeu ou percebeu a manipulação. Por outro lado, variáveis relacionadas ao indivíduo, como necessidade de conclusão cognitiva, foco regulatório, ceticismo, depleção do ego ou autoestima, são manipuladas por aproximação e são mais difíceis de serem verificadas. Nestes casos, em geral usam-se escalas para conferir a manipulação, ou pergunta-se indiretamente sobre o estado do indivíduo. Por exemplo, existem várias tarefas que provocam a depleção do ego, mas não existe uma escala que mensure esse estado. Portanto, o que se costuma fazer para verificar a manipulação da depleção do ego é perguntar ao indivíduo quanta dificuldade ele teve para realizar a tarefa a que foi submetido.

Levando-se em conta o custo e o tempo gastos para desenvolver uma pesquisa experimental, é fortemente recomendável que as manipulações sejam pré-testadas, até apresentarem resultados aceitáveis. Embora o pré-teste não garanta o sucesso da manipulação na aplicação do experimento, ele diminui a chance de fracasso.

Variáveis estranhas são aquelas que podem influenciar o resultado de um experimento, ou aquelas que proporcionam explicações alternativas à variável independente. Desta forma, o controle das variáveis estranhas assegura que a única coisa que muda durante o experimento é o nível do tratamento (Tabachnick & Fidell, 2006). Em experimentos das ciências naturais é relativamente mais fácil controlar o efeito das variáveis estranhas, porque os experimentos podem ser realizados em laboratório - ambientes totalmente isolados. Entretanto, nos experimentos de ciências sociais, o controle das variáveis estranhas é muito mais difícil.

Existem três maneiras de controlar o efeito das variáveis estranhas. A primeira maneira é manter constantes, durante a realização do experimento, todas as variáveis que possam influenciar os resultados. a realização de experimentos em laboratórios permite manter as variáveis ambientais (luz, temperatura e barulho, entre outras) e as condições do teste (por exemplo, aplicação dos estímulos e forma de coleta dos dados) relativamente controladas. Embora o efeito das variáveis individuais (por exemplo, autoestima, autoconfiança, experiência) não possa ser controlado pelo desenho, geralmente eles são “controlados” pela atribuição aleatória dos sujeitos às condições experimentais.

Outra maneira de controlar os efeitos das variáveis estranhas é contrabalançando as condições experimentais. Por exemplo, suponha um experimento em que os sujeitos serão expostos a três propagandas diferentes. Se todos os sujeitos forem expostos às propagandas na mesma ordem, é possível que haja um efeito provocado pela ordem em que eles foram apresentados. Como existem seis formas diferentes de apresentar as três propagandas, cada sexta parte da amostra as verá em uma ordem diferente. Ao final, é necessário verificar se a ordem de apresentação teve algum efeito sobre a variável dependente. Caso negativo, a ordem de apresentação deve ser descartada da análise.

A última maneira de controlar o efeito das variáveis estranhas é por meio do controle estatístico. Na grande maioria dos casos, o controle estatístico é feito incluindo-se as variáveis estranhas no modelo analisado, como se fossem mais uma variável independente, para que seu efeito sobre a variável dependente possa ser isolado. Variáveis estranhas ao modelo também são classificadas como covariáveis em livros e manuais de análise multivariada de dados.

Um grande cuidado que se deve ter ao realizar um experimento é evitar que os sujeitos tenham consciência sobre a natureza dos estímulos e dos efeitos esperados pelo pesquisador, uma vez que esta consciência pode se tornar uma variável estranha. Por exemplo, um dos efeitos mais conhecidos em estudos envolvendo tratamentos médicos é o efeito placebo, ou seja, o efeito provocado pela expectativa e motivação do paciente de melhorar após o tratamento. Para se evitar a influência do efeito placebo sobre os resultados,

parte da amostra recebe um medicamento com princípio ativo e a outra parte recebe um placebo, medicamento sem princípio ativo. Embora os pacientes sejam informados disso, é fundamental garantir que eles não saibam o que estão recebendo. Em geral, o mais indicado é que os experimentos sejam realizados em condição duplo-cego, ou seja, nem o sujeito nem o pesquisador têm consciência sobre as condições experimentais a que cada sujeito foi exposto.

Para evitar que os sujeitos tenham consciência sobre os objetivos do experimento, geralmente se usa uma história fictícia. Por exemplo, em um estudo sobre o efeito placebo provocado pelo preço de um produto, Wright et al. (2013, experimento 1) disseram aos participantes que uma nova bebida seria lançada no Brasil e o objetivo do estudo era testar os benefícios do produto. Supostamente, a nova bebida aumentava o desempenho intelectual e, para testar estes benefícios, os respondentes deveriam fazer uma tarefa de desempenho cognitivo. Entretanto, para aumentar a veracidade da história, foi também dito que a bebida demorava cerca de 10 minutos para fazer efeito e os participantes deveriam assistir um documentário enquanto esperavam pelos efeitos da bebida. De fato, a bebida era uma mistura de guaraná, suco de laranja e refrigerante de limão e não tinha nenhum efeito real.

Do ponto de vista ético, é importante que ao final do experimento seja esclarecido para o participante o verdadeiro objetivo do estudo. Esta prática é universalmente aceita e constitui requisito essencial para que um projeto experimental seja aprovado pela Comissão de Ética das organizações dedicadas à pesquisa.

A mensuração da variável dependente é o elemento mais próximo de outros tipos de desenho de pesquisa. Em estudos de marketing, uma variável-critério pode ter várias naturezas (atitudes, emoções, comportamentos, julgamentos, escolhas, entre outros). Praticamente não existem limites à criatividade para a mensuração da variável dependente. Por exemplo, Waber, Shiv, Carmon e Ariely (2008) queriam verificar se o preço de um analgésico tem influência sobre seu efeito terapêutico. Para tanto, recrutaram voluntários para o teste de um medicamento e os dividiram em dois grupos. Foi dito a todos os participantes que o estudo tinha como objetivo verificar o efeito terapêutico de um analgésico. A um dos grupos foi dito que cada pílula do medicamento custava 10 centavos de dólar, enquanto a outro grupo foi dito que o preço da pílula era dois dólares e cinquenta centavos. Cada participante do estudo passou por duas sessões de choques elétricos no pulso, uma antes e outra depois de ter tomado o medicamento; o nível de dor percebida foi mensurado nas duas sessões. Os resultados mostraram que a redução da dor percebida nas duas sessões, a variável dependente do experimento, foi maior para o grupo que tomou o medicamento com preço maior; além disso, a redução foi mais acentuada quanto mais intenso foi o nível do choque.

Uma vez que o experimento tenha sido planejado e os instrumentos de coleta de dados estejam prontos, o próximo passo é a coleta de dados. Independentemente de os dados serem coletados em laboratórios ou em ambientes naturais, nesta fase é importante garantir que os casos sejam atribuídos às condições experimentais de forma aleatória, ou seja, todos os indivíduos devem possuir a mesma probabilidade de serem selecionados para uma ou outra condição experimental. A atribuição aleatória dos sujeitos às condições do experimento tem dois objetivos (Tabatchnik & Fidell, 2006). O primeiro é evitar que o pesquisador atribua (ainda que inconscientemente) os sujeitos melhores, maiores, mais inteligentes ou mais aptos à condição experimental que ele acredita ser a mais efetiva. O segundo objetivo é eliminar as eventuais diferenças individuais (por exemplo, gênero, idade, experiência, motivação e autoestima) entre os sujeitos. Ao tornar aleatória a atribuição dos sujeitos às condições experimentais, espera-se que as diferenças individuais estejam distribuídas de tal forma que cada condição experimental possa ser considerada equivalente, antes que os sujeitos sejam expostos ao estímulo. Para se certificar da distribuição aleatória, testes de diferenças demográficas ou de outras características podem ser feitos entre os grupos experimentais. Se o pesquisador não encontrar uma associação entre tais características e os grupos, pode-se considerar que a amostra foi aleatoriamente distribuída.

Uma vez que os dados sejam coletados em uma amostra previamente selecionada, o último passo do experimento é testar a hipótese. Uma característica que distingue o experimento dos outros métodos de pesquisa é que seu objetivo sempre é o teste de uma hipótese que permita inferir algo sobre a população representada pela amostra.

Concluindo, se existe teoria suficiente para esperar uma associação entre uma variável independente e uma dependente; se os sujeitos forem alocados às condições experimentais aleatoriamente, de tal forma que não exista associação entre as características individuais e a variável independente; se o pesquisador tem controle sobre todas as variáveis ambientais do experimento; se o pesquisador tem controle do momento em que o estímulo é apresentado e do momento em que a variável dependente é mensurada; então diferenças na variável dependente são naturalmente atribuídas às diferenças nos níveis dos tratamentos, já que não existe nenhuma outra explicação para a variação sistemática na variável dependente. Por isso diz-se que o experimento é a única forma de se demonstrar uma relação de causa-e-efeito.

3 CLASSIFICAÇÕES DE EXPERIMENTOS

Os experimentos podem ser diferenciados quanto ao tipo, sendo classificados em estudos experimentais, quase-experimentais ou pré-experimentais; ao ambiente de realização, podendo serem realizados em laboratório (ambiente controlado) ou no campo (ambiente real); e quanto aos desenhos experimentais utilizados, ou seja, se as manipulações são realizadas entre-sujeitos, dentro-sujeitos ou em ambas as formas. Na sequência, essas classificações serão discutidas e ao final, será apresentado o desenho experimental fatorial amplamente utilizado nos estudos em marketing.

