

Cenários experimentais em pesquisas contábeis: pontos de destaque no desenvolvimento metodológico

Myrna Modolon Lima¹

 <https://orcid.org/0000-0003-2084-481X>

E-mail: myrnalima@alumni.usp.br

¹ Universidade de São Paulo, Faculdade de Economia, Administração, Contabilidade e Atuária, Departamento de Contabilidade e Atuária, São Paulo, SP, Brasil

Recebido em 08.12.2021 – Desk aceite em 06.01.2022 – 5ª versão aprovada em 12.03.2023

Editor-Chefe: Fábio Frezatti

Editora Associada: Jacqueline Veneroso Alves da Cunha

RESUMO

O objetivo principal deste artigo consiste em guiar o pesquisador, por meio da exposição de pontos de destaque e exemplos de literaturas anteriores, no desenvolvimento de um cenário experimental para a investigação de sua pergunta de pesquisa. O estudo enfoca a utilização de cenários experimentais na pesquisa contábil, trazendo pontos de destaque específicos que devem ser observados quanto ao desenvolvimento teórico e metodológico do instrumento. Sistematizam-se pontos de destaque de caráter metodológico pertinentes à utilização de cenários experimentais que visem a investigar perguntas de pesquisa contábeis com foco no julgamento e na tomada de decisão dos indivíduos. O impacto desta pesquisa consiste em sistematizar aspectos relevantes ao desenvolvimento de cenários experimentais que podem afetar o rigor metodológico e as validades interna, de constructo e de conclusão estatística. Consiste em um trabalho de proposição de método, com exemplos trazidos por meio de uma revisão narrativa. O estudo sistematiza aspectos relevantes a serem considerados nas escolhas dos pesquisadores quanto ao desenvolvimento de cenários experimentais, como o papel da randomização, a escolha de participantes, a minimização do viés de desejabilidade social, maneiras específicas de coleta da variável dependente e possibilidades de coleta do *pool* de participantes. Adicionalmente, discute-se o papel da utilização dos cenários experimentais diante da natureza teórica da pergunta de pesquisa, principalmente quanto à avaliação de julgamento e tomada de decisão.

Palavras-chave: experimento, cenários experimentais, pesquisa contábil, causalidade, jogos experimentais.

Endereço para correspondência

Myrna Modolon Lima

Universidade de São Paulo, Faculdade de Economia, Administração, Contabilidade e Atuária, Departamento de Contabilidade e Atuária

Avenida Professor Luciano Gualberto, 908 – CEP 05508-010

Butantã – São Paulo – SP – Brasil

Este é um texto bilíngue. Este artigo também foi traduzido para o idioma inglês, publicado sob o DOI <https://doi.org/10.1590/1808-057x20231594.en>. Trabalho apresentado no XV Congresso ANPCONT, dezembro de 2021.



1. INTRODUÇÃO

Este artigo tem por objetivo sistematizar pontos de destaque para o desenvolvimento de cenários experimentais em pesquisas contábeis no contexto brasileiro. Cenários experimentais, ou *vignettes*, são comumente usados em estudos contábeis internacionais, como visto em Hartmann e Maas (2010) e Kadous e Zhou (2019), todavia, ainda são pouco explorados na literatura nacional (Aguilar, 2017; Frezatti et al., 2015). Os cenários permitem que os pesquisadores se beneficiem da validade interna, condizente com o uso da abordagem experimental, e do teste de teorias por meio de situações mais próximas à realidade (Eckerd et al., 2021). E os cenários experimentais também são uma maneira eficiente de conciliar a validade interna do experimento com a validade externa das situações exploradas nos cenários (Aguinis & Bradley, 2014). Nesse sentido, assume-se que o uso de cenários possa trazer maiores contribuições a áreas com caráter aplicado, como a pesquisa contábil.

Estudos como Aguiar (2017) trazem contribuições substanciais ao discutir aspectos centrais do uso de abordagem experimental na pesquisa contábil. O autor aborda aspectos pertinentes ao método experimental, exemplificando possíveis desenhos da pesquisa e sua execução. Adicionalmente, Suave et al. (2021) exploram, especificamente para a pesquisa em educação contábil, o uso de metodologias experimentais como promotoras de melhorias nas práticas de ensino. Apesar das contribuições para a pesquisa contábil brasileira no tocante à exposição do método e aos apontamentos de perspectivas futuras de pesquisas, ambos os estudos abordam o emprego de experimentos como um todo, sem se aprofundarem no desenvolvimento e nas escolhas de pesquisa no cenário experimental.

2. CENÁRIOS EXPERIMENTAIS OU VIGNETTES

Cenários experimentais, também conhecidos como *vignettes*, são descrições de situações realísticas, frequentemente hipotéticas, apresentadas aos respondentes para que eles possam inferir acerca de seu comportamento, intenções e atitudes diante das informações expostas (Aguinis & Bradley, 2014; Trevino, 1992; Wason et al., 2002). Dada sua natureza experimental, os cenários introduzem informações manipuladas e randomizadas, enquanto mantêm todos os outros itens do cenário constantes, o que permite a aferição de causalidade entre as variáveis de pesquisa (Shadish et al., 2002). Não obstante, suas características mais realistas trazem

No contexto internacional, estudos como Aguinis e Bradley (2014), Collett e Childs (2011), Eckerd et al. (2021) e Lonati et al. (2018) abordaram o uso do cenário experimental ou *vignette* na pesquisa. Os autores buscaram apontar conceitos, desenvolver melhores práticas e abordar possíveis problemas do uso de cenários experimentais. Todavia, essas pesquisas possuem um foco mais amplo, abordando o uso dos cenários como método para ciências sociais ou pesquisas organizacionais.

Dessa forma, este artigo busca complementar a lacuna de pesquisa introduzida por esses dois estudos nacionais e pelas pesquisas internacionais, de maneira a aprofundar o conhecimento sobre cenários experimentais. Dito isso, o texto tem como objetivo principal sistematizar pontos de desenvolvimento de cenários experimentais em pesquisas contábeis brasileiras. Para que consiga atingir seu objetivo principal, conta com os seguintes objetivos específicos: definir cenários experimentais, trazer as principais diferenças referentes à escolha do cenário experimental diante do jogo experimental e discutir pontos de destaque pertinentes no desenvolvimento de cenários experimentais.

Em suma, busca-se contribuir com pesquisadores brasileiros no desenvolvimento de cenários experimentais como método de pesquisa, apresentando exemplos da utilização do método em pesquisas anteriores da literatura internacional e nacional. Ainda, visto que a não observância dos pontos de destaque trazidos no artigo podem prejudicar a validade interna, de constructo e externa do estudo, busca-se trazer à luz especificidades e exemplos que podem auxiliar os pesquisadores a minimizar problemas de validade em seus estudos.

maior generalização de resultados, comparativamente a outros tipos de métodos experimentais como o jogo experimental (Aguinis & Bradley, 2014). Em suma, a utilização de cenários experimentais possibilita aos pesquisadores criarem ou reproduzirem cenários reais em situações hipotéticas, aliando o realismo necessário para garantir a credibilidade do estudo e a identificação dos/das respondentes com o tema pesquisado.

