

Uma Abordagem para Avaliação da Qualidade do Processo de Pesquisa em Contabilidade

José Renato Sena Oliveira

<http://orcid.org/0000-0001-9820-4682> | E-mail: jrsenna@uefs.br

Gilberto de Andrade Martins

<http://orcid.org/0000-0001-5144-5286> | E-mail: martins@usp.br

Resumo

Objetivo: Associar os atributos de qualidade da pesquisa com aqueles percebidos no processo de construção da produção científica em Contabilidade.

Método: Foi utilizada a técnica Delphi Modificada, em duas rodadas, com questionários online e acesso personalizado. O painel foi composto de 41 docentes de 19 Programas de Pós-Graduação em Contabilidade brasileiros. Foi elaborada uma matriz de orientação com 53 atributos/relações vinculados a critérios gerais de qualidade e a nove características-chave, avaliados pelos respondentes de acordo com a aderência às suas práticas de pesquisa.

Resultados: Aproximadamente 3/4 das proposições receberam nível forte de concordância. Entretanto, atributos que atingiram níveis baixo ou moderado incluem itens que podem comprometer a qualidade e a integridade da pesquisa, como os relacionados aos princípios éticos, aos erros e vieses e aos impactos da participação da equipe sobre os resultados.

Contribuições: O estudo pode impactar o campo, ao ajudar pesquisadores a autoavaliarem suas investigações, identificarem deficiências e limitações, resultando em maior aceitação e encurtamento do processo de submissão aos periódicos. Consequentemente, pode criar oportunidades de financiamento e de parcerias internacionais, bem como aumentar a relevância da produção científica da área.

Palavras-chave: contabilidade – pesquisa, produção científica, pós-graduação, qualidade da pesquisa, critérios de qualidade.

Editado em Português e Inglês. Versão original em Português.

Recebido em 25/11/2019. Pedido de Revisão em 29/11/2019. Resubmetido em 4/12/2019. Aceito em 9/12/2019 por Dra. Bruna Camargos Avelino (Editora associada) e por Dr. Gerlando Augusto Sampaio Franco de Lima (Editor). Publicado em 20/12/2019. Organização responsável pelo periódico: Abracicon.

1. Introdução

A pesquisa é uma atividade importante nas instituições educacionais, sendo capaz de desenvolver e melhorar técnicas, estratégias de trabalho e cursos de ação em vários ramos de atividade, além de ser essencial para avançar o conhecimento e construir soluções para os problemas sociais. A pesquisa científica é uma atividade destinada a testar uma hipótese, permitir tirar conclusões e, assim, contribuir para o conhecimento generalizável, expresso em teorias, princípios e declarações de relações (National Institutes of Health, 1979).

O ambiente institucional da pesquisa no Brasil estabelece padrões básicos de produtividade para Programas de Pós-Graduação (PPG) (mestrado/doutorado). Esses padrões são consistentes com os requisitos das agências governamentais de fomento a pesquisadores e à infraestrutura institucional para pesquisa e são estabelecidos pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). Se, por um lado, esses padrões encorajam a produção científica em vários campos, por outro, como efeito da ânsia de publicar os resultados da pesquisa, a atenção ao rigor científico e à relevância podem ser deixados de lado. Assim, o ambiente institucional justifica estudos que analisem a qualidade da produção científica, de “como é feita”, também conhecido como processo de pesquisa.

Estudos sobre a avaliação da pesquisa contábil em nível internacional abordam vários aspectos, como a produtividade e a qualidade, envolvendo os departamentos de negócios universitários (Chan, Chang, Tong, & Zhang, 2012; Jones, Brinn, & Pendlebury, 1996; Lowe & Locke, 2005). Outros abordam critérios de qualidade de pesquisa, como o impacto (Brown & Gardner, 1985; Carmona, 2006); relevância (Reiter & Williams, 2002); rigor (Evans, Feng, Hoffman, Moser, & Van der Stede, 2015; Williams, 2014); e validade (Libby, Bloomfield, & Nelson, 2002). Finalmente, outros exploram a análise de citações (Brown & Gardner, 1985; Dunbar & Weber, 2014) e percepções de qualidade de periódicos contábeis (Ballas & Theoharakis, 2003; Brinn, Jones, & Pendlebury, 2001; Brown & Huefner, 1994; Lowe & Locke, 2005; Lowensohn & Samelson, 2006; Taylor, 2011). Embora não focados em atributos de qualidade de uma boa pesquisa, esses estudos exploram a produtividade e a avaliação da qualidade do que é publicado em periódicos científicos da área.

No contexto brasileiro, estudos sobre a pesquisa contábil adotaram diversas abordagens. Alguns apresentam uma perspectiva bibliométrica e analisam aspectos de caminhos metodológicos utilizados em teses e outras publicações em periódicos e eventos (Mendonça Riccio, & Sakata, 2009; Miranda, Azevedo, & Martins, 2011), questões éticas na pesquisa (Antunes, Mendonça, Oyadomari & Okimura, 2011) e análise de citações em periódicos (Aragão, Oliveira & Lima, 2014). Por outro lado, outro estudo analisa a produtividade acadêmica (Martins & Lucena, 2014) e, especificamente, a produtividade docente nos PPGs brasileiros, ao investigar o perfil e as principais práticas de produção científica desses Programas, destacando a ocorrência de partes dos resultados de um único estudo publicadas em vários artigos (“ciência salame”) e artigos rejeitados por revistas de maior impacto enviados às menos prestigiadas, até serem aceitos.

No entanto, a caracterização de natureza bibliométrica, baseada em relatórios e publicações, é inadequada para revelar falhas no processo de pesquisa em Contabilidade, o qual deve ser refletido desde a escolha do assunto até a submissão do relatório para disseminação científica. É importante uma reflexão sobre os elementos que contribuem para o aumento da qualidade e do caráter científico das investigações. Nessa direção, boas práticas de pesquisa são definidas como regras que os pesquisadores podem seguir para assegurar a qualidade, objetividade e integridade dos dados, com escolhas razoáveis e explícitas (Denscombe, 2010; Shamoo & Resnik, 2003).

Esses elementos sugerem que uma boa pesquisa envolve as bases e práticas apropriadas de investigação científica para alcançar a resposta mais adequada ao problema proposto. Diante desse contexto, o objetivo do estudo foi associar os atributos de qualidade da pesquisa com aqueles percebidos no processo de construção da produção científica em Contabilidade. O processo de construção da produção científica envolve as características-chave ou diferentes estágios (Mays & Pope, 2006; Spencer, Ritchie, Lewis, & Dillon, 2003) e alcança o planejamento, a execução e a avaliação de resultados (Brinberg & McGrath, 1985). Adotou-se como hipótese que determinadas práticas correntes utilizadas no processo de construção da produção científica em Contabilidade no Brasil não atendem adequadamente aos atributos de qualidade da pesquisa descritos na literatura.

Os achados do presente estudo evidenciaram que cerca de 3/4 das proposições relativas aos 53 atributos/relações de qualidade analisados receberam nível forte de concordância pelos pesquisadores da área contábil. Entretanto, 6 atributos atingiram o nível moderado e 3 o nível baixo, os quais incluem itens que podem comprometer a qualidade e a integridade da pesquisa, como os relacionados aos princípios éticos, aos erros e vieses e aos impactos da participação da equipe sobre os resultados. Além disso, 5 atributos associados aos critérios de confiabilidade, integridade e validade interna não tiveram consenso formado, achado que sugere que pode haver falhas no atendimento a esses critérios em determinadas etapas do processo de pesquisa.