3.1 Tipo: Experimentais vs. quase-experimentais vs. pré-experimentais

Os estudos experimentais pressupõem o uso da atribuição aleatória dos sujeitos aos tratamentos. Estes estudos reduzem a possibilidade de existirem explicações alternativas para os resultados encontrados, uma vez que vieses de atribuição dos sujeitos às condições experimentais são eliminados pela aleatoriedade, algo que não ocorre em estudos pré- ou quase-experimentais.

Diferentemente dos estudos experimentais, os quase-experimentais não envolvem a distribuição aleatória dos sujeitos aos tratamentos. Eles geralmente são utilizados quando o pesquisador não possui controle sobre o processo de aleatoriedade, ou quando os estudos são realizados no campo. Além disso, os estudos quase-experimentais pressupõem que o pesquisador não possua controle pleno sobre a exposição dos sujeitos às manipulações da variável independente. Por exemplo, Chae, Li e Zhu (2013, estudo 3), utilizaram um estudo quase-experimental para verificar os efeitos da localização da imagem de um produto em um anúncio (perto ou longe dos efeitos retratados) sobre o julgamento de efetividade deste produto. No procedimento utilizado, como variáveis independentes, os autores manipularam um fator (proximidade espacial) e mediram outro (nível de conhecimento). Essa mensuração permitiu aos pesquisadores classificar os sujeitos em dois grupos (alto e baixo conhecimento). Essa forma de criar o segundo fator não permite a distribuição aleatória dos sujeitos, uma vez que a medida já estabelece em qual grupo o sujeito irá ser enquadrado, o que leva o estudo a ser classificado como quase-experimental.

Os estudos pré-experimentais, por sua vez, caracterizam-se por não possuírem grupo de controle. Em outras palavras, nos estudos pré-experimentais apenas um grupo é estudado, sem que sejam feitas comparações com um grupo equivalente que não tenha recebido a manipulação (grupo de controle). Suponha-se que um pesquisador deseja analisar a injustiça percebida pelos clientes após uma alteração de preço. Para tanto, o pesquisador promove uma alteração de preço (X) e, em seguida, mensura a injustiça percebida

pelos sujeitos (Y) obtendo com isso o nível de injustiça, mas sem ter uma base para comparação com outro grupo (o grupo de controle), que tenha sido exposto a uma condição de preço normal.

Em outro exemplo, Estrela e Botelho (2006) utilizaram um pré-experimento para verificar o efeito do desconto de preços nas vendas do varejo. Para tanto, eles reduziram o preço dos produtos em uma semana, e na semana seguinte retornaram os preços ao patamar anterior. Nesse experimento não houve controle sobre quando os sujeitos foram expostos ao tratamento, nem se sabe se os sujeitos perceberam a variação de preços. Também não houve um grupo de controle (não exposto à variação de preços) que pudesse ser comparado com o grupo que foi exposto à variação. Estas características definem o estudo pré-experimental.

3.2 Ambiente de realização: Campo vs. Laboratório

Os estudos experimentais podem ser realizados em um ambiente real (experimento de campo) ou em um ambiente artificial (experimento de laboratório). Para decidir em que ambiente fazer o experimento, Bonoma (1985) sugere que o pesquisador deva compreender se o fenômeno pode ser estudado de forma satisfatória fora do seu ambiente natural uma vez que, retirado do local de ocorrência natural, o fenômeno pode apresentar distorções. Já Calder, Phillips e Tybout (1981) argumentam que, quando o objetivo do estudo é a aplicação e teste de proposições teóricas, o experimento laboratorial é mais indicado, uma vez que pode prover um maior controle sobre as variáveis, além de possibilitar a atribuição aleatória dos participantes aos grupos experimentais.

Em um experimento para investigar o efeito do tamanho e da distorção do formato do produto no comportamento de reciclagem, Trudel e Argo (2013) fizeram uso de estudos experimentais em laboratório. No entanto, convém mencionar que, embora o local de aplicação utilizado por Trudel e Argo (2013) tenha sido o laboratório, os procedimentos experimentais utilizados fizeram com que os sujeitos se envolvessem e simulassem o fenômeno o mais próximo possível da realidade. Para tanto, os autores entregaram aos participantes diversas folhas de papel e uma tesoura sob o pretexto de que o estudo envolvia o teste da tesoura. Durante o estudo, os participantes foram envolvidos em tarefas de cortar o papel em diferentes tamanhos, forma encontrada para manipular a forma e o tamanho do produto. Ao final, os participantes deveriam avaliar o desempenho da tesoura por meio de uma escala. Ao final de uma série de outras tarefas não relacionadas com o objetivo principal do estudo, os participantes deixaram o laboratório e se depararam com duas lixeiras, uma sem identificação e outra com a identificação de reciclagem, onde poderiam descartar os papéis. O montante descartado em cada lixeira serviu como medida comportamental de reciclagem.

Para Sawyer, Worthing e Sendak (1979), os experimentos laboratoriais têm diversas vantagens, como a maior flexibilidade na verificação de medidas não comportamentais como atitudes e crenças, maior capacidade de especificar e determinar os limites de atuação do fenômeno, menor custo, menor tempo gasto na pesquisa e maior sigilo em relação ao experimento e aos dados. Por outro lado, os estudos experimentais de campo proporcionam um maior realismo, uma vez que envolvem o ambiente real de ocorrência do fenômeno.

Um exemplo de estudo realizado no campo foi desenvolvido por Baca-Motes, Brown, Gneezy, Keenan e Nelson (2013) para avaliar se o comprometimento de um hotel com ações de responsabilidade ambiental gera uma mudança comportamental nos hóspedes. Para tanto, os autores treinaram os recepcionistas de um hotel para apresentar aos hóspedes, no momento do *check-in*, uma carta de compromisso do hotel com o meio ambiente. Por meio da carta os autores manipularam a variável independente (comprometimento ambiental) em dois níveis: utilizando termos gerais (ser ambientalmente responsável) ou utilizando termos específicos (reutilizar as toalhas durante a estada). Eles também obtiveram um grupo de controle, que não foi exposto à manipulação do comprometimento ambiental nem à manipulação dos sinais do compromisso. O estudo durou 31 dias e 4.325 pessoas foram expostas às diferentes condições experimentais.

3.3 Desenhos entre-sujeitos, dentre-sujeitos e mistos

Quando envolvido em um experimento com mais de um tratamento experimental, o pesquisador pode defrontar-se com três opções: 1) expor cada sujeito a apenas um tratamento experimental e comparar as medidas entre os sujeitos expostos a diferentes tratamentos (desenho entre-sujeitos); 2) expor cada sujeito a todos os tratamentos experimentais e comparar as medidas dentre os sujeitos (desenho dentre-sujeitos); 3) misturar as duas opções anteriores, expondo os sujeitos a diferentes tratamentos de um ou mais fatores e a todos os tratamentos de outro(s) fator(es).

Os desenhos experimentais entre-sujeitos, mais frequentemente utilizados na literatura de marketing, são mais recomendados quando é provável ocorrer o efeito de demanda. Esse desenho experimental é utilizado quando o contexto experimental de um tratamento pode influenciar outro tratamento, quando o contexto experimental envolve uma única decisão ou quando há diversas variáveis a serem manipuladas e/ou controladas (Greenwald, 1976; Charness, Gneezy & Kuhn, 2012). Por exemplo, para verificar a influência da forma de apresentação do alimento e do tipo de alimento sobre a percepção de sabor das pessoas, Poor, Duhachek e Krishnan (2013) utilizaram um desenho experimental entre-sujeitos. Para tanto, os sujeitos foram expostos a imagens de

alimentos saudáveis (amêndoas) ou não saudáveis (batatas fritas) e a situações em que esses alimentos eram apresentados sozinhos ou em uma situação de consumo. Nesse estudo, cada sujeito foi exposto somente a uma das quatro possíveis situações.

Por sua vez, os desenhos experimentais dentre sujeitos, embora menos utilizados que os desenhos entre-sujeitos, são geralmente empregados: em estudos em que se deseja eliminar o viés de resposta dos sujeitos, uma vez que tratamentos diferentes não são confundidos com sujeitos diferentes; quando há poucos sujeitos para compor a amostra do estudo; quando o contexto experimental envolver uma série de decisões; quando o pesquisador busca aumentar a validade externa dos resultados; e quando se busca a validade interna, independente dos processos de distribuição aleatória dos sujeitos aos tratamentos (Greenwald, 1976; Charness et al., 2012). Por exemplo, Ge, Häubl e Elrod (2012), testaram se a escolha de uma determinada alternativa é maior quando uma informação favorável ao produto é entregue com atraso (vs sem atraso) durante o processo de triagem pré-compra. Para tanto, os autores manipularam o atraso no fornecimento da informação, expondo os sujeitos tanto à condição em que a informação, no processo pré-compra, era fornecida com atraso (decisão A) quanto à condição em que a informação era fornecida junto com as demais informações (decisão B), portanto, dentre sujeitos.

Embora o desenho dentre sujeitos possua uma série de possibilidades de aplicações, os efeitos psicológicos gerados pelos sucessivos tratamentos podem impor vieses aos resultados. Dentre esses efeitos psicológicos, uma grande preocupação recai sobre o efeito de demanda (Charness et al., 2012). Neste efeito, a inclusão de um tratamento influencia como os sujeitos irão se comportar e responder aos tratamentos subsequentes. De acordo com Charness et al. (2012) e Greenwald (1976), os sujeitos tendem a responder aos estímulos subsequentes baseados em como perceberam os estímulos antecedentes, ou seja, os participantes manipulam suas respostas de acordo com as hipóteses percebidas no experimento. A menos que o pesquisador queira estudar os efeitos da prática e do aprendizado, este efeito pode ser prejudicial para os resultados do estudo.