Os cenários experimentais também podem ser pensados como uma forma de simulação, em que criam-se condições “dentro de um experimento que simulam ou duplicam o ambiente natural no qual os comportamentos

sendo examinados normalmente ocorreriam” (Gravetter & Forzano, 2012, p. 219, tradução nossa). Por exemplo, no caso de pesquisas em auditoria, pode-se averiguar o julgamento do(a) auditor(a) em uma avaliação de valor justo do teste de *impairment* do *goodwill* de um(a) cliente, na qual manipula-se sua motivação (intrínseca *versus* extrínseca *versus* controle) (Kadous & Zhou, 2019).

Por fim, entende-se que o objetivo dos cenários experimentais é que o(a) respondente sinta-se, psicologicamente, dentro da situação exposta, de modo a emular o tipo de comportamento ou intenção de comportamento que experimentaria nessa situação ou em situações similares (Gravetter & Forzano, 2012). Os cenários criados não servem para que se avalie o comportamento da respondente para aquele único cenário, mas para que, diante do ambiente simulado, consiga-se averiguar o comportamento da respondente em um cenário semelhante (Gravetter & Forzano, 2012). Voltando aos exemplos mencionados, uma auditora precisa tomar decisões acerca de estimativas complexas e o teste de *impairment* é uma representação fidedigna de uma dessas decisões. Ou seja, o cenário possui especificidades, mas emula um contexto que está além da situação descrita.

2.1 As Principais Diferenças entre o Uso de Cenários Experimentais *versus* os Jogos Experimentais e Suas Limitações

Uma das principais decisões dos pesquisadores ao desenvolver seu experimento é se irão usar cenários experimentais e/ou jogos experimentais. A escolha deverá ser direcionada pela pergunta de pesquisa (Kinney, 2019), dessa forma, é importante que se conheçam as diferenças e limitações do uso de cenários experimentais diante dos jogos experimentais, visto que isso fará com que os pesquisadores consigam desenvolver uns e/ou os outros de maneira mais informada.

A primeira grande diferença entre os cenários experimentais e jogos experimentais se refere ao contexto, mais precisamente, quão real é o contexto. Um ponto de convergência entre os dois tipos de experimento é acerca do realismo experimental. No realismo experimental se introduzem elementos que consigam capturar a atenção do indivíduo para um determinado tópico e que cause impacto, mas que não ofereçam detalhes que emulem a realidade (Liyanarachchi, 2007). Todavia, adicionalmente, os cenários experimentais também usam maiores detalhes, que emulam um contexto real, denominado *realismo mundano*. Define-se realismo mundano como a possibilidade de que os eventos que ocorrem no experimento serem passíveis de ocorrerem na prática (“mundo real”) (Swieringa & Weick, 1982).

No tocante ao desenvolvimento do experimento, o realismo mundano é o responsável pelo nível de detalhamento do contexto do cenário experimental, que deve ficar a critério dos pesquisadores. Esse nível de detalhamento é comumente encontrado na caracterização do(a) participante como, por exemplo, gerente ou funcionário(a) da empresa ou nos detalhes fornecidos sobre a empresa que ajudam o(a) participante a se colocar naquela posição e entender o contexto com maior facilidade, como visto em Aguiar (2021), o qual pede aos participantes que se coloquem na posição de gerentes de uma indústria de médio porte responsável pela emissão de gases tóxicos. Em contrapartida, os jogos experimentais tendem a ser mais artificiais (Arnold & Triki, 2018), inclusive na denominação dos participantes, que frequentemente são designados *jogadores*, como visto em Santos et al. (2021).

Existem discussões na literatura quanto à extensão da aplicação do realismo mundano. Certos pesquisadores afirmam que se deve procurar mitigar as possibilidades de realismo mundano, focando-se primordialmente no realismo experimental, visto que o primeiro pode interferir na qualidade da inferência ao adicionar distrações e contextos que seriam ignorados em situações da vida real (Gravetter & Forzano, 2012; Kadous & Zhou, 2018). Por exemplo, o acréscimo de contextos extremamente detalhados, que possam confundir o(a) respondente e saturá-lo(a) com a quantidade de informações, não são recomendáveis pois apenas criam um cenário irreal, devendo-se focar em aspectos que sejam relevantes para o entendimento da situação proposta (Kadous & Zhou, 2018).

Em contrapartida, outros pesquisadores abordam uma necessidade de equilíbrio entre o realismo mundano e o realismo experimental, ou de que os pesquisadores devam avaliar a necessidade do acréscimo de detalhes específicos para o desenvolvimento de sua teoria (Haynes & Kachelmeier, 1998), considerando a necessidade de que o(a) respondente se sinta, psicologicamente, dentro da situação exposta de modo a emular o tipo de comportamento que experimentaria nessa situação ou em situações similares (Gravetter & Forzano, 2012). Dessa forma, entende-se que o(a) pesquisador(a) deva fazer uma avaliação acerca dos benefícios e malefícios da inclusão de elementos contextuais ao cenário, focando-se, majoritariamente, no tipo de pesquisa realizada (Haynes & Kachelmeier, 1998; Kadous & Zhou, 2018).

Existem pontos positivos no uso de experimentos mais abstratos, mais comuns aos jogos experimentais, principalmente para o teste de teorias (Lucas, 2003), visto que uma teoria “não consegue incorporar todos os fatores que estão presentes no ambiente do indivíduo e nem consegue prever como todos esses fatores podem simultaneamente impactar o comportamento

do indivíduo” (Arnold & Triki, 2018, p. 290, tradução nossa). Todavia, não existe unanimidade na literatura ou na comunidade acadêmica acerca da superioridade de um tipo de experimento perante outro tipo.

A discussão entre realismo mundano e realismo experimental se interliga às discussões entre pesquisadores da comunidade de economia experimental e da comunidade de psicologia. No caso das pesquisas contábeis, vemos essa mesma discussão entre essas diferentes comunidades no tocante ao desenvolvimento do experimento. Esse juízo é corroborado por Haynes e Kachelmeier (1998), os quais afirmam que pesquisadores de economia experimental tendem a classificar os cenários experimentais como muito confusos, enquanto os pesquisadores de psicologia classificam os jogos experimentais como abstratos demais. Em suma, entende-se que os pesquisadores contábeis que optarem por usarem cenários experimentais devem criar um contexto que seja adequado para testar sua teoria, fazendo uso tanto do realismo experimental, necessário para qualquer tipo de experimento, quanto do realismo mundano (com parcimônia).

A segunda grande diferença se refere ao foco da pesquisa, ou seja, se o foco é econômico e/ou comportamental/cognitivo. Tal qual a discussão entre realismo mundano e realismo experimental, também é comum, em pesquisas contábeis, vermos discussões similares entre pesquisadores que enfocam teorias econômicas ou teorias psicológicas, que têm estímulos diferentes a serem investigados (Libby et al., 2002). Cenários experimentais são comumente usados em questões de pesquisa que foquem em processos de julgamento e tomada de decisão, dado que os cenários são hipotéticos e podem ser pensados para que se consigam capturar essas questões (Trevino, 1992).