A avaliação da produção científica focada nos resultados é limitada para evidenciar a PPGs e pesquisadores as possíveis falhas no processo de pesquisa. Estudos da avaliação de práticas de pesquisa podem impactar o campo, ao ajudar pesquisadores a auto avaliarem suas investigações, identificarem deficiências, limitações e pontos a serem considerados para melhorar suas escolhas metodológicas. Como resultado, pode-se alcançar maior aceitação e encurtamento do processo de submissão aos periódicos, ao diminuir o número de revisões. Estudos de maior qualidade, que fazem uma contribuição clara e têm um impacto, são levados em conta na alocação de recursos e recrutamento de professores, entre outros fatores (Chan *et al.*, 2012). Como impacto esperado, a melhoria da qualidade da produção científica dos PPGs da área contribui para aumentar a chance de publicação dos artigos em periódicos relevantes. Além disso, pode criar novas oportunidades de financiamento, colaboração em parcerias internacionais e maior relevância internacional da produção científica da área.

2. Revisão da Literatura

A revisão da literatura está dividida em três subtópicos: o primeiro apresenta uma visão geral sobre a qualidade e o que são boas práticas no contexto da pesquisa. O seguinte destaca os estágios do processo de pesquisa a partir do Esquema da Rede de Validade (Brinberg & McGrath, 1985), bem como do Framework de Avaliação de Evidências de Pesquisa (Mays & Pope, 2006; Spencer *et al.*, 2003). Finalmente, mostra evidências da literatura acerca de critérios e indicadores de qualidade na avaliação do processo de pesquisa.

2.1 Qualidade e as Boas Práticas no Contexto da Pesquisa

Para entender o objeto deste estudo, é importante ter claro o significado de seus elementos centrais: qualidade, atributo e processo de pesquisa. Em sentido genérico, a qualidade é descrita como um atributo, condição natural, propriedade pela qual algo ou alguém é individualizado que o distingue dos outros. Também como uma forma de ser, essência ou natureza, ou grau de perfeição, precisão e conformidade com um determinado padrão (Michaelis, 2012). Na pesquisa científica, Valentine (2009) diz que a resposta à pergunta “quais são as características de um estudo de alta qualidade?” depende, em parte, de porque o julgamento está sendo feito, pois diferentes atores atribuem significados distintos aos seus julgamentos. O autor defende que a qualidade se refere a um link adequado entre os objetivos do estudo e o desenho e finalidade envolvidos na implementação dele. Em relação à pesquisa contábil, Clarkson (2012) argumenta que a qualidade está associada a três fatores fundamentais: (i) contribuição — a importância do foco do estudo e seu grau de inovação; (ii) o rigor com que o estudo foi conduzido (credibilidade científica); e (iii) a capacidade do próprio documento de transmitir o estudo de forma transparente e acessível (comunicação).

O conceito de atributo também tem uma variedade de significados, incluindo os seguintes: (1) o que é próprio e peculiar a alguém ou algo; e (2) uma condição, propriedade ou qualidade de alguma coisa. Gramaticalmente, um atributo modifica um substantivo expressando sua qualidade ou determinação (Michaelis, 2012). Já o processo de pesquisa é definido como a identificação, combinação e uso de elementos e relações dos domínios conceitual, metodológico e substantivo, dividido em três etapas e com diferentes perspectivas de validade (Brinberg & McGrath, 1985). Além disso, o processo de pesquisa pode ser descrito com base nas seguintes etapas (ou estágios): achados, desenho, amostra, coleta de dados, análise e reporte (Mays & Pope, 2006; Spencer *et al.*, 2003). Face a essas definições, os atributos de qualidade no processo de pesquisa descritos neste estudo incluem um conjunto de características que demonstram as ligações entre o objeto de interesse, a evidência, a teoria e as escolhas metodológicas do pesquisador. A qualidade pode ocorrer em vários estágios de planejamento e execução da pesquisa, permeados pela tradição de pesquisa subjacente.

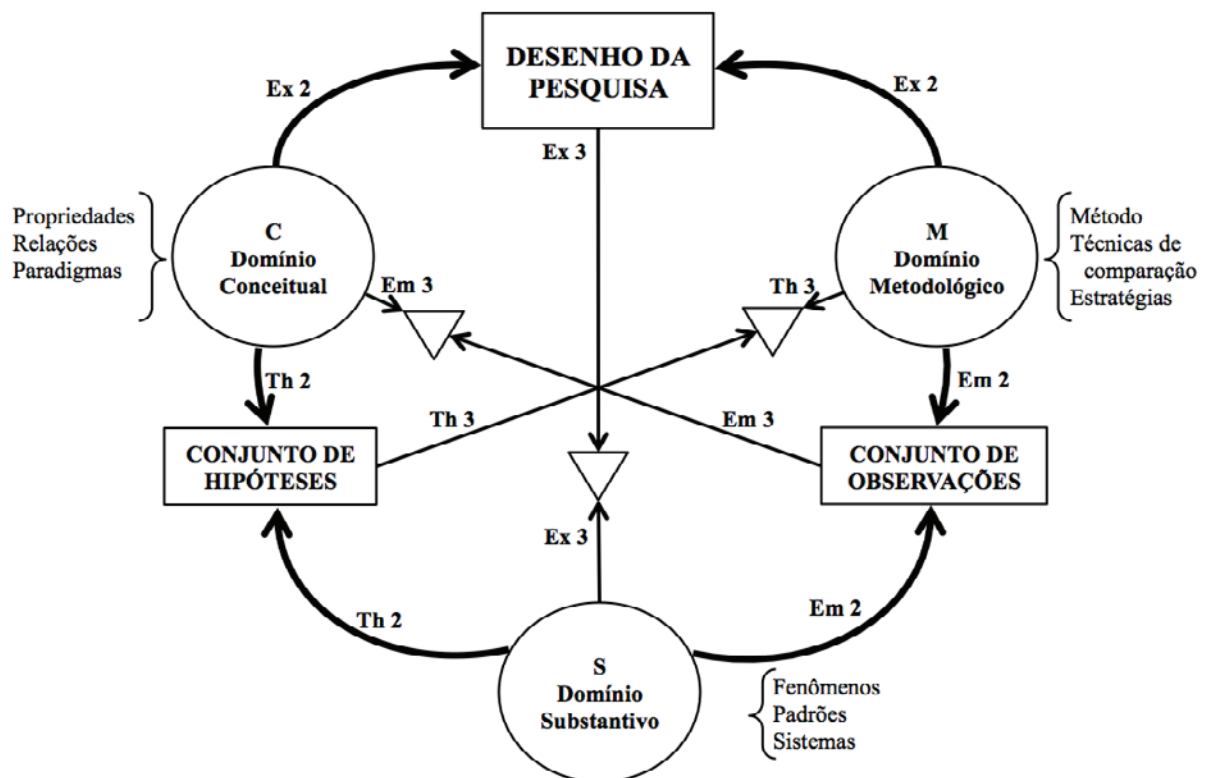
Dado o seu caráter subjetivo e valorativo, a definição de boa pesquisa não é tarefa fácil, tampouco consensual. As pessoas em posições de liderança na pesquisa podem desempenhar um papel-chave no desenvolvimento de uma cultura em que as atitudes éticas e as boas práticas de pesquisa prevalecem. Esses autores argumentam que se os principais pesquisadores, gerentes, empresas e agências governamentais demonstram e toleram atitudes antiéticas e práticas de pesquisa precárias, então essas atitudes e práticas prevalecerão (Shamoo & Resnik, 2003). Para os autores, a percepção de boas práticas de pesquisa é que as mesmas representam regras que os pesquisadores podem seguir para ajudar a garantir a qualidade, objetividade e integridade de seus dados.