Para alinhar os benefícios de cada tipo de desenho experimental, alguns pesquisadores têm aplicado os dois tipos, seja no mesmo estudo (p. ex. McQuarrie & Mick, 2003; Thomas, Dsai & Seenivsan, 2011; Thomas & Tsai, 2012) ou em estudos diferentes (p. ex. Chan & Sengupta, 2010).

3.4 Desenhos fatoriais

Os desenhos fatoriais são utilizados em estudos que visam testar o efeito de duas ou mais variáveis independentes sobre as variáveis dependentes. Um estudo fatorial completo envolve a combinação de todas as condições dos fatores. Assim, um estudo envolvendo dois fatores, um com 3

condições e outro com 2 condições, resulta em 6 combinações possíveis, ou seja, seis grupos experimentais. Os estudos fatoriais permitem ao pesquisador explorar todas as combinações possíveis entre todos os níveis dos fatores envolvidos, permitindo a análise tanto dos efeitos principais de cada um dos fatores quanto das possíveis interações entre eles. Recomenda-se o uso desse tipo de experimento quando a análise da interação é necessária para o teste das hipóteses propostas.

Tomemos a hipótese de que consumidores com poder elevado (uma medida de influência social) percebem o preço como mais injusto quando eles pagam mais do que outros consumidores (comparação social), enquanto consumidores com baixo poder percebem maior injustiça no preço quando eles pagam mais do que pagaram anteriormente (auto-comparação). Jin, He e Zhang (2014) realizaram um estudo fatorial 2 (poder: baixo vs. alto) X 2 (referência: auto-referência vs. referência nos outros) para testar tal hipótese e a combinação dos níveis dos dois fatores gerou quatro condições experimentais. Cada participante foi exposto a apenas uma das condições experimentais em um desenho entre-sujeitos. Os autores identificaram que há um efeito significativo da interação entre os fatores na percepção de injustiça de preço, permitindo-lhes comprovar a hipótese proposta pelo estudo.

Uma vez que a complexidade de um estudo experimental passa pela quantidade de fatores e pela quantidade de níveis em cada fator, o pesquisador deve estar ciente de que seu estudo pode tornar-se inviável em termos de aplicação, de custo e de complexidade da execução. Portanto, sugere-se que os pesquisadores comecem com estudos experimentais mais simples, procurando analisar o efeito principal que se pretende investigar; posteriormente, podem acrescentar outras variáveis que auxiliem a explicar o fenômeno.

Neste sentido, Chan e Sengupta (2010), para verificar o efeito da lisonja nas atitudes explícitas e implícitas, realizaram dois estudos (1A e 1B) menos complexos em termos de manipulação de fatores. Na sequência, os autores realizaram outros estudos fatoriais em que há uma interação de fatores, aumentando a explicação do fenômeno. Essa estratégia permite ao pesquisador compreender e adquirir um conhecimento empírico sobre a ocorrência do fenômeno, para posteriormente aprofundar este conhecimento.

4 MEDIAÇÃO E MODERAÇÃO

Além de saber se existe uma relação de causa e efeito entre X e Y, muitas vezes o pesquisador também quer saber como e quando X exerce influência sobre Y. A primeira pergunta (como?) refere-se aos processos psicológicos, cognitivos e biológicos que relacionam X a Y e o tipo de análise que procura responder a essa pergunta chama-se mediação. Já a

segunda pergunta (quando?) relaciona-se às condições de fronteira do fenômeno, isto é, em que condições ou para que tipos de pessoas X exerce (ou não) influência sobre Y (Hayes, 2013). A análise que responde a essa pergunta é chamada de moderação. A seguir discutimos cada um destes dois tipos de análise.

4.1 Mediação

Quando o pesquisador deseja estabelecer o mecanismo de um fenômeno, isto é, por qual processo uma variável independente X exerce seu efeito sobre uma variável dependente Y, em geral ele propõe um modelo em que uma ou mais variáveis intervenientes M_1, M_2, \dots, M_k mantêm uma relação de causa e efeito com X e Y (ver a Figura 2). Na Figura 2, a, b e c representam o efeito correspondente a cada um dos caminhos. Esse efeito pode ser mensurado, por exemplo, por meio de regressões lineares.

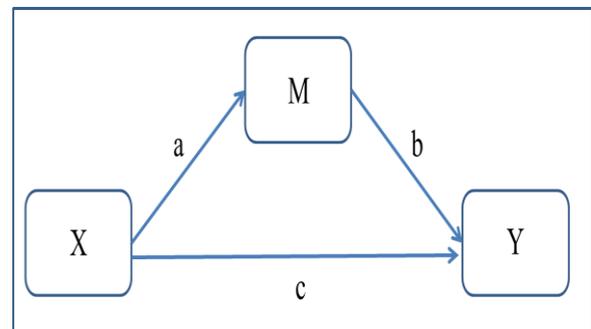


Figura 2 - Esquema conceitual de mediação

Lee, Keller e Sternthal (2010) verificaram que consumidores com foco regulatório em estado de prevenção formam atitudes mais positivas em relação a uma marca quando a propaganda da marca tem uma linguagem mais concreta (baixo nível de representação), enquanto consumidores com foco regulatório em estado de promoção formam atitudes mais positivas quando a mensagem da propaganda tem uma linguagem mais abstrata (alto nível de representação). Para explicar o mecanismo de formação de atitude em relação à marca, Lee, Keller e Sternthal (2010, experimento 4) testaram o efeito de duas variáveis mediadoras: fluência de processamento (grau de dificuldade de compreensão de uma mensagem) e engajamento (motivação para compreender a mensagem). Os resultados sugerem que, quando há adequação entre o foco regulatório e a linguagem da mensagem (foco regulatório prevenção com baixo nível de representação ou foco regulatório promoção com alto nível de representação), tanto a fluência de processamento quanto o engajamento aumentam, levando à formação de atitudes mais favoráveis em relação à marca; quando não há adequação entre o foco regulatório e a linguagem (foco regulatório promoção com baixo nível de representação ou foco regulatório prevenção com alto nível de representação), a fluência de processamento e o

engajamento diminuam, levando à formação de atitudes em relação à marca menos favoráveis.

Baron e Kenny (1986), proponentes do método mais usado pelos pesquisadores de marketing para testar o efeito de mediação, sugerem que “uma variável mediadora explica como eventos físicos externos assumem significado psicológico interno” (p. 1176). De acordo com o método proposto por Baron e Kenny, uma variável M serve como mediadora entre X e Y quando: a) variações nos níveis da variável independente X são responsáveis por variações na variável mediadora M (caminho a , Figura 2); b) variações na variável mediadora M são responsáveis por variações na variável dependente Y (caminho b , Figura 2); c) quando se controla o efeito de M , a relação previamente significativa entre X e Y (caminho c , Figura 2) deixa de ser significativa. A evidência mais forte de mediação surge quando o caminho c (Figura 2) passa a ser zero ou não significativo, podendo-se inferir que a mediação é total; quando o caminho c continua significativo, infere-se que a mediação é parcial.

Entretanto, há cerca de 10 anos esse procedimento para testar mediação vem sendo substituído por um método alternativo (ver, por exemplo, Hayes, 2009; Zhao, Lynch & Chen, 2010). O principal problema do método sugerido por Baron e Kenny (1986) é que ele pressupõe que necessariamente haja uma relação causal entre X e Y , ou seja, que o caminho c da Figura 2 seja significativo. Em outras palavras, se não há relação causal entre X e Y , não há porque se falar em mediador entre X e Y . Entretanto, segundo Bollen (1989, p. 52), “a falta de correlação (entre duas variáveis) não invalida a causalidade” e “correlação não é uma condição necessária nem suficiente para se provar causalidade”. Embora fosse totalmente contrária ao que se acreditava até então (ver a segunda condição de variação concomitante para se demonstrar causalidade, discutida na primeira seção deste artigo), esta nova visão foi rapidamente incorporada. O principal argumento é que, se for possível demonstrar que existe uma ou mais variáveis M_1, M_2, \dots, M_k que medeiam a relação entre X e Y , pode-se concluir que há uma relação causal entre X e Y (ver também a terceira condição de correlação espúria para se demonstrar causalidade, discutida na primeira seção deste artigo).

Assim, para que a variável M seja um dos mediadores da relação entre X e Y é necessário apenas que o caminho $a*b$ da figura 2, correspondente ao efeito indireto de X sobre Y , seja diferente de zero. O teste de Sobel (Sobel, 1982) permite o cálculo do intervalo de confiança para o produto $a*b$. Este teste é usado para realizar inferências sobre o seu valor populacional e vem sendo usado por pesquisadores há bastante tempo, não raro complementando o método proposto por Baron e Kenny. A limitação do teste de Sobel é pressupor que a distribuição de $a*b$ seja normal, quando já se demonstrou que esta distribuição é irregular para pequenas amostras – em geral, as que são usadas em experimentos (Bollen & Stine, 1990). Uma vez que não

é possível saber com certeza se a distribuição de $a*b$ se aproxima de uma distribuição normal em uma dada situação, atualmente usa-se o método de *bootstrap* para calcular o intervalo de confiança do valor de $a*b$.