Em contrapartida, os jogos experimentais se desenvolvem, em sua maioria, a partir da Teoria Econômica Clássica, principalmente a partir da Teoria dos Jogos, a qual foi encabeçada por Von Neumann e Morgenstern (2004) e Nash (1950,1951) – responsável pelo conceito denominado Equilíbrio de Nash (Onderstal, 2014, p. 60). A premissa basilar dos vários modelos da Teoria dos Jogos é a da Teoria da Escolha Racional, que prevê que “[...] uma tomadora de decisão escolhe a melhor ação de acordo com suas preferências, diante de todas as ações disponíveis para ela” (Osborne, 2000, p. 4, tradução nossa). A Teoria dos jogos busca resolver problemas econômicos que extrapolam a teoria econômica, mais especificamente, enfoca o entendimento do comportamento econômico e “[...] a tentativa de encontrar uma descrição exata do empenho do indivíduo em obter o máximo de utilidade ou [...] o máximo de lucro” (Von Neumann & Morgenstern, 2004, p. 1, tradução nossa).

Por exemplo, a Teoria da Agência possui premissas da Teoria Econômica Clássica, que avalia o homem/a mulher como indivíduo racional (“*homo economicus*”), capaz de buscar maximizar seus ganhos em cada escolha (Jensen & Meckling, 1976), o que facilita sua adequação a diferentes desenhos experimentais. Em termos práticos, é comum que pesquisas contábeis que abordam teorias econômicas usem jogos adaptados como tarefas experimentais, como o jogo do ditador (Ham et al., 2017), o jogo dos bens públicos (Tayler & Bloomfield, 2011) e o dilema do prisioneiro (Grasser et al., 2021), ao invés de usarem cenários experimentais.

Já os cenários experimentais focam-se em trazer aspectos contextuais que possam emular estímulos comportamentais/cognitivos que são necessários para diversos tipos de pesquisas contábeis, visto que, provavelmente, a respondente não conseguiria ser estimulada da mesma forma em um jogo (Libby et al., 2002). Assim, estudos que necessitam de conhecimentos específicos ou que foquem nas diferentes motivações da respondente, como auditores ou investidores, podem ser mais interessantes quando da utilização de cenários experimentais, dado seu caráter mais realista.

Ressalta-se que, ainda que os estudos contábeis enfoquem teorias econômicas majoritariamente (Hesford et al., 2006), as explicações para os fenômenos contam com perspectivas comportamentais/cognitivas. Nesse caso, pode-se averiguar que toda a literatura de folga orçamentária, por exemplo, envolve a utilização de duas ou mais teorias, sendo uma econômica e outra(s) comportamental(is). Como visto em Hartmann e Maas (2010), que une a perspectiva do conflito de agência da *controller* da unidade de negócios ao elaborar a proposta de orçamento às perspectivas comportamentais da pressão social e do nível maquiavelismo da *controller*.

A terceira grande diferença se refere a como os participantes são caracterizados em ambas as modalidades de experimento e as consequências dessa caracterização. Como dito anteriormente, os participantes dos cenários experimentais e dos jogos experimentais são classificados, respectivamente, de modo mais e menos contextual. Dada essa caracterização, o(a) pesquisador(a) deve tomar o devido cuidado em emular, de alguma forma, o realismo experimental necessário para o entendimento dos participantes. Por exemplo, situações em que existam diferenças de poder como visto entre gerentes e funcionários, auditores juniores e seniores, e entre auditores e gerentes são facilmente captadas em um cenário experimental, visto que os participantes recebem suas instruções acerca de qual papel devem assumir.

Todavia, no jogo experimental, essa relação precisa ser emulada de modo que os jogadores entendam que há uma

hierarquização entre eles, como visto em Schuhmacher et al. (2021). Ao falar sobre os efeitos do exemplo da líder em organizações que possuem objetivos sociais nas contribuições dos seguidores quanto à conquista de um objetivo em comum, os autores selecionam de modo randomizado os participantes que assumem o papel de

líder ou de seguidor, mas dão liberdade aos líderes para que possam escolher seu nível de contribuição para essa conquista, enquanto randomizam a qual tipo de contribuição da líder (baixa vs. alta) os seguidores serão expostos. Essa escolha de desenho experimental ajuda a emular as diferenças hierárquicas de cada posição.

3. CENÁRIOS EXPERIMENTAIS OU VIGNETTES NA PRÁTICA: PONTOS DE DESTAQUE QUANTO AO DESENVOLVIMENTO

Como em qualquer desenho experimental, os cenários experimentais tem como objetivo primordial testar relações de causalidade entre variáveis (Shadish et al., 2002). Ou seja, o(a) pesquisador(a), ao manipular uma variável independente vs. de modo a eliminar possíveis fatores confundidores, estabelece uma relação causal entre essa variável e uma variável dependente y (Lonati et al., 2018).

O primeiro ponto de destaque é que, tanto no caso dos cenários como em todo desenho experimental, deve-se garantir que a única mudança no experimento sejam as manipulações das variáveis independentes (juntamente à randomização dos participantes), de modo a garantir sua validade interna (Shadish et al., 2002). Em termos práticos, para o cenário experimental, isso significa que o(a) pesquisador(a) deve planejar a parte do cenário que permanecerá constante a todas as condições e manipular, seja entre-participantes ou intra-participantes, somente as variáveis independentes (Aguinis & Bradley, 2014; Kadous & Zhou, 2018).

Por exemplo, Bonner et al. (2014) usam uma série de experimentos com manipulações entre-participantes (em que cada participante faz parte de somente uma condição) e intra-participantes (em que os participantes fazem parte de mais de uma condição) para avaliar os efeitos da contabilidade mental (Thaler, 1999): (i) na escolha da agregação/desagregação de itens da demonstração de resultados pelas gerentes; e (ii) na avaliação das empresas pelos investidores diante da agregação/desagregação dos itens.

Assumindo o papel de *chief executive officer* (CEO), os participantes avaliam entre agregar ou desagregar os itens conforme as condições apresentadas no experimento um (condição de pequenos ganhos vs. condição de grandes perdas), no experimento dois (condição de grandes ganhos vs. condição de pequenas perdas) e experimento três (condição de transparência alta vs. condição de transparência extremamente alta). Já no experimento quatro, os participantes assumem o papel de investidor responsável por escolher entre duas companhias que apresentam seus resultados de modo agregado ou desagregado (intra-participantes) para realizar seus

investimentos. Adicionalmente, os participantes-investidores recebem manipulações acerca dos resultados financeiros das empresas (entre-participantes).

Em suma, no que concerne as diferenças de cenários e as manipulações, Bonner et al. (2014) mantêm os cenários apresentados aos participantes constantes, em que todas as informações permanecem iguais, independentes da condição a qual o(a) participante é alocado(a), alterando-se somente as informações pertinentes a cada condição intra-participantes e entre-participantes. Por fim, fica a critério do(a) pesquisador(a) usar um cenário experimental previamente validado, ligeiramente modificar um cenário já validado ou criar um cenário próprio que se adeque ao seu problema de pesquisa (lembrando de validá-los por meio de pré-testes), todavia, deve-se garantir a observância do primeiro ponto de destaque para que a validade interna do estudo não seja questionada.

O segundo ponto de destaque é acerca dos participantes usados no experimento, principalmente no caso dos cenários experimentais. Existe uma discussão bastante recorrente acerca de quem devem ser os participantes dos estudos experimentais (Ashton & Kramer, 2006; Trottier & Gordon, 2018). É comum que estudantes, por exemplo, sejam usadas pelos pesquisadores para a pesquisa experimental, inclusive para os cenários experimentais (Ashton & Kramer, 2006). Os motivos para a escolha são claros: a disponibilidade de estudantes é maior do que de auditores, contadores ou gerentes e sua responsabilidade ao teste das teorias pesquisadas tende a evitar elementos *confounding* na pesquisa (como diferentes experiências que levam a interpretações distintas, por exemplo) (Libby et al., 2002; Mortensen et al., 2012).