Esses aspectos podem orientar uma avaliação da qualidade da pesquisa, pois representam atributos gerais de uma boa pesquisa. Elementos semelhantes constam em um estudo (Webster & Watson, 2002) que apresenta questões associadas a cada um deles: (a) contribuição (“o que há de novo?”); (b) impacto (“e daí?”); (c) lógica (“por quê?”); e (d) rigor (“bem feito?”). Os dois primeiros aspectos referem-se à “substância” de um estudo, isto é, seu problema de pesquisa, seus fundamentos, a lacuna identificada, sua justificativa e seus possíveis efeitos sobre um campo. A lógica do estudo é associada não somente à construção de um quadro teórico, mas também à escolha do desenho de pesquisa. A última envolve o rigor e os aspectos éticos associados à condução da pesquisa.

Em suma, a noção de boa pesquisa está associada a regras que os pesquisadores podem seguir para ajudá-los a garantir a qualidade, objetividade e integridade dos dados. Para isso, é necessário incluir escolhas razoáveis e explícitas como parte de cada componente da pesquisa e nesse contexto, o rigor e a qualidade apresentam um forte relacionamento, pois não é possível obter investigações de alta qualidade se o processo for mal executado, razão pela qual cada tarefa ao longo das fases (estágios) do processo de pesquisa deve ser cuidadosamente pensada.

2.2 Estágios do Processo de Pesquisa: o Esquema da Rede de Validade e o Framework de Avaliação de Evidências de Pesquisa

A pesquisa é basicamente um estudo de relações, ou seja, é sempre feita em termos de relações entre unidades (Brinberg & McGrath, 1985). Esses autores estudaram a validade no processo de pesquisa e defendem que essa não é uma *commodity* que pode ser comprada com técnicas, mas um estado ideal, um conceito a ser perseguido e que se aplica em diferentes estágios do processo de pesquisa. Dada essa caracterização, os autores apresentam o Esquema da Rede de Validade (ERV) (Validity Network Schema – VNS), em que assumem que a pesquisa envolve três domínios inter-relacionados, mas analiticamente distintos: o conceitual, o metodológico e o substantivo. Enquanto o fenômeno é o objeto central na perspectiva do domínio substantivo, o domínio conceitual se preocupa em identificar os conceitos e atributos que podem contribuir para explicar o fenômeno estudado, ou seja, uma teoria (ou lente teórica) que sustente a análise das observações coletadas. O domínio metodológico, por sua vez, se ocupa em apresentar e descrever os meios para a realização do estudo (Figura 1).



- | | | |
|---------------------------|---|---|
| Caminho Experimental (Ex) | → | Construção de um projeto, e implementá-lo usando um conjunto de eventos substantivos. |
| Caminho Teórico (Th) | → | Construção de um conjunto de hipóteses, e testá-las e avaliá-las com um apropriado conjunto de métodos. |
| Caminho Empírico (Em) | → | Construção de um conjunto de observações, e explicá-las e interpretá-las com base em um conjunto de conceitos significativos. |

Figura 1. O sistema ERV: domínios, níveis e caminhos.
Adaptada de Brinberg e McGrath (1985, p. 22). Tradução livre.

Para Brinberg e McGrath (1985), o processo completo de pesquisa envolve três grandes estágios (etapas): o Estágio 1 – Validade como Valor, denominado Pré-Estudo ou Estágio Preparatório, diz respeito à fase de planejamento da pesquisa, quando são estabelecidos os critérios para avaliação das relações dos domínios, isto é, há um julgamento de mérito da proposta de pesquisa; o Estágio 2 – Validade como Correspondência, também chamado de Estágio Central, se materializa na etapa de execução do estudo e nela se estabelece o uso de elementos e relações dos três domínios para produzir um conjunto de achados empíricos a partir dos diferentes caminhos; e o Estágio 3 – Validade como Robustez, denominado Monitoramento dos Achados do Estágio 2, etapa em que ocorre a avaliação dos resultados e se verifica a extensão e o delineamento dos achados, a fim de evidenciar as fronteiras e a contribuição do estudo. Há três caminhos alternativos para executar o estágio central do processo de pesquisa: o Experimental (Ex), o Teórico (Th) e o Empírico (Em), os quais refletem diferentes estilos de fazer pesquisa e encontrar diferentes questões de validade. Em essência, tem-se que a ligação entre o objeto de interesse (representado pelo domínio substantivo) e as relações conceituais denotadas pelas propriedades do mesmo (a teoria) sustenta a geração de um conjunto de hipóteses que poderão explicar o fenômeno. De outro lado, a conexão entre o objeto e o domínio metodológico está relacionada com as estratégias utilizadas para a coleta sistemática de um conjunto de observações que deem suporte para as evidências e achados.

Outra perspectiva sobre a realização do processo de pesquisa é o Framework de Avaliação de Evidências de Pesquisa (Mays & Pope, 2006; Spencer *et al.*, 2003), orientado para avaliar evidências em relação a nove principais características e processos:

- (1) Achados – característica associada ao julgamento de elementos como a credibilidade dos achados, como o conhecimento é expandido pelo estudo, o escopo da inferência e o cumprimento de objetivos e propósitos originais.
- (2) Desenho – justificativa para o desenho da pesquisa (escolhas metodológicas).
- (3) Amostra – avaliação de critérios para a concepção e seleção da amostra, bem como casos de inclusão e exclusão.
- (4) Coleta de dados – avaliação da condução da coleta de dados.
- (5) Análise – profundidade e complexidade dos dados, escolha da abordagem e análise da formulação, contexto das fontes de dados e diversidade de perspectivas.
- (6) Reporte – avaliação das ligações entre dados, interpretações e conclusões, bem como a coerência global do relatório.
- (7) Reflexividade e neutralidade – clareza sobre os pressupostos, perspectivas teóricas e valores que orientam o estudo e cuidado dos erros e vieses.
- (8) Ética – avaliação de como o pesquisador e sua equipe lidam com questões éticas.
- (9) Auditabilidade – um julgamento sobre o cumprimento dos procedimentos formais e documentação do processo de condução do estudo para inspeções futuras.

Estas duas estruturas balizaram a análise realizada neste estudo e são orientadoras do julgamento da qualidade no processo de pesquisa a partir de critérios e indicadores.