O *bootstrap* é especialmente indicado para fazer inferências sobre um valor cuja distribuição é desconhecida. Neste método, a amostra original de tamanho n é tratada como uma pequena representação da população original. As observações da amostra são re-amostradas, com reposição, um grande número de vezes (geralmente criando-se entre cinco e dez mil re-amostras) e a estatística de interesse (no presente caso, o valor de $a*b$) é calculada para cada uma das re-amostras. Desta forma, pode-se construir uma distribuição hipotética da estatística de interesse e o intervalo de confiança desejado pode ser obtido com base nessa distribuição hipotética. Estritamente falando, este procedimento não tem o mesmo significado de um teste de hipótese. Na prática, entretanto, a interpretação do intervalo de confiança construído por meio do método *bootstrap* é o mesmo de um teste de hipótese: se ele contém o zero, então não se pode rejeitar a hipótese nula de que a estatística seja igual a zero; caso o intervalo não contenha o zero, rejeita-se a hipótese nula e conclui-se que a estatística é diferente de zero.

Ferraro, Kirmani e Matherly (2013) usaram o método descrito acima para testar a hipótese de que a atitude em relação ao usuário de uma marca medeia a relação entre o uso conspícuo (ostensivo) da marca e a atitude em relação a ela. A hipótese é de que, se a marca é usada de forma conspícua por uma pessoa, outras pessoas podem desenvolver uma atitude negativa em relação a essa pessoa e, conseqüentemente, em relação à marca. Os resultados do estudo mostraram que o uso conspícuo da marca tem impacto negativo sobre a atitude em relação ao usuário e desta sobre a atitude em relação à marca apenas para pessoas que têm baixa conexão com a marca. Para as pessoas com alta conexão com a marca, o uso conspícuo da marca não exerce efeito nem direto nem indireto sobre a atitude em relação à marca. Desta forma, Ferraro e colegas puderam concluir que o uso conspícuo de uma marca pode ser prejudicial para a marca apenas para as pessoas que mantêm baixa conexão com a marca e que o mecanismo desse efeito é mediado pela atitude em relação ao usuário da marca.

Um ponto que deve ser ressaltado no estudo de Ferraro e colegas (2013) é que os resultados explicaram não apenas o “como”, mas também o “quando”, ou seja, o efeito do uso conspícuo sobre a atitude em relação à marca é mediado pela atitude em relação ao usuário, apenas quando a conexão com a marca é baixa - mas não quando é alta. Quando se examina o “quando” de um fenômeno, a análise envolvida é a moderação, tópico discutido a seguir.

4.2 Moderação

A análise de moderação é usada quando se deseja descobrir as condições de fronteira de um fenômeno. Uma associação é dita moderada quando a natureza (grandeza e sinal) da relação entre duas variáveis X e Y depende de uma terceira variável W. A representação usual de moderação é descrita na Figura 3.

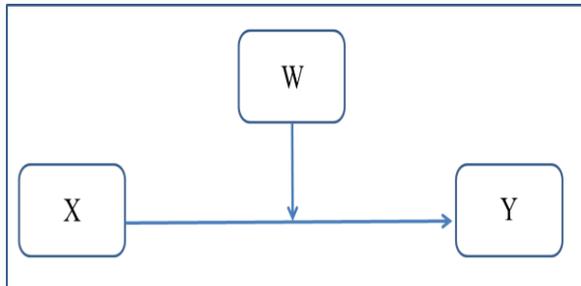


Figura 3 - Esquema conceitual de moderação

O efeito de moderação é também conhecido por interação e pode ser verificada por meio de uma análise de regressão, quando X ou W forem intervalares, ou por meio de uma ANOVA, quando X e W foram nominais ou ordinais. Uma vez que o pesquisador tenha verificado a existência de um efeito de interação, o próximo passo é testar como a moderação ocorre verificando-se qual o efeito de X sobre Y para cada valor de W (este processo é conhecido em inglês por *probing*).

Hernandez, Han e Kardes (2014) testaram a hipótese de que o efeito da familiaridade com uma marca sobre a avaliação de um produto descrito por meio de poucos atributos é moderado pelo conhecimento objetivo do consumidor. O argumento é que sujeitos com baixo conhecimento objetivo fazem uma avaliação heurística do produto, levando em conta principalmente a marca e desconsiderando os demais atributos. A consequência disso é uma avaliação mais favorável do produto com uma marca familiar do que com uma marca desconhecida. Por outro lado, sujeitos com alto conhecimento objetivo fazem uma avaliação sistemática, levando em consideração tanto a marca quanto os demais atributos. Portanto, se os atributos forem insuficientes para uma avaliação adequada, um sujeito com alto conhecimento objetivo fará a mesma avaliação tanto do produto com marca familiar quanto daquele com marca não familiar. Para testar esta hipótese, Hernandez e colegas pediram aos sujeitos que avaliassem uma câmera digital cuja descrição continha quatro atributos. Metade dos sujeitos viu uma descrição com a marca Sony e a outra metade com uma marca desconhecida. Em seguida, os respondentes tiveram que responder dez perguntas para medir seu conhecimento objetivo sobre câmeras digitais. Os resultados mostraram que os sujeitos com baixo conhecimento sobre câmeras digitais avaliaram a câmera digital Sony mais favoravelmente do que a câmera digital desconhecida, enquanto os sujeitos com

alto conhecimento deram avaliações similares às duas câmeras. Uma análise posterior mostrou que o efeito da familiaridade sobre a avaliação era mediada pela suficiência das informações contidas nas descrições.

Quando as variáveis independentes do modelo (X e W) são discretas, a análise de moderação é relativamente simples. Por meio de uma análise de variância tendo Y como variável dependente, verifica-se se a interação entre X e W é significativa. Caso seja significativa, o próximo passo é testar as diferenças de Y para os diferentes valores de X, mantendo fixos cada um dos valores de W.

Entretanto, a solução não é tão fácil quando as variáveis independente ou moderadora são contínuas. Neste caso, o problema não pode ser resolvido por meio de uma análise de variância método com o qual a maioria dos pesquisadores está familiarizada. Até não muito tempo atrás, costumava-se dicotomizar a variável contínua pela média ou pela mediana. O maior problema deste procedimento é a perda de informação decorrente da dicotomização, ou seja, a transformação de uma variável mensurada com cinco, sete ou nove pontos, para apenas dois níveis diminui a força da análise. Além do mais, já se verificou que os resultados da dicotomização podem não ser os mesmos daqueles analisados com os dados completos (para uma discussão mais completa, ver Irwin & McClelland, 2001, 2003; MacCallum, Zhang, Preacher & Rucker, 2002; Maxwell & Delaney, 1993; Fitzsimons, 2008).

Por outro lado, a análise sem a dicotomização, embora não seja exatamente trivial, está longe de ser complicada. Suponha um exemplo simples, no qual a variável independente X de um experimento tenha sido manipulada em apenas dois níveis, e no qual tanto o moderador W quanto a variável dependente Y tenham sido mensurados em escalas que possam ser consideradas intervalares (por exemplo, uma escala do tipo Likert ou diferencial semântico). O primeiro passo é fazer uma análise de regressão tomando a variável Y como dependente e X, W e $X*W$ (a interação entre X e W), como variáveis independentes. Caso o coeficiente associado à interação seja significativo, está demonstrado que existe interação entre X e W ou, em outras palavras, W modera o efeito de X sobre Y.

Uma vez demonstrado que existe o efeito de moderação, o próximo passo é investigar a natureza dessa moderação (*probing*). Isso se obtém executando duas novas regressões similares à anterior, mas transformando os valores da variável moderadora em valores convenientes, geralmente um desvio-padrão abaixo e acima da média da variável moderadora. Mudanças na significância e na grandeza do coeficiente associado à variável independente X indicariam a natureza da moderação. Este tipo de análise tem sido chamada na literatura de marketing de análise *spotlight* e mais detalhes sobre como pode ser feita podem ser encontrados em Aiken e West (1991), Irwin e McClelland (2001) ou em Spiller, Fitzsimons, Lynch Jr. e McClelland (2013). Hayes (2013) desenvolveu PROCESS, uma sintaxe para SPSS ou SAS que pode

ser usada tanto para análises de mediação quanto para análises de moderação (ver <http://www.afhayes.com/>).

5 AMEAÇAS AOS EXPERIMENTOS

Os experimentos possuem diversos pontos de críticas e de discussão entre os pesquisadores. Além disso, diversos fatores podem reduzir o impacto dos estudos experimentais para a ciência e, conseqüentemente, sua capacidade de contribuição teórica. Dentre esses fatores incluem-se a validade dos experimentos, o poder da análise e a seleção da amostra. Pesquisadores preocupados com a aceitação e a disseminação dos resultados de seus estudos experimentais devem atentar para tais fatores, tanto na elaboração quanto na execução de seus estudos.

5.1 Validade em experimentos

Validade é um termo utilizado para referir-se à aproximação entre uma inferência e a verdade. A validade é uma propriedade das inferências e o uso de diferentes métodos exerce distintos impactos sobre a validade de uma inferência (Shadish, Cook & Campbell, 2001).

Quatro tipos de validade podem ser identificados (Calder, Phillips & Tybout, 1982; Shadish et al., 2001): validade estatística, relativa à validade da relação entre os tratamentos e os resultados; validade de construto, ou a capacidade dos itens utilizados para mensuração ou da manipulação realizada representarem o conceito teórico relacionado; validade interna, que representa o quanto as inferências sobre o relacionamento causal são seguras; e a validade externa, relacionada com a capacidade de extensão ou de generalização do relacionamento causal encontrado para outros contextos, pessoas ou situações.