Existe certo consenso acerca dos contextos em que a utilização de estudantes-participantes é possível, como para a investigação de respostas comportamentais (tomada de decisão), que focam em respostas concernentes à situação apresentada (seja um cenário experimental ou um jogo experimental) (Mortensen et al., 2012). Ainda, estudos como Houghton e Hronsky (1993) demonstram empiricamente a comparabilidade cognitiva entre grupos de estudantes e de profissionais de contabilidade.

Adicionalmente, Liyanarachchi e Milne (2005) também encontraram resultados similares entre comportamentos de estudantes e de investidores quanto ao padrão de alocação de recursos em investimentos de curto e longo prazo de estudos. Na mesma linha, Elliott et al. (2007) encontram resultados que asseveram o uso de estudantes de MBA como uma boa *proxy* para investidores individuais quando a complexidade da tarefa é adequada para o conhecimento dos estudantes. Por fim, Trottier e Gordon (2018) também encontram resultados demonstrando que estudantes tomam decisões similares a gerentes acerca de avaliações de *impairment* de ativos. Todavia, alguns cuidados e preocupações precisam ser tomados ao usar estudantes de modo indiscriminado em pesquisas. Assim, é necessário que o(a) pesquisador(a) avalie o quanto de conhecimento é necessário para a interpretação e avaliação correta do cenário experimental (Libby et al., 2002; Liyanarachchi, 2007).

Kadous et al. (2019) usaram três experimentos para averiguar os efeitos de motivação intrínseca, ênfase de liderança nas metas e ambiguidade da questão de auditoria na possibilidade do membro do time de auditoria endereçar suas preocupações (“*speak up*”) sobre questões de auditoria, mais especificamente, problemas que possam prejudicar uma auditoria eficaz. Os autores adotam diferentes amostras para os três experimentos, avaliando o conhecimento necessário da participante para cada estudo. No primeiro experimento, para avaliar os efeitos da motivação intrínseca (interna vs. externa vs. controle), os autores usam alunos de graduação (seniores) e pós-graduação que tinham pelo menos um ano de experiência como estagiários de auditoria e que majoritariamente tinham aceitado empregos em firmas de auditoria. Esses alunos serviram como *proxy* para *staff auditors*, visto que possuíam o conhecimento necessário para o desempenho das mesmas tarefas desses auditores.

Perceba que os autores usam uma variável cognitiva (motivação intrínseca) para averiguar uma pergunta de pesquisa que depende do conhecimento formal da participante quanto a essas questões de auditoria. Nessa instância, a amostra de participantes é suficiente para que se garanta a compreensibilidade do caso apresentado, uma vez que todos já haviam realizado tarefas semelhantes.

Aguiar & Suave (2021) abordam as preferências dos gerentes por agregação de informações em reportes internos de resultados (ganhos vs. perdas) e o propósito do relatório (revisão de metas vs. avaliação de desempenho). Os autores explicam que por usarem a teoria da contabilidade mental (Thaler, 1999) para explicar as relações entre as variáveis, que consiste em uma teoria cognitiva, não há necessidade de uma amostra sofisticada (com profissionais), alinhando-se às recomendações de Libby et al. (2002).

Adicionalmente, é necessário avaliar outros pontos que podem influenciar a escolha dos participantes, como o conhecimento adquirido por meio de experiência (Mortensen et al., 2012). Os experimentos dois e três de Kadous et al. (2019) usam *staff auditors* de uma empresa de auditoria, tendo em vista a necessidade de que os participantes tenham experiências com auditores seniores como supervisores de modo a emular o comportamento da participante na mudança de ênfase da liderança (focada em metas intrínsecas vs. focada em metas extrínsecas). Da mesma forma, elementos como idade, cultura e gênero também devem ser observados na escolha de participantes quando existe uma ênfase do(a) pesquisador(a) nesses elementos (Trottier & Gordon, 2018). Por fim, entende-se que a escolha por usar estudantes ou profissionais na pesquisa deve ser pautada pela pergunta de pesquisa e pela teoria que explica o fenômeno analisado pelo(a) pesquisador(a).

O terceiro ponto de destaque é a preocupação com o viés de desejabilidade social (Eckerd et al., 2021; Lonati et al., 2018). O viés de desejabilidade social está alinhado à tendência humana de desejar parecer mais altruísta, mais inteligente, mais honesta, dentre outros adjetivos, também pode ser uma característica da personalidade do indivíduo, em que existe uma necessidade de aprovação social, ou uma resposta a um tema/tópico, fundamentado na questão mais específica à pesquisa (Randall & Fernandes, 1991). Apesar das diferenças das fontes de desejabilidade, a literatura aponta maneiras similares de lidar com o viés.

Apesar de não ser exclusivo para a pesquisa contábil, o(a) pesquisador(a) da área deve se preocupar com esse viés dada a característica de muitas pesquisas em avaliar a tomada de decisão em contextos éticos/antiéticos e honestos/desonestos (Chung & Monroe, 2003). Em termos práticos, o viés de desejabilidade social pode afetar a forma como o(a) participante se vê e, por consequência, como ela responde acerca de seu comportamento ou intenção de comportamento (Wason et al., 2002). Ou seja, é necessário que o(a) pesquisador(a) busque maneiras de mitigar o problema para que possa garantir que seus resultados não sejam afetados pelo viés da participante.

As possibilidades para lidar com o problema não são excludentes, podendo ser usadas simultaneamente. A primeira providência possível para reduzir o viés de desejabilidade social é garantir ao/a participante a anonimidade da pesquisa. Pesquisas mostram que o nível de viés de desejabilidade social é inversamente proporcional à anonimidade da pesquisa (Nederhof, 1985). O(A) pesquisador(a) deve, portanto, garantir à participante que suas respostas serão anonimizadas, que a identificação do(a) participante será impossível e que o caso é fictício ou hipotético. Esses cuidados trazem maior

garantia ao/à participante quanto a falta de relação do cenário com situações pessoais ou consequências reais (Randall & Fernandes, 1991). Alguns pesquisadores também argumentam acerca da possibilidade de substituir a descrição detalhada do propósito do estudo por uma alternativa mais genérica (Larson, 2019), por exemplo, explicando que o objetivo do estudo é inquirir sobre a tomada de decisão em organizações. O(A) pesquisador(a), todavia, deve tomar cuidado para não deliberadamente enganar o(a) participante, dado que o ato é, no mínimo, mal visto pelos comitês de ética de pesquisa (Eckerd et al., 2021).

Outra possibilidade para dirimir o viés de desejabilidade social envolve a decisão do(a) pesquisador(a) quanto à construção do cenário experimental em primeira pessoa (forma direta) ou terceira pessoa (forma indireta) do ponto de vista do(a) participante. Pesquisas mostram que indivíduos tendem a acreditar que são mais éticos do que seus colegas (Randall & Fernandes, 1991). Isso faz com que perguntas de auto-reporte (em primeira pessoa) sejam passíveis de questionamento quanto à confiabilidade dos dados, uma vez que o(a) participante pode estar sujeita ao viés de desejabilidade social. Ou seja, o(a) participante pode acreditar, ao responder, que teria atitudes mais éticas e/ou honestas ao comparar suas atitudes em uma situação real similar.