2.3 Critérios e Indicadores de Qualidade na Avaliação do Processo de Pesquisa

De acordo com a definição do Comitê Conjunto de Normas para Avaliação Educacional (Yarbrough, Shulha, Hopson, & Caruthers, 2011), um critério é um padrão pelo qual alguma coisa pode ser julgada. Outra definição envolve a dimensão do mérito, isto é, se algo é valioso, ou não, ou se é bom ou ruim (Davidson, 2005). Um terceiro conceito descreve os critérios como um conjunto de padrões que definem a aceitabilidade (Brinberg & McGrath, 1985). No entanto, o julgamento de cada avaliador tende a ser diferente com base no seu propósito. Os avaliadores usam indicadores diferentes no processo de avaliação e, embora a palavra “avaliação” abranja o julgamento com base em perspectivas amplas, ela é essencialmente focada no mérito. Em essência, o indicador é descrito como uma bitola (uma escala) ou uma medida de uma variável (Weiss, 1997). O *framework* para avaliar a qualidade das evidências de pesquisa qualitativa (Mays & Pope, 2006; Spencer *et al.*, 2003) enfatiza que pode ser possível usar “(...) uma série de indicadores de qualidade que apontam para os tipos de informação necessários para avaliar se a característica de qualidade em questão foi alcançada ou não”.

Na literatura não há consenso em relação a quais critérios são os melhores para avaliar a qualidade da pesquisa. Por exemplo, alguns autores argumentam que é impossível ter apenas um conjunto de critérios para fazer isso na pesquisa qualitativa e quantitativa, devido às diferenças envolvidas na conduta desses dois tipos de pesquisa (Whittemore, Chase & Mandle, 2001). Outros autores, entretanto, defendem a possibilidade de utilizar critérios comuns para a pesquisa qualitativa e quantitativa (LeCompte & Goetz, 1982).

A avaliação dos estágios da pesquisa é uma tarefa que envolve vários *stakeholders*. Por um lado, as instituições de fomento, interessadas nos resultados da fase de planejamento (o projeto, em si) que identifica critérios como relevância, impacto e viabilidade para definir o que deve ser financiado. Por outro, o pesquisador e sua equipe, preocupados em observar os atributos de qualidade que permitam conduzir sistemática e rigorosamente o estudo, ou seja, a execução da pesquisa para obter a melhor evidência. Por último, outros interessados centram-se na avaliação da qualidade do “produto de investigação”, ou seja, publicações científicas. A comunidade acadêmica tem usado basicamente duas formas para identificar a qualidade dos periódicos: *surveys* com membros da comunidade científica (referees, editores, pesquisadores, etc.) para identificar suas percepções de qualidade (Ballas & Theoharakis, 2003; Brinn *et al.*, 2001; Lowe & Locke, 2005; Lowensohn & Samelson, 2006; Northcott & Linacre, 2010; Van der Stede, Young, & Chen, 2005) e medidas baseadas em citações dos estudos (a ressonância da produção científica entre os pares) (Aragão *et al.*, 2014; Doyle & Arthurs, 1995).

Como o foco deste trabalho está relacionado ao estudo da prática, utilizou-se uma combinação de diferentes fontes para identificar um conjunto de critérios gerais que foram julgados como possíveis de aplicação, conforme sumariados na Tabela 1. Assim, com base na literatura e nos *frameworks* estudados (Brinberg & McGrath, 1985; Mays & Pope, 2006; Spencer *et al.*, 2003), foi escolhido um conjunto de critérios mais gerais para orientar a análise do processo de pesquisa em Contabilidade.

3. Metodologia

A pesquisa foi realizada compreendendo 4 etapas. A primeira foi uma pesquisa bibliográfica, com o propósito de identificar os atributos de uma boa pesquisa descritos na literatura. Após a definição da lista de critérios gerais e considerando os domínios do ERV e as características-chave apresentadas no *Framework* de Avaliação de Evidências de Pesquisa, foi construída uma matriz para orientar a estruturação dos instrumentos de coleta.

A segunda etapa foi a aplicação da Técnica Delphi, baseada no tipo para formação de consensos. Usou-se a Delphi Modificada, em que as proposições da primeira rodada são elaboradas com base na literatura e não em entradas sugeridas pelos participantes (Kelbaugh, 2003). Foram duas rodadas com docentes dos programas acadêmicos de PPGs da área contábil reconhecidos pela Capes. Ao todo, 318 docentes foram convidados, 41 participaram na primeira rodada e 37 na segunda. Houve representantes de 19 dos 23 PPGs da área contábil reconhecidos pela Capes e ativos no período da coleta de dados.

Com o uso de escala numérica de 10 pontos, o instrumento de coleta teve 53 proposições com atributos/relações relativos à qualidade no processo de pesquisa e à aderência dos mesmos à prática dos respondentes e foi baseado em um estudo da área clínica (Elwyn *et al.*, 2006). A fim de aumentar a validade interna, o instrumento passou por pré-teste com 5 avaliadores, sendo 3 doutorandos e 1 doutor em Contabilidade e 1 doutor em Educação. A coleta de dados ocorreu entre outubro de 2015 e fevereiro de 2016, através da plataforma on-line *SurveyMonkey*, com senha e código de identificação personalizado encaminhados por e-mail-convite aos especialistas que compuseram o painel. O resultado da primeira rodada apresentou um Alfa de *Cronbach* de 92,2% e 28 proposições não tiveram o consenso formado, tendo sido reavaliadas na segunda rodada. Como apenas 5 proposições (menos de 10% do total) não formaram consenso após a segunda rodada e por ter havido baixa variação no desvio padrão destas entre as rodadas, optou-se por dispensar a terceira rodada. A análise de dados considerou medidas de posição e dispersão, além dos critérios para formação do consenso.

Identificados os atributos (Etapa 1) e de posse dos resultados da Delphi com o consenso dos respondentes (Etapa 2), a terceira etapa consistiu na confrontação destas evidências com a classificação por níveis de concordância e identificação de possíveis padrões e/ou variações. Os critérios para formação de consenso foram: (a) 75% ou mais dos respondentes atribuíram notas 1-3 ou 8-10 (forte concordância); (b) 25% ou mais dos respondentes atribuíram notas 1-3 (discordância); e (c) variação do desvio-padrão $\leq 25\%$ entre as rodadas. A classificação das proposições de acordo com o nível de concordância está detalhada na Tabela 2.

Tabela 1

Resumo dos critérios científicos gerais e suas características.

Critério	Significado/características	Fontes
Contribuição/qualidade da perspectiva teórica	A extensão pela qual o público-alvo pode utilizar diretamente os achados; a extensão em que o estudo avança o conhecimento ou traz uma compreensão mais ampliada sobre o objeto. Caracterizada pela identificação de lacunas de conhecimento e pode oferecer soluções para os problemas do público-alvo.	(Mays & Pope, 2006; Meyrick, 2006; Spencer et al., 2003; Webster & Watson, 2002)
Validade externa ou generalização/opportunidade	A aplicação de resultados em outros contextos para generalizar a outras populações. Em uma abordagem qualitativa, a oportunidade é a possibilidade de perceber quão bem as hipóteses poderiam caber em um contexto diferente do que foi gerado. Fator: aplicabilidade.	(Brinberg & McGrath, 1985; Spencer et al., 2003; Valentine, 2009; Yarbrough et al., 2011)
Viabilidade	É “em que medida os recursos e outros fatores permitem que uma avaliação seja conduzida de forma satisfatória” (Yarbrough et al., 2011). Para avaliar se um estudo é viável, considerar o acesso a dados, recursos disponíveis, habilidades da equipe de pesquisa, restrições de tempo e outros, bem como o custo-benefício.	(Davidson, 2005; Rossi, Lipsey, & Freeman, 2004; Weiss, 1997; Yarbrough et al., 2011)
Impacto	Na avaliação de um programa, o impacto pode ser definido como “uma mudança na população-alvo ou condições sociais que tenha sido provocada pelo programa, ou seja, uma mudança que não teria ocorrido se o programa estivesse ausente”. (Rossi et al., 2004). Similarmente, no caso da pesquisa científica, o impacto refere-se às mudanças que ocorrem devido aos resultados do estudo ou foram induzidas por eles.	(Carmona, 2006; Rossi et al., 2004; Webster & Watson, 2002; Weiss, 1997)