Validade estatística. Dois principais erros podem interferir na validade estatística dos estudos experimentais: o erro do Tipo I e o erro do Tipo II. Montgomery (2005) define o erro do Tipo I quando o pesquisador erroneamente rejeita a hipótese nula, inferindo que existe uma relação causal, quando de fato ela não existe. Para não incorrer no erro do Tipo I, na psicologia e áreas afins (p. ex. comportamento do consumidor), Kelley e Preacher (2012) lembram a convenção estabelecida de que, se o valor de p for igual ou inferior a 0,05 (nível de significância), a hipótese nula deve ser rejeitada.

Já em relação ao erro do Tipo II, Montgomery (2005) sugere que ele ocorre quando o pesquisador não rejeita uma hipótese nula que é falsa, e, conseqüentemente, infere que a relação causal não existe, quando de fato ela existe. Diante disso, quanto menor o nível de significância utilizado (para reduzir o erro do Tipo I), maior será a probabilidade de ocorrência de erros do Tipo II. No mesmo sentido, Fern e Monroe (1996) e Ye, Marionova e Singh (2011) afirmam que amostras pequenas podem aumentar a

chance de ocorrência de erros do Tipo II, uma vez que maiores tamanhos amostrais possuem uma clara relação com a significância estatística.

Validade de construto. Em marketing, vários estudos fazem uso de construtos (para uma discussão mais ampla sobre o conceito de construtos, ver Peter, 1981). Nos experimentos, a validade dos construtos refere-se a como o construto é compreendido e avaliado, seja quando utilizado como variável independente ou dependente. Quando o construto utilizado constitui uma variável independente categórica, ou seja, um fator que está sendo manipulado no experimento, o pesquisador pode avaliar a validade do construto utilizando a análise de validade de conteúdo, que pode ser feita por pares. Além disso, a checagem da manipulação também funciona como um instrumento para garantir que o construto que se deseja manipular está, de fato, sendo manipulado.

Quando o construto for uma variável dependente, o pesquisador pode avaliar a validade não somente por meio da análise do conteúdo dos itens da escala que mensuram o construto, mas também usando testes estatísticos. Para tanto, sugere-se que o pesquisador avalie a validade dos construtos por meio da análise de consistência interna (Cronbach, 1951; Garver & Mentzer, 1999; Iacobucci & Duhachek, 2003); de validade convergente (Bagozzi, Yi & Phillips, 1991; Garver & Mentzer, 1999); e de validade discriminante (Fornell & Larcker, 1981; Bagozzi et al., 1991; Garver & Mentzer, 1999). Embora essas análises sejam recomendadas para verificar a validade dos construtos, a maioria dos estudos experimentais em marketing faz uso apenas da medida de consistência interna, principalmente a medida denominada alpha de Cronbach (p. ex. Mattila & Wirtz, 2001; McFerran, Dahl, Fitzsimons & Morales, 2010).

Validade interna. No que tange à validade interna, Winer (1999) afirma que é o principal foco dos estudos experimentais. A validade interna refere-se, principalmente, às inferências de causalidade (McQuarrie, 2004), ou seja, a capacidade do pesquisador inferir que o tratamento X causou o efeito Y observado. Neste sentido, de acordo com Anderson e Bushman (1997), se o estudo conclui que a variável independente causou mudanças na variável dependente, então pode-se afirmar que há alta validade interna; por outro lado, se o estudo deixa em aberto possíveis explicações alternativas para a relação observada, então tem-se baixa validade interna. Anderson e Bushman (1997) ainda mencionam que uma forma para reduzir as explicações alternativas é utilizar processos de mediação, explicando o mecanismo pelo qual a variável independente causa os efeitos observados na variável dependente.

De acordo com Zhang e Shrum (2009), estudos de campo possuem menos validade interna do que estudos laboratoriais, executados com maior controle e com distribuição aleatória dos sujeitos às condições experimentais. No mesmo sentido, Zourrig, Chebat e Toffoli (2009) afirmam que estudos usando

cenários e controles laboratoriais podem ser uma ameaça à validade externa. No entanto, eles permitem controlar os efeitos dos estímulos entre a amostra de participantes, e também proporcionam um estímulo padronizado para todos os sujeitos, o que aumenta a capacidade do pesquisador de inferir sobre a relação causal.

Para Schram (2005) e Roe e Just (2009), há um *trade-off* entre a validade interna e a validade externa, ou seja, um aumento na validade interna faz com que o estudo perca realismo e capacidade de generalização (menor validade externa), enquanto uma maior aproximação da realidade e a busca pela generalização (maior validade externa) pode conduzir a uma falta de controle e, conseqüentemente, ao surgimento de explicações alternativas para os resultados encontrados (menor validade interna).

Validade externa. A validade externa ou ecológica (Shadish et al., 2001) é o tipo de validade mais discutida nos estudos em marketing (Winer, 1999; McQuarrie, 2004). Segundo Winer (1999) e Lynch Jr. (1999) há três principais perspectivas que interferem na validade externa dos experimentos. Primeiro, a generalização estatística dos resultados, ou seja, o problema de usar uma amostra limitada e considerar que os efeitos encontrados nela possam ser generalizados para a população de interesse. Uma segunda perspectiva recai sobre a robustez dos efeitos encontrados, o que condiz com a possibilidade de expandir os resultados para outros contextos, outros sujeitos ou até outros momentos temporais. Por fim, considera-se que o realismo também possa afetar a validade externa dos experimentos, uma vez que tanto tarefas, estímulos e tratamentos podem ser distantes da realidade e, por isso, tornar difícil a transposição dos resultados para o mundo real.

No que tange à generalização estatística dos resultados de um estudo experimental para uma população, Lynch Jr. (1999; 1982) sugere que ela pode ser inviável, devido aos tamanhos amostrais comumente usados. No entanto, quando o pesquisador seleciona a amostra, pode escolher sujeitos com características que têm maior probabilidade de influenciar as variáveis dependentes. Contemplar estas características usando amostragem por quotas (Lynch Jr., 1999) pode ser uma maneira efetiva de aumentar a validade externa dos resultados.

A maioria dos estudos experimentais mantém fixo o contexto de pesquisa. No entanto, de acordo com Lynch Jr. (1999; 1982), o contexto no qual os indivíduos se encontram ou são expostos pode impor um viés aos resultados. Diante disso, os pesquisadores podem alternar os contextos utilizados entre um estudo e outro, ou até mesmo dentro de um mesmo estudo, procurando replicar os resultados em diferentes contextos de interesse. É importante que o pesquisador controle o contexto e teste as suas interações com as demais variáveis independentes utilizadas no estudo.

Em relação ao realismo, Aronson, Ellsworth, Carlsmith e Gonzales (1990) afirmam que os

tratamentos experimentais devem, além de representar as variáveis conceituais, atentar para a realidade percebida pelo sujeito. Assim, tanto na área de marketing quanto na área da psicologia social, os experimentos conseguem amenizar as suposições e interpretações dos participantes sobre o assunto da pesquisa ao mimetizarem situações naturais. No entanto, sugerir ao participante pensar numa situação real (como o tipo de relacionamento que ele mantém com uma marca) pode fazer com que a variabilidade advinda de experiências anteriores interfira nas relações observadas pelo estudo experimental. Diante disso, Lynch Jr. (1999) defende que o realismo não seja confundido com a realização de estudos de campo ou de laboratório, mas sim com a realidade envolvida no estudo e com o realismo do contexto.

Pan e Siemens (2011, experimento 2) manipularam a densidade do varejo por meio de fotos com diferentes números de pessoas para poder verificar seus efeitos sobre os consumidores. Para tanto, os autores desenvolveram um pré-teste com a apresentação de três fotos, sendo cada uma com um nível de densidade – baixa, média e alta. Em seguida, os participantes responderam a uma escala de percepção de densidade utilizada em estudos anteriores. A utilização de fotos para manipular a densidade possui realismo ou validade ecológica (para uma discussão sobre validade ecológica em manipulação de ambiente de loja, veja Bateson & Hui, 1992).

Em síntese, Winer (1999) sugere que: (1) sejam realizados múltiplos estudos, com alterações na forma de manipulação, nos sujeitos e nos procedimentos, procurando replicar os resultados; (2) sejam identificadas condições que limitam os efeitos ou que revertam os efeitos encontrados; (3) sejam mensurados a motivação e o envolvimento do sujeito com a tarefa experimental, removendo aqueles que não demonstrarem engajamento com o estudo; (4) o pesquisador atente para a realidade da tarefa experimental; e (5) sejam utilizadas covariáveis para controlar o efeito de características individuais e situacionais sobre as variáveis dependentes do estudo.

5.2 Poder

A análise do poder é apresentada por livros e artigos tanto na área de ciências sociais, comportamentais e da saúde. De acordo com Faul, Erdfelder, Lang e Buchner (2007), o poder de um teste estatístico é a probabilidade de que o teste irá rejeitar a hipótese nula quando ela realmente for falsa, reduzindo os erros do Tipo II. Se β refere-se à probabilidade de ocorrência do erro do Tipo II, o poder pode ser calculado através de $1 - \beta$ (Sawyer & Ball, 1981). Segundo Faul et al. (2007), o poder depende: do nível de probabilidade de erro do Tipo I escolhido; do tamanho da amostra utilizada para o teste e do tamanho do efeito para a hipótese a ser testada.

De acordo com Sawyer e Ball (1981), os estudos em marketing devem procurar responder a duas questões muito negligenciadas pelos pesquisadores: o estudo é sensível o suficiente para permitir a identificação de covariação? Se há evidência de covariação, qual é a força da covariação entre a causa e o efeito?