Assim, é comum que muitos pesquisadores escrevam o cenário experimental e questionem o(a) participante acerca de seu comportamento de modo indireto, especificamente na terceira pessoa (Fisher, 1993; Wason et al., 2002). Por exemplo, ao invés de formular o cenário pedindo ao/à participante que assuma o papel de gerente, como em Johnson et al. (2012), o(a) pesquisador(a) pode pedir aos participantes que imaginem um(a) gerente na mesma situação. Alguns pesquisadores, ao criar o cenário e inquirir o(a) participante em primeira pessoa também adicionam uma segunda mensuração da variável dependente questionando o(a) participante sobre como ela imagina que outras pessoas na situação se comportariam, como visto em Murphy et al. (2019).

Outra forma de mitigar a possibilidade de viés de desejabilidade social é trazer consequências reais à decisão do(a) participante (Lonati et al., 2018). Por exemplo, em Church et al. (2019), o(a) participante assume o papel de um(a) gerente de divisão que precisa tomar uma decisão acerca de investimentos relacionados a aspectos de *corporate social responsibility* (CSR). Diante da manipulação da base da mensuração do investimento (financeira vs. não financeira) e da mensuração da norma pessoal do(a) participante acerca de CSR (apoiador vs. não apoiador), os autores capturam o comportamento da mesma ao fazê-la sofrer uma consequência econômica real conforme seu investimento. No cenário, o(a) gerente possui fundos

discricionários que podem ser gastos (US\$ 0 a US\$ 1.000) para que se plantem árvores. De modo a operacionalizar consequências econômicas reais, os autores estabeleceram aos participantes que 1% do valor dos investimentos para o plantio de árvores seria doado a uma organização não governamental (ONG) que se dedicasse a esse fim e que 1% de todo o valor que não fosse investido no plantio seria dado ao participante no final do experimento.

Outra possível forma de averiguar o viés de desejabilidade social é mensurar esse efeito por meio de alguma escala. Isso acontece em Johnson et al. (2012), os quais usaram uma versão reduzida da escala de Marlowe-Crowne (Crowne & Marlowe, 1960). A utilização da escala de desejabilidade social se alinha também à possibilidade da desejabilidade ser uma característica da respondente e não necessariamente um viés quanto ao tema do(a) pesquisador(a) (Randall & Fernandes, 1991). Na prática, a mensuração pela escala permite a utilização do resultado como variável de controle para os resultados do(a) pesquisador(a).

Por fim, diante das diversas formas de mensuração e controle de viés de desejabilidade social, é importante que o(a) pesquisador(a) enfoque a minimização do viés e avalie qual o melhor método a ser usado diante de sua própria pergunta de pesquisa, de modo a evitar problemas com a validade de constructo visto que a mensuração da variável foi incorreta.

O quarto ponto de destaque é a escolha entre inquirir acerca de um comportamento da participante ou de uma intenção de comportamento. Diferentemente dos jogos experimentais, que avaliam o comportamento do(a) participante diante das intervenções, o cenário experimental frequentemente avalia intenções de comportamento (Wason et al., 2002). Isso significa que o(a) pesquisador(a) pergunta ao/à participante qual é sua intenção, dadas as intervenções do experimento, em engajar em algum tipo de comportamento. Como visto em Austin et al. (2021), as quais averiguavam a intenção do(a) participante em investir um valor hipotético entre US\$ 0 a US\$ 10.000 na ação da empresa apresentada.

É importante compreender que a maioria dos cenários experimentais usa a mensuração de intenções diante da impossibilidade de aferir consequências reais ao/à participante (Eckerd et al., 2021). Ainda, o código de ética de pesquisa brasileiro não permite que os participantes sejam remunerados pela sua participação em qualquer tipo de estudo (inclusive os experimentos em ciências sociais/ciências sociais aplicadas) visto que se pauta na Resolução n. 466/2012 do Ministério da Saúde. Essa resolução torna mais difícil a realização de estudos em que os participantes tenham uma remuneração pela participação ou uma remuneração variável dependente do seu comportamento no experimento. Dessa forma,

a utilização da intenção como medida para a variável dependente pode ser uma boa alternativa.

O quinto ponto de destaque é sobre onde buscar o pool de participantes desejados para a pesquisa. Ou seja, quem são os participantes que responderão à pesquisa. A escolha do pool de participantes interliga-se a uma série de avaliações posteriores como a possibilidade do(a) pesquisador(a) administrar pessoalmente o experimento ou fazê-lo *on-line*, e, se precisará de contatos dos participantes ou se usará plataformas de *crowdsourcing* para aplicação do experimento (p. ex., *MTurk*, *Prolific.ac*, *LinkedIn* etc.). Em suma, o(a) pesquisador(a) deve avaliar a disponibilidade e os custos para a realização da pesquisa.

Diante disso, dada a possibilidade de utilização de estudantes de graduação ou pós-graduação, conforme visto no segundo ponto de destaque, é comum que pesquisadores conduzam seus experimentos em laboratórios ou em sala de aula. Laboratórios experimentais para as áreas de pesquisas sociais aplicadas, sociais e humanas são bem diferentes dos laboratórios usados para pesquisas das áreas de biológicas, com os instrumentos, jalecos e materiais necessários. Na prática, o laboratório experimental para ciências sociais-aplicadas é uma sala com vários computadores, que possam ser acessados pelo(a) pesquisador(a) e que tenham uma distância que permita que o(a) pesquisador(a) consiga observar os participantes enquanto participam do experimento e para que os próprios participantes não “espiem” o computador ao lado (geralmente dividindo-se os participantes em baias individuais).

O uso de estudantes também permite que o(a) pesquisador(a) aplique o experimento em sala de aula, com alunos de cursos de graduação, mestrado ou MBA. Esses experimentos são frequentemente feitos com o cenário experimental impresso em papel. Novamente, é importante garantir a distância dos participantes para que não haja nenhum contato ou conversa durante a condução do estudo. Nesse caso, o(a) pesquisador(a) também se mantém na sala durante a aplicação do experimento. Ambos os experimentos, no laboratório experimental e na sala de aula, conseguem garantir maior validade interna dado o controle do(a) pesquisador(a) durante a condução, impedindo que os participantes conversem entre si ou que “espiem” o que o colega está respondendo (Bloomfield et al., 2016; Lourenço, 2019).

Dito isso, um aspecto importante da pesquisa com cenários experimentais é que, independente da onde o experimento está sendo conduzido, seja no espaço físico do laboratório, em sala de aula, ou *on-line*, considera-se todo cenário experimental como um experimento de laboratório (Bloomfield et al., 2016). Apesar de pesquisadores como Floyd e List (2016) classificarem qualquer experimento que não envolva uma amostra de

estudantes como experimentos de campo ou experimento *artefactual*, a classificação do experimento envolve como o contexto foi definido e não sua amostra (Bloomfield et al., 2016). Dessa forma, qualquer cenário experimental, o qual foi desenvolvido por um pesquisador, é classificado como de laboratório. Outro aspecto que diferencia o experimento de laboratório do de campo é a ciência dos participantes quanto a participação no experimento. Diferentemente do primeiro, que exige um termo de livre consentimento esclarecido antes do início do estudo, o segundo, geralmente, não conta com a ciência dos participantes quanto à sua participação (Bloomfield et al., 2016; Lourenço, 2019).