Critério	Significado/características	Fontes
Integridade	O rigor das questões de pesquisa, design, conduta e teorização. Revela princípios e regras de conduta ou códigos de prática que pessoas e organizações envolvidas na pesquisa científica devem seguir, (a) honestidade na comunicação; (b) confiabilidade na realização de pesquisas; (c) objetividade; (d) imparcialidade e independência; (e) abertura e acessibilidade; (f) O dever de diligência; (g) equidade em fornecer referências e dar crédito; e (h) responsabilidade para futuros cientistas e pesquisadores (European Science Foundation, 2011). Falha na integridade pode ser evidência de má-conduta na pesquisa.	(Antunes et al., 2011; European Science Foundation, 2011; FAPESP, 2012; OADS, 2012)
Validade interna / credibilidade ou defensibilidade	Uma medida da significância da variável independente em relação à variável dependente, ou quão bem e fielmente representado o fenômeno é. Nos estudos qualitativos, é a credibilidade, ou “quão vívida e fiel é a descrição do fenômeno” (Beck, 1993). Em um sentido geral, a validade é definida como “a extensão em que uma medida realmente mede o que se pretende medir” (Rossi et al., 2004). Fator: valor da verdade.	(Beck, 1993; Brinberg & McGrath, 1985; Libby et al., 2002; Mays & Pope, 2006; Spencer et al., 2003)
Relevância	A capacidade da pesquisa de ajudar um grupo de profissionais a resolver problemas. A relevância da pesquisa depende dos usos potenciais para fins científicos, educacionais ou aplicados. Schwartzman (1988) defende que a avaliação da relevância científica de um estudo depende dos próprios cientistas, notadamente pela revisão de pares; outras formas de relevância requerem a participação de outros setores interessados.	(Mays & Pope, 2006; Reiter & Williams, 2002; Schwartzman, 1988; Weiss, 1997; Yarbrough et al., 2011)
Confiabilidade/auditabilidade	Medida da consistência do instrumento na obtenção de resultados semelhantes ou da capacidade de outro investigador seguir a pista de auditoria. Em outras palavras, é “até que ponto uma medida produz os mesmos resultados quando usada repetidamente para medir a mesma coisa” (Rossi et al., 2004). Fator: consistência.	(Beck, 1993; LeCompte & Goetz, 1982; Mays & Pope, 2006; Rossi et al., 2004; Yarbrough et al., 2011)
Rigor/meticulosidade	Registro completo e fiel. Dados ricos, detalhados e complexos. O rigor envolve boas práticas de coleta, análise e transparência de dados. No entanto, a percepção de rigor é diversa, não-consensual, e pode depender do método de pesquisa adotado pelo pesquisador.	(Denscombe, 2010; Evans et al., 2015; Webster & Watson, 2002; Williams, 2014)
Adequação	Escolha de uma estratégia que provavelmente trará sucesso para alcançar os objetivos da pesquisa e que justifique tal escolha clara e explicitamente. Segundo Denscombe (2010), não é uma questão de definição se é boa ou má – ou correta ou errada –, mas se é útil e apropriada para resolver o problema de pesquisa proposto.	(Denscombe, 2010)

Tabela 2

Critérios para classificar as proposições de acordo com níveis de concordância.

Rodríguez-Mañas et al. (2013)	Modelo ajustado	Nível de concordância
>80% das respostas ≥ 8 ou ≤ 3	>75% das respostas ≥ 8 ou ≤ 3	Forte
70% – 80%	65% – 75%	Moderado
50% – 70%	50% – 65%	Baixo
<50%	<50%	Consenso não formado

Com base nas proposições e no uso do modelo lógico, a quarta etapa versou sobre a elaboração de uma abordagem para avaliar o processo de pesquisa. A expressão *logic model* (modelo lógico, em tradução livre) se refere a uma técnica utilizada para avaliar programas e projetos, tanto com foco no andamento (*process*) como nos resultados (*outcomes*) (Kellog Foundation, 2014; Ladd & Jernigan, 2006). No caso deste estudo, o modelo lógico foi estruturado a partir da adaptação de exemplos estudados durante a disciplina Introdução à Teoria da Avaliação, ministrada pelo Professor Thomas A. Schwandt, na Faculdade de Educação da Universidade de Illinois em Urbana-Champaign (UIUC), Estados Unidos, em setembro de 2014. Foram considerados os atributos/relações identificados na literatura e avaliados pelos *experts* no painel da Delphi. Adicionalmente, foram apresentadas possíveis implicações da inobservância dos critérios nos domínios substantivo, conceitual e metodológico do ERV.

Levando em conta os domínios da investigação científica apresentados no Esquema da Rede de Validade, o desenho desta pesquisa é sumarizado conforme Figura 2.