Para medir a força de um efeito, o pesquisador pode usar o *d* de Cohen (1988) e/ou o *eta* quadrado parcial para verificar o tamanho do efeito da variável independente sobre a variável dependente. De acordo com Kelley e Preacher (2012, p. 140), o tamanho do efeito é “um reflexo quantitativo da magnitude de algum fenômeno”. O *d* de Cohen é calculado pela diferença da média de dois grupos, dividida pelo desvio-padrão de ambos (Thalheimer & Cook, 2002). De acordo com Cohen (1992), os efeitos devem obedecer determinadas convenções quanto ao tamanho: valores até 0,20 são considerados pequenos, tendo 0,50 como referência desejada valores entre 0,20 e 0,80 são considerados médios e valores acima de 0,80, grandes.

Já o *eta* quadrado parcial é uma medida de intensidade de associação entre a(s) variável(is) independente(s) e a variável dependente (Cohen, 1988). Para o cálculo, a soma dos quadrados do efeito deve ser dividida pela soma dos quadrados do efeito mais a soma dos quadrados dos erros. De acordo com Cohen (1988), valores de *eta* quadrado parcial em torno de 0,01 podem ser considerados efeitos de tamanho pequeno, valores em torno de 0,06 são considerados efeitos médios e valores em torno de 0,13 são considerados grandes.

Faul et al. (2007) apresentam o G*Power 3, um programa que fornece análises de poder para testes estatísticos que utilizam distribuições normais (testes paramétricos) t, F, χ^2 . O download do programa pode ser feito através do seguinte endereço: <http://www.gpower.hhu.de/>.

5.3 Seleção da amostra

Calder et al. (1981) recomendam, quando a aplicação de uma ou mais teorias for o objetivo de um estudo, a utilização de uma amostra de respondentes homogênea, exceto quando as características individuais sejam necessárias e não possam ser manipuladas (p. ex. traços de personalidade). Além disso, duas razões justificam o uso de uma amostra homogênea na aplicação de experimentos. Primeiro, por minimizar a ocorrência de fatores que podem interferir nos resultados, uma amostra homogênea permite que o teste de hipóteses teóricas seja feito com maior exatidão do que quando são utilizadas amostras heterogêneas (Calder et al., 1981; Lynch Jr., 1982). Segundo, quando os respondentes são heterogêneos em relação a características que podem afetar as respostas, o erro de variância pode aumentar, e com isso a sensibilidade dos testes estatísticos para identificar a significância das relações pode diminuir (Calder et al., 1981).

Falk e Heckman (2009) citam que uma objeção comum aos estudos experimentais em ciências sociais é o uso de estudantes na composição da amostra. Diante disso, após os testes das proposições teóricas, e seguindo as recomendações de Calder et al. (1981), replicações do estudo podem ser feitas com amostras heterogêneas, que possam corresponder com maior proximidade à população. Essa replicação com amostra heterogênea pode garantir uma maior capacidade de generalização dos resultados (Calder et al., 1981).

Em relação aos tamanhos amostrais, embora atualmente as ferramentas estatísticas (p. ex. análises não paramétricas) apresentem soluções para amostras pequenas, recomenda-se que os estudos utilizem amostras com mais casos por grupo, observando as discussões sobre tamanhos amostrais dispostas neste artigo (p. ex. efeitos da amostra no poder). Hair, Anderson, Tatham e Black (2005) afirmam que amostras com mais de 30 casos já tendem a evidenciar características de distribuição normal; assim, sugere-se que os pesquisadores utilizem ao menos 30 casos em cada condição experimental.

6 PERSPECTIVAS DO USO DE EXPERIMENTOS EM MARKETING

Nesta seção apresentamos algumas tendências e desafios do uso de experimentos em marketing, elaborados com base na observação daquilo que é publicado nos principais periódicos de marketing ou nos comentários de estudiosos críticos da área. Nossas observações não têm a pretensão de serem aceitas com unanimidade, uma vez que faltam bases empíricas para sustentá-las cientificamente. Entretanto, consideramos existirem evidências interessantes de que estas observações fazem bastante sentido no atual contexto da disciplina de marketing.

6.1 Tendências do uso de experimentos em marketing

Não é nenhuma novidade para os acadêmicos que publicar em periódicos conceituados é uma atividade cada vez mais competitiva. A partir da observação dos trabalhos publicados nos últimos 30 anos nos principais periódicos de marketing e psicologia do consumidor (JM, JCR, JMR, JAMS, JCP, IJMR), detectamos algumas tendências, como: o aumento de estudos experimentais reportados em um mesmo artigo; o uso de desenhos experimentais mais simples, ainda que conduzidos de forma cada vez mais rigorosa; e aumento do número de estudos de campo. Discutimos a seguir cada uma dessas tendências. logo após, discutimos os desafios do uso de experimentos em estudos de marketing.

6.1.1 Aumento do número de estudos experimentais publicados em um mesmo artigo

Embora artigos contendo apenas um único experimento costumassem ser a regra, mesmo nos principais periódicos, hoje em dia os artigos neles publicados descrevem no mínimo três experimentos que, além de mostrar o fenômeno principal, tratam de eliminar diversas possíveis explicações por meio de análises cada vez mais complexas (Pham, 2013).

Para ilustrar nosso argumento, comparamos a edição de dezembro de 1985 (volume 12) com a edição de dezembro de 2013 (volume 40) do *Journal of Consumer Research*. Dos 8 artigos publicados no volume de 1985, apenas 2 artigos reportavam resultados de experimentos, sendo que cada um apresentava os resultados de apenas 1 experimento. No último volume de 2013, dos 12 artigos publicados, 9 relatavam os resultados de 41 experimentos, o que dá uma média de 4,5 experimentos por artigo. E o JCR não é uma exceção, pois observamos o mesmo movimento em periódicos que publicam estudos do consumidor como JM, JMR, JCP e IJMR.

Tabachnick e Fidell (2006) alertam: “mesmo que um estudo experimental seja executado de forma apropriada, inferências sobre uma relação causal são perigosas até que o estudo seja replicado uma ou mais vezes. Os pesquisadores só começam a acreditar que existe uma relação causal entre duas variáveis quando o estudo é repetido e resultados similares são encontrados. O argumento para uma relação causal é fortalecido quando se faz a mesma pergunta de pesquisa usando várias estratégias investigatórias e atingindo os mesmos resultados todas as vezes” (p.4).

Ao contrário do que ocorre em diversas disciplinas, replicações de estudos não são aceitas para publicação nos periódicos mais conceituados de marketing (uma exceção é o IJMR que recentemente abriu uma seção exclusiva para replicações). Logo, a saída que os pesquisadores encontraram é replicar os seus próprios estudos várias vezes. Um artigo típico relata no mínimo três experimentos: no primeiro, é mostrado o efeito causal; no segundo, usando outra estratégia (uma manipulação diferente da variável independente, uma variável dependente mensurada de outra forma, ou os dados coletados de outra maneira), mostra-se o efeito de uma variável moderadora; o terceiro estudo, usando outra estratégia, demonstra o efeito de uma variável mediadora. Um quarto e quinto estudo poderiam eliminar explicações alternativas. Portanto, é natural que o número de experimentos reportados tenha aumentado de 1 ou 1,5 para 4,5 a 5 por artigo nos últimos 30 anos.

6.1.2 Experimentos mais simples

Se, por um lado, o número de experimentos por artigo aumentou, também parece que os experimentos se tornaram um pouco mais simples. O que não quer dizer menos rigorosos. Apenas para

ilustrar nosso argumento, tomamos dois artigos bastante citados na área de memória publicados na década de 80. Em seu estudo sobre a influência da memória de atributos sobre a escolha de um produto, Lynch Jr., Marmorstein e Weigold (1988) realizaram um experimento em que foram manipuladas apenas 2 variáveis independentes. Entretanto, os 108 sujeitos foram expostos a 8 condições experimentais arranjadas de maneira bastante complexa para que as quatro hipóteses do estudo pudessem ser testadas. Os dados foram coletados em dois momentos diferentes, com um intervalo de 2 dias e foram analisadas 4 variáveis dependentes. Em outro estudo da mesma época, Keller (1987) queria verificar o efeito que dicas que facilitam a recuperação de uma propaganda da memória têm sobre a avaliação de uma marca. Para tanto, o autor conduziu um experimento em que foram manipuladas 3 variáveis independentes usando um desenho misto (duas variáveis entre-sujeitos e 1 dentro-sujeitos) com 2 replicações. Foram produzidas 16 propagandas de produtos que serviram como estímulos para 4 condições experimentais e um grupo de controle. Os 200 participantes, adultos recrutados entre um grupo de pais de alunos, foram expostos a 12 propagandas de marcas diferentes pertencentes a 4 categorias de produtos. Foram mensuradas e analisadas 6 variáveis dependentes.

Por outro lado, é mais provável que um artigo publicado recentemente descreva entre 3 e 6 experimentos com desenhos relativamente mais simples, nos quais são testadas uma ou duas hipóteses por experimento. Entretanto, o rigor no relato dos experimentos não diminuiu. Embora o número de experimentos por artigo tenha aumentado e o número máximo de páginas do artigo não tenha mudado substancialmente, a saída encontrada por um número cada vez maior de periódicos é publicar parte do artigo em seus websites. Por exemplo, recentemente o JCP publicou novas diretrizes para publicação no periódico, que exigem expressamente que os autores submetam “uma descrição detalhada dos métodos, análises e resultados do estudo dentro do texto ou em um apêndice que permita a replicação do estudo; um apêndice metodológico contendo os estímulos e medidas usadas no estudo e uma declaração do papel que cada um dos autores desempenhou no estudo” (Pechmann, 2013, p.1). Nossa conclusão é que, para ser aceito em um periódico conceituado, são necessários vários estudos com diferentes níveis de complexidade, relatados de forma minuciosa para que possam ser replicados.