Caso o(a) pesquisador(a) necessite de participantes especialistas, o experimento conduzido em sala de aula e laboratório se tornam mais complicados devido à pouca disponibilidade de tempo desses profissionais. Assim, o mais comum é que o(a) pesquisador(a) busque em redes sociais ou plataformas de *crowdsourcing* por profissionais que tenham características que se enquadrem nas suas necessidades de pesquisa. Estudos como Brandon et al. (2014) citam as redes sociais como *Twitter*, *Facebook* e *LinkedIn* como formas de distribuição para o estudo experimental, todavia, pela mesma lógica, a busca por participantes que se enquadrem em especificações do(a) pesquisador(a) também pode se dar pelas mesmas redes sociais. O *LinkedIn*, por exemplo, pode ser uma ferramenta útil para a procura de participantes profissionais, com a ressalva de que existe uma demanda de tempo considerável para coletar os contatos, contatá-los e garantir que participem da pesquisa. Ainda, o(a) pesquisador(a) deve confiar nos participantes para que respondam com atenção e pessoalmente a pesquisa, perdendo parte do controle do experimento e, por consequência, parte da validade interna.

Adicionalmente, é comum que pesquisadores usem plataformas de *crowdsourcing* como *Amazon Mechanical Turk* (*MTurk*) e *Prolific.ac* para aplicação de suas pesquisas (Aguinis & Ramani, 2021). Na prática, milhares de pessoas têm a possibilidade de cadastrar-se nas plataformas, provendo informações pessoais no cadastramento, como: idade, gênero, escolaridade, profissão, nacionalidade. E uma série de informações mais específicas, como: se trabalha com mais pessoas, se trabalha em regime de *home office*, além de outras informações. Com isso, o(a) pesquisador(a) pode selecionar, antes da realização do estudo, o *pool* específico para o qual deseja disponibilizar sua pesquisa. A quantidade de pessoas cadastradas nas plataformas traz benefícios como a facilidade de acesso e rapidez da coleta de dados (Aguinis & Ramani, 2021). Todavia, esses benefícios têm um custo considerável, visto que o pagamento é dado pelo número de participantes

e pode ser em dólares, no caso do *MTurk* ou em libras esterlinas, no caso do *Prolific.ac*.

Outro ponto negativo é a confiabilidade dos dados. No caso do *MTurk*, participantes podem responder à pesquisa sem atenção, podem mentir acerca de detalhes demográficos, podem não ter fluência suficiente de inglês para entender o cenário experimental e podem ter tido contato com cenários similares ao apresentado anteriormente (Aguinis & Ramani, 2021). Os mesmos problemas também são comuns ao *Prolific*. Todavia, por consistir em uma plataforma mais voltada à pesquisa e pelo menor volume de estudos disponibilizados diariamente, é menos provável que os participantes usem a plataforma como um meio de obter uma remuneração.

Peer et al. (2017) avaliam as diferenças entre as plataformas *Amazon MTurk*, *CloudResearch* e *Prolific*. Os pesquisadores encontraram resultados que corroboram a superioridade do *Prolific* e da *CloudResearch* diante do *MTurk* quanto à qualidade dos dados, melhor taxa de resposta e menos desonestidade.

Na prática, o(a) pesquisador(a) pode usar algumas premissas na escolha dos participantes dessas plataformas para garantir respostas de maior qualidade. Em relação ao problema de falsa identidade ou problemas de fluência em

inglês, o(a) pesquisador(a) pode elaborar perguntas que garantam a filtragem do(a) participante e, posteriormente, convidar somente participantes aprovados para um estudo específico (Aguinis & Ramani, 2021).

Outra possibilidade consiste na utilização das mesmas perguntas para validação da participante, eliminando-se do estudo os participantes que não acertarem as questões, em uma espécie de verificação de confirmação quanto às informações fornecidas pelo(a) participante. Ainda, dentro das plataformas existem alguns índices que podem ser válidos para os pesquisadores, como os índices de aprovação dos participantes (*approval rate*) que, tanto para o *MTurk* quanto para o *Prolific*, são ferramentas de filtragem para a determinação dos participantes do estudo.

Por fim, deve-se entender que existe um *trade-off* quanto ao nível de controle que o(a) pesquisador(a) possui ao usar um *pool* de participantes de uma plataforma de *crowdsourcing* em comparação a possibilidade de fazer um experimento em um laboratório físico. Todavia, a literatura anterior mostra aceitação a utilização de participantes de plataformas como *Amazon MTurk* se garantida a confiabilidade dos dados.

A Tabela 1 reúne os pontos de destaque mencionados nesse tópico.

Tabela 1

Resumo dos pontos de destaque quanto ao desenvolvimento dos cenários experimentais

Pontos de destaque	Decisões acerca dos pontos de destaque nos cenários experimentais
Manipulações	Ao construir o cenário experimental, os pesquisadores devem tomar cuidado para que a única mudança nos diferentes tratamentos (condições) sejam as manipulações.
Participantes	Os participantes devem ser selecionados de acordo com a necessidade dos pesquisadores, garantindo-se que possuam o conhecimento necessário para o entendimento do cenário.
Viés de desejabilidade social	O viés de desejabilidade social é comum em cenários experimentais que tratam, principalmente, de questões éticas. Medidas como o desenvolvimento do cenário em terceira pessoa, a mensuração da desejabilidade social por meio de uma escala ou trazer consequências reais para a decisão dos participantes ajudam a mitigar os efeitos do viés.
Comportamento/ação ou intenção de comportamento	É comum que cenários experimentais inquiram tanto sobre intenção de comportamento como sobre um comportamento.
Pool de participantes	O <i>pool</i> de participantes pode ser selecionado a partir de uma série de possibilidades como: alunos de graduação ou pós-graduação, redes sociais (p. ex., <i>LinkedIn</i>) ou plataformas de <i>crowdsourcing</i> (p. ex., <i>Amazon MTurk</i> , <i>Prolific</i>). A escolha de onde buscar o pool de participantes deve ser pautada na necessidade de conhecimentos específicos e nos custos associados a cada escolha.

Fonte: Elaborada pela autora.

4. CONCLUSÃO E POSSIBILIDADES DE PESQUISAS FUTURAS

O principal objetivo desse artigo consiste na sistematização de pontos relevantes ao desenvolvimento de cenários experimentais como método de pesquisa para os pesquisadores contábeis brasileiros. No estudo, foram exploradas as principais diferenças entre o uso dos cenários experimentais diante dos jogos experimentais, como o nível de contextualização do instrumento, o foco da pesquisa e

as escolhas pertinentes ao entendimento dos participantes quando na falta de contextualização (comum aos jogos experimentais). A seguir, foram discutidos os pontos de destaque referentes ao desenvolvimento do cenário experimental, como: a introdução das manipulações, escolha dos participantes, preocupações com viés de desejabilidade social, aferição de comportamento ou

intenção de comportamento como variável dependente, e, por fim, onde buscar seu pool de participantes. Todos esses pontos de destaque devem ser avaliados pelos pesquisadores de acordo com a temática de seus estudos e sua disponibilidade de recursos.