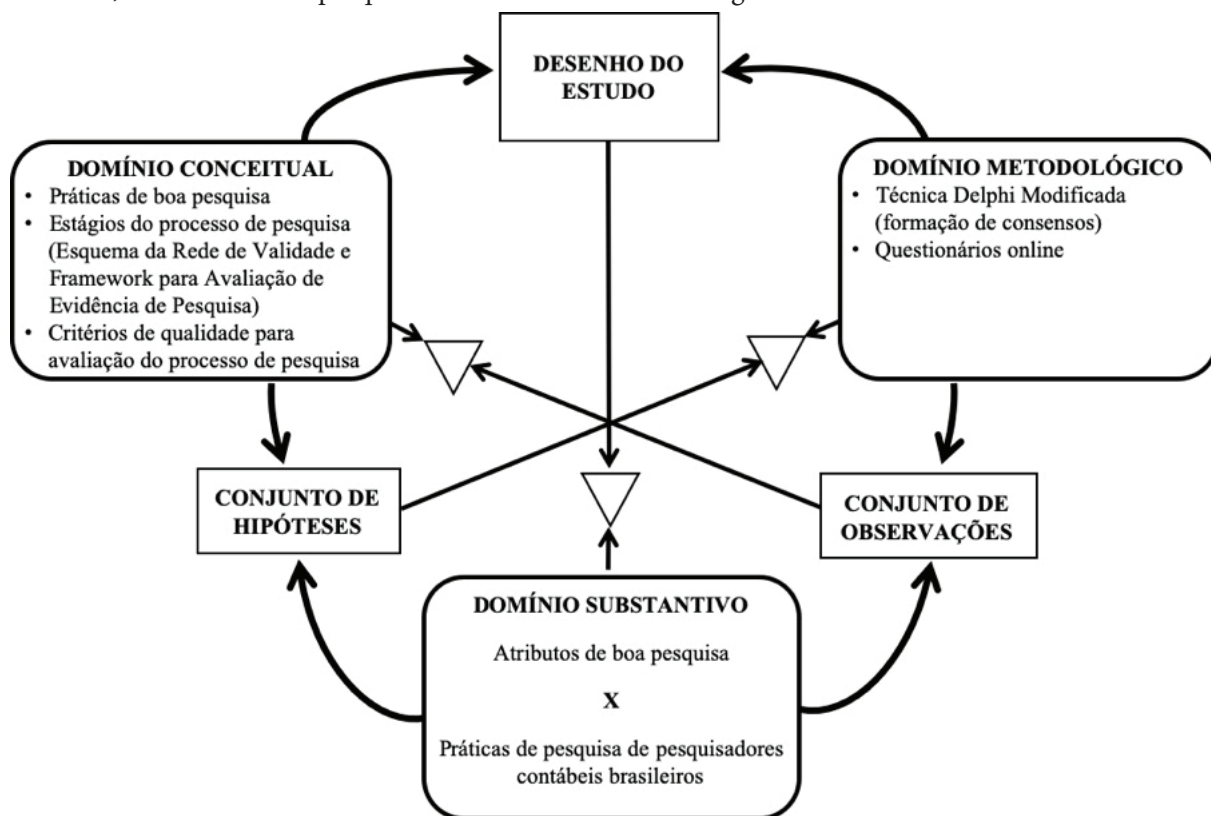


Figura 2. Resumo do desenho da pesquisa com base no ERV (Brinberg & McGrath, 1985).

O estudo foi submetido ao Comitê de Ética em Pesquisa (CEP) e está registrado na Plataforma Brasil.

4. Resultados e Discussão

O painel de especialistas da Delphi contou com 41 membros na primeira rodada e 37 na segunda. A fim de se conhecer o perfil, os respondentes foram questionados sobre suas qualificações profissionais e experiências. A maioria tinha Doutorado em Contabilidade, feito em uma instituição brasileira diferente da que atuava e menos de 15% fizeram doutorado sanduíche ou pós-doutorado em instituição estrangeira, o que sugere uma baixa experiência em atividades internacionais e que pode se refletir em um baixo nível de inserção em redes internacionais de pesquisa. Apenas 9,76% fizeram doutorado no mesmo PPG onde atuavam, uma evidência de reduzida endogeneidade, considerado um elemento de melhor qualidade do Programa pela CAPES. Além disso, PPGs com experiências educacionais diversificadas tendem a ter pesquisadores com habilidades e visões de mundo diferentes, que podem contribuir para o surgimento de estudos mais abrangentes e inovadores.

Sobre a experiência de pesquisa dos respondentes, fator importante para caracterizar a participação como experts no painel, observou-se que 34,14% tinham até 2 anos de experiência em PPG e outros 26,83% tinham entre 5 e 7 anos. Além disso, 56,10% já tiveram/têm pesquisa financiada, cerca de 1/4 possui bolsa de produtividade em pesquisa e apenas 2,44% declararam não ter publicado papers nos estratos A1, A2 ou B1 na última avaliação da CAPES. A maioria participa de Programa com nota 4 e atua em mestrado e doutorado. Estes dados revelam que o grupo possui experiência de pesquisa que o qualifica para tal julgamento.

4.1 Aplicação da Técnica Delphi Modificada

O resultado final da aplicação da Delphi Modificada está detalhado no Apêndice A. Houve variação no número de respostas válidas em função da quantidade de respostas negativas relacionadas ao uso de pesquisas envolvendo seres humanos. Avaliou-se, ainda, que como apenas cinco proposições não atingiram o consenso, não justificava o esforço de realizar outra rodada, dada a baixa variabilidade no desvio-padrão entre as rodadas.

Os achados revelaram que, das 53 proposições, 73,58% (39) receberam o nível de forte de concordância, 11,32% (6) o nível moderado, 5,66% (3) o nível baixo e 9,43% (5) não formaram consenso. O atributo “participação voluntária de sujeitos”, vinculado à integridade, obteve a maior pontuação relativa e 10 itens alcançaram percentual superior a 90%. O item que obteve a maior pontuação absoluta foi “objetivo/problema mostrado com precisão”, que está relacionado ao critério de adequação e obteve 395 pontos dos 410 possíveis.

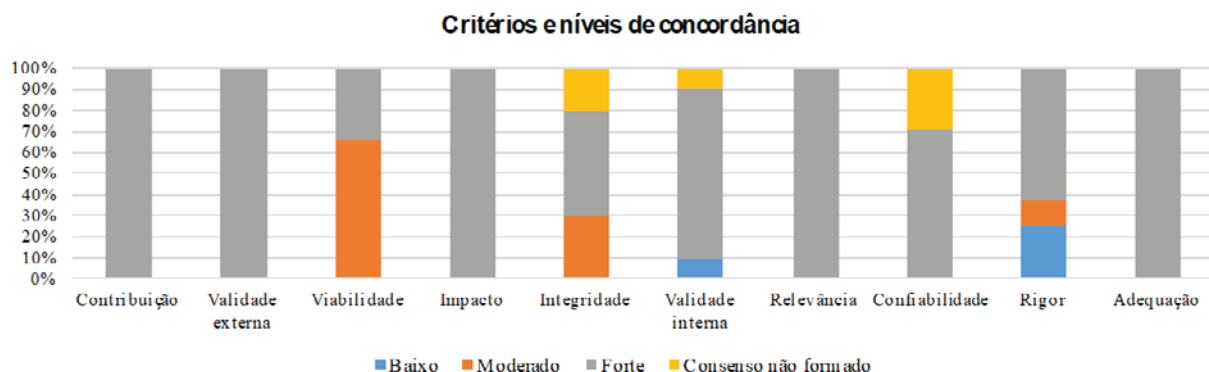


Figura 3. Critérios e níveis de concordância.

A Figura 3 revela que as proposições sobre a contribuição/qualidade da perspectiva teórica, validade externa ou generalização/oportunidade, impacto, relevância e adequação mostraram forte concordância na opinião dos respondentes. Aproximadamente 65% dos itens relacionados ao critério de viabilidade estão em nível moderado. Embora não indiquem, necessariamente, a ocorrência de falhas graves, esses resultados devem ser observados no planejamento dos estudos, pois estão associados a elementos como o dimensionamento do projeto de pesquisa face às restrições de tempo e ao público-alvo na seleção da estratégia de pesquisa. No ERV (Brinberg & McGrath, 1985), o planejamento tem como premissa a validade como valor, é chamado de estágio preparatório e deve ser feito antes dos estágios seguintes. Conseqüentemente, falhas nesta etapa podem comprometer a viabilidade (Rossi *et al.*, 2004; Weiss, 1997; Yarbrough *et al.*, 2011) e conduzir a dados e achados inconsistentes, com inobservância dos critérios de validade interna e externa, como também de impacto.

A integridade apresentou 35% das proposições com nível moderado, o que requer atenção a questões como a forma como o pesquisador lida com erros e vieses e o impacto da participação da equipe de pesquisa sobre os resultados. A integridade dos dados é um elemento central no que se considera prática de boa pesquisa e sua violação pode comprometer a credibilidade científica dos achados (Clarkson, 2012; European Science Foundation, 2011; FAPESP, 2012; OADS, 2012; Shamo & Resnik, 2003).