6.1.3 Mais experimentos de campo

Ficamos em dúvida se classificávamos este tópico como tendência ou desafio, pois embora o número de experimentos de campo publicados tenha aumentado nos últimos anos, também é verdade que o número de experimentos de laboratório ainda é muito maior. Embora estudos de laboratório ou de campo possam ser utilizados de forma isolada, Falk e

Heckman (2009) sugerem que a combinação de ambos é recomendável para a investigação da maioria dos fenômenos em ciências sociais, pois os experimentos de laboratório complementam as deficiências dos experimentos de campo e vice-versa.

Vários autores têm escrito sobre a necessidade de ampliar o realismo e a validade externa dos resultados dos estudos em marketing (por exemplo, Pham, 2013). Ainda de forma embrionária, cada vez mais autores têm atendido a este clamor com experimentos de campo. Estes, embora tenham maior validade externa, são mais difíceis de serem executados. Por exemplo, para verificar se o ajuste (do inglês *fit*) entre a concepção (do inglês *framing*) e o grau de abstração de uma mensagem incentivando a reciclagem influencia o volume de lixo reciclado pelas pessoas, White, MacDonnell e Dahl (2011) conduziram um experimento de campo em parceria com a cidade de Calgary, no Canadá. Três medidas de volume de reciclagem de lixo foram criadas e observadas nos 390 lares incluídos na amostra, em 3 momentos diferentes: na primeira semana, na terceira semana e no sexto mês de duração do experimento. O estímulo, um folheto impresso contendo um incentivo à reciclagem, foi distribuído nos lares durante a segunda semana do experimento. É importante ressaltar a complexidade deste tipo de experimento que envolveu medidas comportamentais difíceis de serem obtidas (observar quanto e que tipo de lixo foi destinado para reciclagem), observadas em 3 momentos diferentes com uma amostra probabilística de lares relativamente espalhada em um estudo que durou 6 meses.

Embora mais difíceis de serem executados, é muito provável que no futuro cada vez mais experimentos de campo sejam publicados nos principais periódicos de marketing. Mesmo estudos laboratoriais deverão aumentar o realismo dos estímulos a que os sujeitos são expostos. Isto não significa apenas a capacidade do sujeito de imaginar a situação ou de percebê-la como real, mas também como o participante do estudo imerge no estudo, como se envolve com a tarefa e como, através desta imersão, provê resultados mais confiáveis e próximos da realidade de ocorrência do fenômeno.

6.2 Desafios para o uso de experimentos em marketing

Nesta seção apresentamos dois dos principais desafios ao uso de experimentos em marketing: amostras mais representativas e maior uso de comportamentos como variável dependente.

6.2.1 Uso de amostras mais representativas

Para Pham (2013), um dos sete pecados dos estudos da psicologia do consumidor é o emprego de amostras convenientes. Pham não critica apenas a confiança excessiva em amostras compostas por estudantes, mas também o emprego de participantes do

Mturk ou de amostras exclusivamente de origem norte-americana. Segundo ele, a psicologia do consumidor (e, dizemos nós, a área de marketing como um todo) deve encorajar estudos de consumidores reais, com uma ampla variedade de condições socioeconômicas e vivendo em diferentes lugares.

De fato, são bastante raros os estudos publicados que empregam amostras de consumidores reais. Afinal, amostras de estudantes são relativamente fáceis de conseguir e muito mais baratas. Além do mais, estudantes universitários possuem as habilidades cognitivas necessárias para compreender e dar respostas aos cenários cada vez mais complexos que os pesquisadores de marketing desenvolvem. Por outro lado, as críticas aos estudos realizados exclusivamente com amostras de estudantes cresceram a tal ponto, que diversos periódicos não mais aceitam estudos realizados dessa forma (por exemplo, o *Journal of Advertising Research*). Portanto, é de se esperar cada vez mais estudos usando amostras de consumidores reais, e menos estudos com amostras compostas por estudantes. Entretanto, este é um enorme desafio considerando os custos e as dificuldades de se obter amostras de consumidores reais.

6.2.2 Estudos com observação de comportamentos

Há relativamente pouco tempo, Baumeister, Vohs e Funder (2007) afirmaram que era raro encontrar estudos investigando comportamentos verdadeiros e, quando os estudos observam um comportamento, este “é sempre realizado em posição sentada, geralmente em frente a um computador... Os movimentos dos dedos, como os toques no teclado de um computador ou as marcações à caneta, constituem a maior parte da ação humana” (p. 397). Embora Baumeister e colegas (2007) estivessem criticando os estudos publicados na área de psicologia social e da personalidade, sua crítica poderia muito bem ser aplicada aos estudos publicados na área de marketing. Continuam os autores (2007): “embora a psicologia seja a ciência do comportamento, comportamento é o mais difícil de ser encontrado nas publicações da área... A psicologia social se tornou, em anos recentes, o estudo dos tempos de reação e das respostas autopreenchidas. Algumas vezes os questionários pedem que as pessoas relatem o que fizeram, farão ou gostariam de fazer. Mais frequentemente, eles pedem às pessoas relatar o que pensam, como se sentem, ou porque fazem o que fazem”.

Assim como na psicologia, hoje em dia grande parte dos experimentos em marketing pede às pessoas que relatem seus pensamentos, sentimentos, memórias e atitudes. Entretanto, já se sabe, graças a vários estudos e observações pessoais, que as pessoas nem sempre fazem o que dizem que irão fazer ou não fizeram o que dizem ter feito. Em marketing, um dos aspectos mais importantes é conhecer os comportamentos do consumidor, como saber se ele comprou ou não, se usou ou não, se jogou fora e onde e como ou guardou, se

assistiu ou não o comercial, se deu algo de presente ou se comprou para si, ou se falou bem ou mal de um marca. Entretanto, assim como na psicologia, muitas vezes os estudos de marketing relatam medidas de intenções, atitudes, emoções e pensamentos. Este também é um desafio enorme a ser superado, porque os estudos experimentais envolvendo mensuração de comportamentos são mais sujeitos a erros e mais difíceis de serem executados. Apesar disto, é muito provável que no futuro os periódicos passem a cobrar mais estudos envolvendo medidas comportamentais.

7 CONCLUSÃO

Embora tenha como objetivo discutir os principais aspectos de um estudo experimental, este artigo está longe de ensejar um fim da discussão sobre a realização de experimentos em marketing. Uma expectativa tanto dos autores quanto da academia de marketing é que os estudos experimentais tenham cada vez mais rigor sem deixar de lado a relevância. Para tanto, artigos que discutam métodos de pesquisa são fundamentais tanto para embasar novos pesquisadores, quanto para provocar discussões entre pesquisadores mais experientes. Neste sentido, incentiva-se que artigos sobre o método experimental abordem diferentes percepções e mesmo lacunas não abordadas neste trabalho.

Os autores esperam que próximos estudos discutam com mais profundidade testes de moderação e mediação (p. ex., a utilização de variáveis mensuradas por meio de escalas categóricas ou intervalares, manipuladas ou declaradas, testes mais adequados, etc.), bem como o processo envolvido na construção de uma pesquisa com desenho experimental (a necessidade de vários estudos-piloto até se chegar ao desenho ideal para teste das hipóteses, a necessidade de planejamento e mensuração de amostras devido às condições experimentais, entre outros)

Além de discussões sobre o método experimental, ressalte-se também a importância de pesquisadores o utilizarem em suas pesquisas com rigor metodológico com o objetivo de aprofundar a compreensão dos fenômenos mercadológicos. Assim, é importante que os pesquisadores atentem para os elementos mencionados neste trabalho e também para os desafios e tendências dos estudos experimentais em marketing.

REFERÊNCIAS

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.
- Anderson, C. A. & Bushman, B. J. (1997). External validity “trivial” experiments: The case of

laboratory aggression. *Review of General Psychology*, 1(1), 19-41.

- Aronson, E., Ellsworth, P. C., Carlsmith, J. M., & Gonzalez, M. H. (1990). *Methods of Research in Social Psychology*. 2 ed. Reading, MA: Addison-Wesley.
- Baca-Motes, K., Brown, A., Gneezy, A., Keenan, E. A., & Nelson, L. D. (2013). Commitment and behavior change: Evidence from the field. *Journal of Consumer Research*, 39(5), 1070-1084.
- Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991) Assessing construct validity in organizational research. *Administrative Science Quarterly*, 36(3), 421-458.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Bateson, J. E. G., & Hui, M. K. (1992). The ecological validity of photographic slides and videotapes in simulating the service setting. *Journal of Consumer Research*, 19(2), 271-281.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, 2(4), 396-403.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: John Wiley and Sons.
- Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, 20, 115-140.
- Bonoma, T. V. (1985). Case research in marketing: Opportunities, problems, and a process. *Journal of Marketing Research*, 22(2), 199-208.
- Calder, B. J., Phillips, L. W., & Tybout, A. M. (1981). Designing research for application. *Journal of Consumer Research*, 8(2), 197-207.
- Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982). The Concept of External Validity. *Journal of Consumer Research*, 9(3), 240-244.
- Chae, B., Li, X., & Zhu R. (2013). Judging Product Effectiveness from Perceived Spatial Proximity. *Journal of Consumer Research*, 40(2), 317-335.