Temáticas comuns à literatura de contabilidade financeira como o nível de *disclosure*, forma de apresentação e características das informações divulgadas pelas empresas são bastante exploradas na literatura nacional por meio de *archival studies*. Na literatura internacional, essas mesmas temáticas também tendem a ser exploradas por meio de cenários experimentais para averiguar os efeitos dessas escolhas na tomada de decisão dos investidores, como visto em Asay et al. (2017) e Elliott et al. (2015). Nesse sentido, sugere-se a utilização dos estudos experimentais como forma de complementar e triangular os resultados (Bloomfield et al., 2016) que foram previamente encontrados em estudos com empresas brasileiras.

Temáticas comuns à literatura de contabilidade gerencial, como a influência de diferentes controles gerenciais – seja na forma de pacotes ou observados individualmente – que frequentemente são exploradas por meio de *surveys*, podem ser complementadas por experimentos que investiguem os efeitos da combinação

desses controles, como proposto em Choi (2020). Ainda, temáticas comuns à literatura de auditoria, como as relações entre clientes e auditores, julgamentos dos auditores e qualidade da auditoria, que também têm sido exploradas na literatura nacional por meio de estudos de arquivo, *surveys* e entrevistas, podem ser exploradas por meio de experimentos como o visto em Kadous e Zhou (2019).

Por fim, este estudo tem limitações quanto aos pontos de destaque, que não foram apresentados de forma exaustiva, e quanto à literatura anterior apresentada nos exemplos, que foram escolhas pontuais da autora. É essencial que o(a) leitor(a) busque um método de pesquisa que seja o mais coerente com sua pergunta de pesquisa, todavia, o avanço da investigação de teorias e a aferição de causalidade são partes fundamentais da pesquisa e podem trazer grandes avanços à pesquisa contábil brasileira. Adicionalmente, o(a) pesquisador(a) interessado(a) em desenvolver um cenário experimental conta com uma literatura internacional bastante vasta, que pode auxiliar por meio de exemplos e discussões acerca do método. Dessa forma, entende-se que o objetivo desse estudo também é incentivar a busca por informações mais granulares que garantam o desenvolvimento do cenário experimental adequado para cada pergunta de pesquisa.

REFERÊNCIAS

- Aguiar, A. B. (2017). Pesquisa experimental em contabilidade: propósito, desenho e execução. *Advances in Scientific and Applied Accounting*, 10(2), 224-244. <https://doi.org/10.14392/asaa.2017100206>
- Aguiar, A. B. (2021). Is value statement an effective informal control for stimulating pro-environmental behaviors? *Revista de Contabilidade e Finanças*, 32(86), 193-206. <https://doi.org/10.1590/1808-057x202011500>
- Aguiar, A. B., & Suave, R. (2021). Managers' presentation preferences for aggregation: The role of internal reporting outcome, purpose, and Machiavellianism. *Spanish Journal of Finance and Accounting*, 51(1), 21-49. <https://doi.org/10.1080/02102412.2021.1952517>
- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, 17(4), 351-371. <https://doi.org/10.1177/1094428114547952>
- Aguinis, H., & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, 47(4), 823-837. <https://doi.org/10.1177/0149206320969787>
- Arnold, V., & Triki, A. (2018). Use of student and online participants in behavioural accounting research. In T. Libby, & L. Thorne (Eds.), *The Routledge companion to behavioural accounting research* (pp. 287-299). <https://doi.org/10.4324/9781315710129-20>
- Asay, H. S., Elliott, W. B., & Rennekamp, K. (2017). Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *The Accounting Review*, 92(4), 1-25. <https://doi.org/10.2308/accr-51570>
- Ashton, R. H., & Kramer, S. S. (2006). Students as surrogates in behavioral accounting research: Some evidence. *Journal of Accounting Research*, 18(1), 1-15. <https://doi.org/10.2307/2490389>
- Austin, C. R., Bobek, D. D., & Harris, L. L. (2021). Does information about gender pay matter to investors? An experimental investigation. *Accounting, Organizations and Society*, 90, 101193. <https://doi.org/10.1016/j.aos.2020.101193>
- Bloomfield, R., Nelson, M. W., & Soltes, E. (2016). Gathering data for archival, field, survey, and experimental accounting research. *Journal of Accounting Research*, 54(2), 341-395. <https://doi.org/10.1111/1475-679X.12104>
- Bonner, S. E., Clor-Proell, S. M., & Koonce, L. (2014). Mental accounting and disaggregation based on the sign and relative magnitude of income statement items. *The Accounting Review*, 89(6), 2087-2114. <https://doi.org/10.2308/accr-50838>
- Brandon, D. M., Long, J. H., Loraas, T. M., Mueller, J. M., & Vansant, B. (2014). Online instrument delivery and participant recruitment services: Emerging opportunities for behavioral accounting research. *Behavioral Research in Accounting*, 26(1), 1-23.