As proposições com baixo nível foram associadas à validade interna/credibilidade ou defensibilidade e ao rigor, sendo que cerca de 1/4 das desse último obtiveram baixo nível de concordância e devem ser um ponto considerado pelos pesquisadores por sugerir que há fraquezas na condução de estudos na área. Baixos níveis de concordância nesses itens podem indicar uma perda da qualidade da pesquisa, por comprometer a credibilidade dos achados, como também demonstrar pontos de melhoria na formação dos pesquisadores da área, dada a relevância desses critérios apontada na literatura (Evans *et al.*, 2015; Mays & Pope, 2006; OADS, 2012; Spencer *et al.*, 2003; Williams, 2014).

Os itens que não alcançaram o consenso estão associados à integridade, validade interna/credibilidade ou defensibilidade, e confiabilidade/auditabilidade. As notas dessas questões revelaram discrepâncias ou menor aceitação média entre os respondentes, além de maior dispersão que levou à não formação do consenso, mesmo com alto escore relativo. Esses critérios são importantes para a qualidade dos estudos (LeCompte & Goetz, 1982; Libby *et al.*, 2002; Mays & Pope, 2006; Spencer *et al.*, 2003; Yarbrough *et al.*, 2011) e avaliações dissonantes nestes itens podem revelar falta de clareza sobre a extensão dos achados em relação à confiabilidade, a exemplo da necessidade de registrar as razões para mudanças em relação ao planejamento, assim como as razões para as limitações do estudo. Conseqüentemente, pode revelar alguma possível prática de pesquisa que requer modificação, de modo a garantir o atendimento a esses critérios.

Quanto à relação entre características-chave e níveis de concordância, a Figura 4 mostra que os “sujeitos da pesquisa” foram a única com forte nível de aceitação. Por outro lado, reflexividade/neutralidade mostrou nível moderado em todas as proposições. Além disso, “ética” e “reporte” foram aquelas em que as proposições com baixo nível se concentraram. Aproximadamente 25% das proposições relacionadas à ética apresentaram baixo nível de aceitação, e esse achado sugere que os pesquisadores precisam se preocupar com os aspectos legais e podem indicar o não cumprimento das características essenciais da integridade da pesquisa e fraquezas no rigor metodológico (European Science Foundation, 2011; Mays & Pope, 2006; Spencer *et al.*, 2003). Esses níveis de aceitação mais baixos podem sugerir, também, a existência de um possível desalinhamento entre os domínios metodológico, substantivo e teórico propostos pelo ERV (Brinberg & McGrath, 1985).

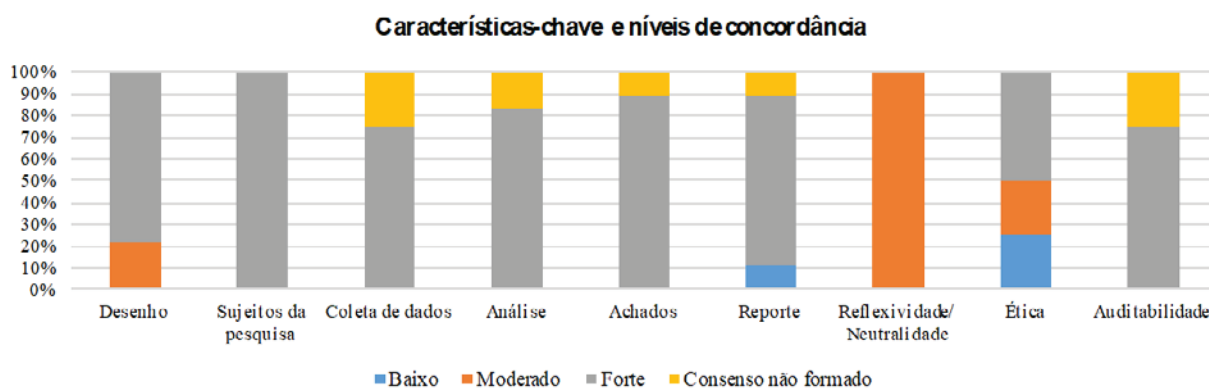


Figura 4. Características-chave e níveis de concordância.

Entre as proposições com baixo nível de aceitação, aquela relativa à apreciação de drafts por pares está associada ao “relatório” e as outras duas estão relacionadas com os aspectos éticos da submissão formal e obrigatória de projetos de pesquisa a um CEP. A avaliação de drafts desempenha um papel significativo na identificação de potenciais problemas e sugere pontos de melhoria, como também ajuda a aumentar a consistência e a coerência lógica de um documento (Mays & Pope, 2006; Spencer *et al.*, 2003). Como resultado, acredita-se que tal avaliação possa aumentar a possibilidade de um manuscrito ser aprovado para publicação porque reduz possíveis falhas identificadas por pares que poderiam levar a uma rejeição na submissão (Carmona, 2006; Martins & Lucena, 2014; Shamoo & Resnik, 2003; Valentine, 2009; Webster & Watson, 2002).

As questões éticas formam a única característica que teve proposições nos três níveis de concordância. Assim como ocorre internacionalmente, a legislação brasileira exige que os estudos envolvendo seres humanos sejam submetidos à avaliação e aprovação prévia por um CEP, a fim de assegurar o atendimento a princípios éticos como o respeito pela dignidade e autonomia de cada participante. O consentimento informado deve descrever o propósito, metas e procedimentos do estudo, esclarecer possível desconforto e riscos de participação e especificar a forma de monitoramento e assistência a que todos os participantes têm direito, mesmo após a conclusão do estudo (Brasil, 2012, 2013). A legislação prevê ainda que, em casos de denúncia e de infração ética que impliquem riscos para os participantes, deve haver um exame que poderá envolver sanções, inclusive a exigência de investigação pelo Ministério Público. Essas práticas são consistentes com princípios internacionais como respeito pelas pessoas, beneficência e justiça (National Institutes of Health, 1979), e de *accountability*, respeito e integridade (European Science Foundation, 2011; OADS, 2012).

O cumprimento de padrões éticos em pesquisas envolvendo seres humanos traz, pelo menos, três efeitos práticos: (a) assegurar que a pesquisa atende aos padrões internacionais de integridade; (b) mitigar riscos ao pesquisador e à instituição relacionados a possíveis ações judiciais de participantes que se sentem prejudicados pela participação nos estudos; e (c) aumentar a validade e a relevância externa do estudo, o que poderá culminar em maior possibilidade de obter financiamento e maior aceitação pela comunidade científica.

4.2 Abordagem para Julgar a Qualidade do Processo de Pesquisa em Contabilidade

Considerando a importância da ligação entre os critérios de qualidade da investigação e as percepções dos respondentes quanto à aderência das proposições às suas práticas, este estudo sugere um conjunto de elementos para avaliar o processo de pesquisa em Contabilidade a partir dos estágios do processo de pesquisa (características-chave), atributos/relações, critérios gerais da qualidade da pesquisa e domínios da investigação.