- Chan, E., & Sengupta, J. (2010). Insincere Flattery Actually Works: A Dual Attitudes Perspective. *Journal of Marketing Research*, 47(1), 122-133.
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1-8.
- Choi, J. A., Koo, M., Choic, I., & Auh, S. (2008). Need for cognitive closure and information search. *Psychology & Marketing*, 25(11), 1027-1042.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2. Ed. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power prime. *Psychological Bulletin*, 112(1), 155-159.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Estrela, V. B., & Botelho, D. (2006). Efeito de atividades promocionais no varejo. *Revista de Economia e Administração*, 5(3), 297-311.
- Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326, 535-538.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. L. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Fern, E. F., & Monroe, K. B. (1996). Effects-Size estimates: Issues and problems in interpretation. *Journal of Consumer Research*, 23(2), 89-105.
- Ferraro, R., Kirmani, A., & Matherly, T. (2013). Look at Me! Look at Me! Conspicuous brand usage, self-brand connection, and dilution. *Journal of Marketing Research*, 50(4), 477-488.
- Fishbein, M., & Ajzen, I. (1972). Attitudes and opinions. *Annual Review of Psychology*, 23, 487-543.
- Fitzsimons, G. J. (2008). Death to dichotomizing. *Journal of Consumer Research*, 35(1), 5-8.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), 382-388.
- Garver, M. S., & Mentzer, J. T. (1999). Logistics research methods: Employing structural equation modeling to test for construct validity. *Journal of Business Logistics*, 20(1), 33-57.
- Ge, X., Häubl, G., & Elrod, T. (2012). What to say when: Influencing consumer choice by delaying the presentation of favorable information. *Journal of Consumer Research*, 38(6), 1004-1021.
- Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin*, 83(2), 314-320.
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise multivariada de dados*. 5. ed. Porto Alegre: Bookman.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408-420.
- Hernandez, J. M. C., Han, X., & Kardes, F. R. (2014). Effects of the perceived diagnosticity of presented attribute and brand name information on sensitivity to missing information. *Journal of Business Research*, 67(5), 874-881.
- Hernandez, J. M. C., & Marinelli, F. M. (2013). Seeing is not believing: The influence of advertising skepticism on brand extension evaluations. In XXXVII Encontro Nacional da Anpad, Anais, Rio de Janeiro.
- Hunt, S. D. (2010). *Marketing theory: Foundations, controversy, strategy, resource-advantage theory*. Armonk, NY: M.E. Sharpe.
- Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology*, 13(4), 478-487.
- Irwin, J. R., & McClelland, G. H. (2001). Misleading heuristics and moderated multiple regression models. *Journal of Marketing Research*, 38(1), 100-109.
- Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, 40(3), 366-371.

- Jin, L., He, Y., & Zhang, Y. (2014). How power states influence consumers' perceptions of price. *Journal of Consumer Research*, 40(5), 818-833.
- Keller, K. L. (1987). Memory factors in advertising: The effect of advertising retrieval cues on brand evaluations. *Journal of Consumer Research*, 14(3), 316-333.
- Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, 17(2), 137-152.
- Lee, A. Y., Keller, P. A., & Sternthal, B. (2010). Value from regulatory construal fit: The persuasive impact of fit between consumer goals and message concreteness. *Journal of Consumer Research*, 36(5), 735-747.
- Lynch Jr., J.G., Marmorstein, H., & Weigold, M.F. (1988). Choices from sets including remembered brands: Use of recalled attributes and prior overall evaluations. *Journal of Consumer Research*, 15(2), 169-184.
- Lynch Jr., J. G. (1982). On the external validity of experiments in consumer research. *Journal of Consumer Research*, 9(3), 225-239.
- Lynch Jr., J. G. (1999). Theory and external validity. *Journal of the Academy of Marketing Science*, 27(3), 367-376.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19-40.
- Mattila, A., & Wirtz, J. (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of Retailing*, 77(2), 273-289.
- Maxwell, S. E. & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, 113(1), 181-190.
- Mazzon, J. A., & Hernandez, J. M. C. (2013). Produção científica brasileira em marketing no período 2000-2009. *Revista de Administração de Empresas*, 53(1), 67-80.
- McFerran, B., Dahl, D. W., Fitzsimons, G. J., & Morales, A. C. (2010). I'll have what she's having: Effects of social influence and body type on the food choices of others. *Journal of Consumer Research*, 36(6), 915-929.
- McQuarrie, E. F. (2004). Integration of construct and external validity by means of proximal similarity: Implications for laboratory experiments in marketing. *Journal of Business Research*, 57(2), 142-153.
- McQuarrie, E. F., & Mick, D. G. (2003). Visual and verbal rhetorical figures under directed processing versus incidental exposure to advertising. *Journal of Consumer Research*, 29(4), 579-587.
- Montgomery, D. C. (2005). *Design and analysis of experiments*. 7. ed. New York: John Wiley & Sons, Inc.
- Pan, Y., & Siemens, J. C. (2011). The differential effects of retail density: An investigation of goods versus service settings. *Journal of Business Research*, 64(2), 105-112.
- Pechmann, C. (2013). Editorial regarding the new submission guidelines at the Journal of Consumer Psychology. *Journal of Consumer Psychology* 24(1), 1-3.
- Peter, J. P. (1981). Construct validity: A review of basic issues and marketing practices. *Journal of Marketing Research*, 18(2), 133-145.
- Peter, J. P., & Olson, J. C. (1983). Is Science Marketing? *Journal of Marketing*, 47(4), 111-125.
- Pham, M. T. (2013). The seven sins of consumer psychology. *Journal of Consumer Psychology*, 23(4), 411-423.
- Poor, M., Duhachek, A., & Krishnan, H. S. (2013). How images of other consumers influence subsequent taste perceptions. *Journal of Marketing*, 77(6), 124-139.
- Ray, M. (1973). Marketing communication and the hierarchy of effects. In Clark, P., *New Models for Mass Communications Research*, v. 2, Beverly Hills: Sage, 147-176.
- Roe, B. E., & Just, D. R. (2009). Internal and external validity in economics research: Tradeoffs between experiments, fields experiments, natural experiments, field data. *American Journal of Agricultural Economics*, 91(5), 1266-1271.
- Sawyer, A. G., & Ball, D. (1981). Statistical power and effect size in marketing research. *Journal of Marketing Research*, 18(3), 275-290.
- Sawyer, A. G., Worthing, P. M., & Sendak, P. E. (1979). The role of laboratory experiments to test marketing strategies. *Journal of Marketing*, 43(3), 60-67.
- Schram, A. (2005). Artificiality: the tension between internal and external validity in economic

- experiments. *Journal of Economic Methodology*, 12(2), 225-237.
- Shadish, W., Cook, T., & Campbell, D. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In Leinhardt, S. (ed), *Sociological methodology*, San Francisco, CA: Jossey-Bass, 290-312.
- Spiller, S. A., Fitzsimons, G. J., Lynch Jr., J. G., & McClelland, G. H. (2013) Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research*, 50(2), 277-288.
- Tabachnick, B. G. e Fidell, L. S. (2006). *Experimental design using ANOVA*. California: Cengage.
- Thalheimer, W., Cook, S. (2002). *How to Calculate Effect Sizes from Published Research articles: A Simplified Methodology*. Documento digital, 2002. Disponível em: <http://www.bwgriffin.com/gsu/courses/edur9131/content/Effect_Sizes_pdf5.pdf>. Acesso em: 01/11/2012.
- Thomas, M., & Tsai, C. I. (2012). Psychological distance and subjective experience: how distancing reduces the feeling of difficulty. *Journal of Consumer Research*, 39(2), 324-340.
- Thomas, M., Desai, K. K., & Seenivasan, S. (2011). How credit card payments increase unhealthy food purchases: Visceral regulation of vices. *Journal of Consumer Research*, 38(1), 126-139.
- Trudel, R., & Argo, J. J. (2013). The effect of product size and form distortion on consumer recycling behavior. *Journal of Consumer Research*, 40(4), 632-643.
- Waber, R. L., Shiv, B., Carmon, Z., & Ariely, D. (2008). Commercial features of placebo and therapeutic efficacy. *Journal of the American Medicine Association*, 299(9), 1016-1017.
- White, K., MacDonnell, R., & Dahl, D. W. (2011). It's the mind-set that matters: The role of construal level and message framing in influencing consumer efficacy and conservation behaviors. *Journal of Marketing Research*, 48(3), 472-485.
- Winer, R. S. (1999). Experimentation in the 21st century: The importance of external validity. *Journal of the Academy of Marketing Science*, 27(3), 349-358.
- Wright, S., Hernandez, J.M.C., Sundar, A., Disnmore, J., & Kardes, F. (2013). If it tastes bad it must be good: Consumer naïve theories and the marketing placebo effect. *International Journal of Research in Marketing*, 30(2), 197-198.
- Ye, J., Marinova, D., & Singh, J. (2012). Bottom-up learning in marketing frontlines: conceptualization, processes, and consequences. *Journal of the Academy of Marketing Science*, 40(6), 821-844.
- Zhang, Y., & Shrum, L. J. (2009). The influence of self-construal on impulsive consumption. *Journal of Consumer Research*, 35(5), 838-850.
- Zhao, X., Lynch Jr., J.G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197-206.
- Zourrig, H., Chebat, J. C., & Toffoli, R. (2009). Consumer revenge behavior: a cross-cultural perspective. *Journal of Business Research*, 62(10), 995-1001.