- Choi, J. W. (2020). Studying “and”: A perspective on studying the interdependence between management control practices. *Accounting, Organizations and Society*, 86, 101188. <https://doi.org/10.1016/j.aos.2020.101188>
- Chung, J., & Monroe, G. S. (2003). Exploring social desirability bias. *Journal of Business Ethics*, 44, 291-302.
- Church, B. K., Jiang, W., Kuang, X. J., & Vitalis, A. (2019). A dollar for a tree or a tree for a dollar? The behavioral effects of measurement basis on managers' CSR investment decision. *The Accounting Review*, 94(5), 117-137.
- Collett, J. L., & Childs, E. (2011). Minding the gap: Meaning, affect, and the potential shortcomings of vignettes. *Social Science Research*, 40(2), 513-522. <https://doi.org/10.1016/j.ssresearch.2010.08.008>
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349-354. <https://doi.org/10.1037/h0047358>
- Eckerd, S., DuHadway, S., Bendoly, E., Carter, C. R., & Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, 67(2), 261-275. <https://doi.org/10.1002/joom.1128>
- Elliott, W. B., Hodge, F. D., & Kennedy, J. J. (2007). Are M.B.A. students a good proxy for nonprofessional investors? *The Accounting Review*, 82(1), 139-168. <https://doi.org/10.2308/accr.2007.82.1.139>
- Elliott, W. B., Rennekamp, K. M., & White, B. J. (2015). Does concrete language in disclosures increase willingness to invest? *Review of Accounting Studies*, 20, 839-865. <https://doi.org/10.1007/s11142-014-9315-6>
- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2), 303-315. <https://doi.org/10.1086/209351>
- Floyd, E., & List, J. A. (2016). Using field experiments in accounting and finance. *Journal of Accounting Research*, 54(2), 437-475. <https://doi.org/10.1111/1475-679X.12113>
- Frezatti, F., Aguiar, A., Wanderley, C., & Malagueño, R. (2015). A pesquisa em contabilidade gerencial no Brasil: Desenvolvimento, dificuldades e oportunidades. *Revista Universo Contábil*, 11(1), 47-68. <https://doi.org/10.4270/ruc.2015103>
- Frezatti, F., Bido, D. S., Cruz, A. P. C., & Machado, M. J. C. (2015). A Estrutura de artefatos de controle gerencial no processo de inovação: Existe associação com o perfil estratégico? *Brazilian Business Review*, 19, 129-156. <https://doi.org/10.15728/bbr.2015.12.1.6>
- Grasser, R. A., Majerczyk, M., Staehle, M., & Sydney, U. (2021). The benefits of deliberative involvement in the design of incomplete feedback systems. *Contemporary Accounting*, 38(3), 2351-2375. <https://doi.org/10.1111/1911-3846.12675>
- Gravetter, F. J., & Forzano, L. A. B. (2012). *Research methods for the behavioral sciences* (4th ed.). Cengage Learning.
- Ham, C., Lang, M., Seybert, N., & Wang, S. (2017). CFO narcissism and financial reporting quality. *Journal of Accounting Research*, 55(5), 1089-1135. <https://doi.org/10.1111/1475-679X.12176>
- Hartmann, F. G. H., & Maas, V. S. (2010). Why business unit controllers create budget slack: Involvement in management, social pressure, and Machiavellianism. *Behavioral Research in Accounting*, 22(2), 27-49. <https://doi.org/10.2308/bria.2010.22.2.27>
- Haynes, C. M., & Kachelmeier, S. J. (1998). The effects of accounting contexts on accounting decisions: A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature*, 17, 97-136.
- Hesford, J. W., Lee, S. H. S., Van der Stede, W. A., & Young, S. M. (2006). Management accounting: A bibliographic study. *Handbooks of Management Accounting Research*, 1, 3-26. [https://doi.org/10.1016/S1751-3243\(06\)01001-7](https://doi.org/10.1016/S1751-3243(06)01001-7)
- Houghton, K. A., & Hronsky, J. J. F. (1993). The sharing of meaning between accounting students and members of the accounting profession. *Accounting & Finance*, 33(2), 131-147. <https://doi.org/10.1111/j.1467-629X.1993.tb00324.x>
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm managerial behaviour, agency costs and ownership structure. *Journal of Financial Economics*, 3, 305-360. http://uclafinance.typepad.com/main/files/jensen_76.pdf
- Johnson, E. N., Fleischman, G. M., Valentine, S., & Walker, K. B. (2012). Managers' ethical evaluations of earnings management and its consequences. *Contemporary Accounting Research*, 29(3), 910-927. <https://doi.org/10.1111/j.1911-3846.2011.01135.x>
- Kadous, K., Proell, C. A., Rich, J., & Zhou, Y. (2019). It goes without saying: The effects of intrinsic motivational orientation, leadership emphasis of intrinsic goals, and audit issue ambiguity on speaking up. *Contemporary Accounting Research*, 36(4), 2113-2141. <https://doi.org/10.1111/1911-3846.12500>
- Kadous, K., & Zhou, Y. (2018). Maximizing the contribution of JDM-style experiments in accounting. In T. Libby & L. Thorne (Eds.), *The Routledge companion to behavioural accounting research* (pp. 175-192). <https://doi.org/10.4324/9781315710129-13>
- Kadous, K., & Zhou, Y. D. (2019). How does intrinsic motivation improve auditor judgment in complex audit tasks? *Contemporary Accounting Research*, 36(1), 108-131. <https://doi.org/10.1111/1911-3846.12431>
- Kinney, W. R. (2019). The Kinney three paragraphs (and more) for accounting Ph.D. students. *Accounting Horizons*, 33(4), 1-14. <https://doi.org/10.2308/acch-52451>
- Larson, R. B. (2019). Controlling social desirability bias. *International Journal of Market Research*, 61(5), 534-547.
- Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, 27, 775-810.
- Liyanarachchi, G. A. (2007). Feasibility of using student subjects in accounting experiments: A review. *Pacific Accounting Review*, 19(1), 47-67. <https://doi.org/10.1108/01140580710754647>

- Liyanarachchi, G. A., & Milne, M. J. (2005). Comparing the investment decisions of accounting practitioners and students: An empirical study on the adequacy of student surrogates. *Accounting Forum*, 29(2), 121-135. <https://doi.org/10.1016/j.accfor.2004.05.001>
- Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, 64, 19-40. <https://doi.org/10.1016/j.jom.2018.10.003>
- Lourenço, S. M. (2019). Field experiments in managerial accounting research. *Foundations and Trends® in Accounting*, 14(1), 1-72. <https://doi.org/10.1561/14000000059>
- Lucas, J. W. (2003). Theory-testing, generalization, and the problem of external validity. *Sociological Theory*, 21(3), 236-253. <https://doi.org/10.1111/1467-9558.00187>
- Mortensen, T., Fisher, R., & Wines, G. (2012). Students as surrogates for practicing accountants: Further evidence. *Accounting Forum*, 36(4), 251-265. <https://doi.org/10.1016/j.accfor.2012.06.003>
- Murphy, P. R., Wynnes, M., Hahn, T. A., & Devine, P. G. (2019). Why are people honest? Internal and external motivations to report honestly. *Contemporary Accounting Research*, 53(9), 1689-1699. <https://doi.org/10.1017/CBO9781107415324.004>
- Nash Jr, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1), 48-49.
- Nash Jr, J. F. (1951). Non-cooperative games. *Annals of Mathematics*, 286-295.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15, 263-280.
- Onderstal, S. (2014). *Economics of organizations and markets*. Pearson.
- Osborne, M. J. (2003). *An introduction to game theory*. Oxford University Press.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163. <https://doi.org/10.1016/j.jesp.2017.01.006>
- Randall, D. M., & Fernandes, M. F. (1991). The social desirability response bias in ethics research. *Journal of Business Ethics*, 10(11), 805-817.
- Santos, L. P. G., Cerqueira, A. J. F., & Carvalho, C. V. O., Júnior. (2021). Uma análise experimental do efeito da manutenção de registros sobre a reciprocidade direta. *Revista Contabilidade & Finanças*, 32(86), 359-375. <https://doi.org/10.1590/1808-057x202010800>
- Schuhmacher, K., Towry, K. L., & Zureich, J. (2021). Leading by example in socially driven organizations: The effect of transparent leader compensation contracts on following. *The Accounting Review*, 97(3), 373-393. <https://doi.org/10.2308/tar-2019-0227>
- Shadish, W. R., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Cengage Learning.
- Suave, R., Altoé, S. M. L., & Ferreira, M. M. (2021). Pesquisas experimentais aplicadas à educação contábil: Panorama atual e oportunidades no cenário brasileiro. *Revista Contemporânea de Contabilidade*, 18(47), 155-176. <https://doi.org/10.5007/2175-8069.2021.e77682>
- Swieringa, R. J., & Weick, K. E. (1982). An assessment of laboratory experiments in accounting. *Journal of Accounting Research*, 20, 56-101.
- Taylor, W. B., & Bloomfield, R. (2011). Norms, Conformity and controls. *Journal of Accountig Research*, 46(2), 1-23. <https://doi.org/10.1111/j.1475-679X.2011.00398.x>
- Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12, 183-206.
- Trevino, L. K. (1992). Experimental approaches to studying ethical-unethical behavior in organizations. *Business Ethics Quarterly*, 2(2), 121-136.
- Trottier, K., & Gordon, I. M. (2018). Students as surrogates for managers: Evidence from a replicated experiment. *Canadian Journal of Administrative Sciences*, 35(1), 146-161. <https://doi.org/10.1002/cjas.1377>
- Von Neumann, J., & Morgenstern, O. (2004). *Theory of games and economic behavior* (6th ed.). Princeton University Press.
- Wason, K. D., Polonsky, M. J., & Hyman, M. R. (2002). Designing vignette research in marketing. *Australasian Marketing Journal*, 10(3), 41-58. [https://doi.org/10.1016/S1441-3582\(02\)70157-2](https://doi.org/10.1016/S1441-3582(02)70157-2)