Com a intenção de traçar a proposta de uma abordagem para avaliar a qualidade no processo de pesquisa em Contabilidade, foi utilizada a estrutura do modelo Lógico (Kellogg Foundation, 2014; Ladd & Jernigan, 2006). Este modelo é amplamente aplicado para avaliar programas e projetos, tanto para processos como para resultados (ou variação). Um modelo Lógico para avaliação de processos é estruturado em insumos, processos/atividades e indicadores/saídas de processo. A abordagem proposta baseada no modelo Lógico considera como insumos as etapas do processo de pesquisa (Figura 5). Inicialmente, esta abordagem considerou os três estágios propostos por Brinberg e McGrath (1985), que neste trabalho são denominados primeiro (planejamento), segundo (execução) e terceiro (monitoramento de resultados ou apenas resultados). Além disso, as nove características-chave (Mays & Pope, 2006; Spencer *et al.*, 2003) foram agrupadas nesses três estágios na seguinte distribuição: (1) desenho e sujeitos da pesquisa; (2) coleta de dados, análise, resultados e reporte; e (3) reflexividade/neutralidade, ética e auditabilidade.

Os atributos e relacionamentos utilizados nos instrumentos de coleta de dados da Delphi foram elaborados a partir da literatura e considerados para caracterizar os processos e atividades propostos no modelo lógico. Considerou-se neste estudo que esses atributos e relacionamentos representam adequadamente os processos, uma vez que a construção do próprio instrumento de coleta se baseou nas práticas de pesquisa dos respondentes, isto é, nas atividades realizadas em diferentes estágios de seus estudos. As proposições foram agrupadas de acordo com a característica-chave e o critério de qualidade a que estão associadas. Com relação aos indicadores de processo/saídas, foram alocados os critérios utilizados para agrupar as proposições, isto é, estão associados aos atributos/relações que são definidos como processos/atividades. Finalmente, foram agregadas algumas implicações potenciais quando esses critérios são violados ou ignorados e, para tanto, utilizou-se a estrutura de domínios do ERV.

Por fim, deve-se enfatizar que esta abordagem (Figura 5) pretende contribuir para a discussão de critérios de qualidade para a construção e realização da pesquisa (o processo) e está menos centrada nos resultados finais (avaliação do produto). A abordagem é apenas uma tentativa de ajudar os pesquisadores da área contábil a auto-avaliarem a qualidade dos seus estudos e não pretende substituir outros *frameworks* para a avaliação da pesquisa.

MODELO LÓGICO PARA AVALIAR O PROCESSO DE PESQUISA						
ENTRADAS	PROCESSOS/ATIVIDADES	INDICADORES DE PROCESSO/SAIDAS				
Estágio	Atributo/relacionamento	Critério	Algumas possíveis implicações			
1º - Planejamento	Desenho <ul style="list-style-type: none"> Estratégia útil ao propósito Visão geral clara Visão geral x estratégia Acesso aos dados Restrições de tempo 	<ul style="list-style-type: none"> Definição do objetivo/problema Visão geral x teoria Argumentos para escolha da técnica Estratégia x público-alvo 	DOMÍNIO SUBSTANTIVO <ul style="list-style-type: none"> Lacuna pouco evidente para justificar a pesquisa. Baixa contribuição para o avanço do conhecimento. Uso de estratégia de pesquisa inadequada. Desperdício de recursos pelo uso inadequado do tempo. Falhas em mostrar os impactos da pesquisa. DOMÍNIO CONCEITUAL <ul style="list-style-type: none"> Falta de clareza dos principais conceitos relacionados ao estudo. Ausência de uma teoria para suportar a análise de dados. Dificuldade para justificar as hipóteses com base na literatura. DOMÍNIO METODOLÓGICO <ul style="list-style-type: none"> Falhas nos procedimentos éticos. Dificuldades para permitir replicação. Limitações para gerar comparação. Contexto da pesquisa descrito inadequadamente. Registros insuficientes dos passos seguidos para realizar a pesquisa. Conclusões não suportadas pelos achados. Risco de dificultar a publicação dos resultados. 			
	Sujeitos da pesquisa <ul style="list-style-type: none"> Critérios p/ desenho/seleção de sujeitos 	<ul style="list-style-type: none"> Representatividade dos sujeitos 		<ul style="list-style-type: none"> Validade interna 		
2º - Execução	Coleta de dados <ul style="list-style-type: none"> Notas de cada passo da pesquisa Participação voluntária dos sujeitos 	<ul style="list-style-type: none"> Anotações sobre eventos divergentes Concordância formal dos sujeitos 		<ul style="list-style-type: none"> Rigor Integridade 		
	Análise <ul style="list-style-type: none"> Descrição da natureza e forma dos dados Descrição das ferram. e procedimentos Contexto x impacto p/ análise dos dados Significância dos dados para alcançar o objetivo 	<ul style="list-style-type: none"> Links impl./explicit.- achados x objetivos Outras visões para conhecer o contexto 		<ul style="list-style-type: none"> Integridade Rigor Confiabilidade Relevância 		
	Achados <ul style="list-style-type: none"> Caminho para chegar às conclusões Link achados x evidências Comparar resultados com outros estudos Novas áreas bascadas nos achados Achados anteriores x hipóteses Contexto para permitir replicação 	<ul style="list-style-type: none"> Checkagem de links achados x propósito Impacto da natureza das divergências 		<ul style="list-style-type: none"> Validade interna Validade externa Contribuição Relevância Confiabilidade 		
		Reporte <ul style="list-style-type: none"> Discussão do impacto p/ o conhecimento Evidenciação das limitações Razões para as limitações Revisão da liter. x principais conceitos Teoria para suportar proposições Possibilidade explícita de generalização 		<ul style="list-style-type: none"> Fronteiras do estudo Conclusões x objetivo Apreciação de drafts por colegas 	<ul style="list-style-type: none"> Impacto Contribuição Integridade Validade interna Validade externa 	
				Reflexividade/Neutral. <ul style="list-style-type: none"> Como lidar com erros e vieses 	<ul style="list-style-type: none"> Impacto da participação da equipe 	<ul style="list-style-type: none"> Integridade
				Ética <ul style="list-style-type: none"> Comitê de Ética para registrar os estudos Regras para registro no CEP Uso de um código de ética Respeito formal pelos seres humanos 	<ul style="list-style-type: none"> Submissão formal ao CEP Submissão compulsória ao CEP Estratégia para mitigar danos Confidencialid. de dados de participantes 	<ul style="list-style-type: none"> Rigor Integridade
3º - Resultados	Auditabilidade <ul style="list-style-type: none"> Registros de mudanças no desenho Registros de razões para mudanças 			<ul style="list-style-type: none"> Proteção da base de dados p/ checkagens Guarda de documentos para reduzir riscos 	<ul style="list-style-type: none"> Confiabilidade 	

Figura 5. Modelo lógico para avaliar o processo de pesquisa.

5. Conclusões

Uma das conclusões deste estudo é que determinadas práticas correntes usadas no processo de construção da produção científica em Contabilidade no Brasil não atendem adequadamente aos atributos de qualidade de uma boa pesquisa descritos na literatura. Isso se evidencia nas elevadas amplitudes entre as respostas obtidas em diversos atributos, bem como na baixa aceitação de elementos relacionados à ética, para os quais a legislação brasileira não admite flexibilidade. Outra conclusão é que falta clareza sobre critérios éticos e a necessidade de cumpri-los nas pesquisas em Contabilidade. Ainda, concluiu-se que o ambiente institucional brasileiro contribui para uma menor qualidade da produção científica em Contabilidade como resultado de falhas institucionais que podem impactar a integridade da pesquisa, suportado em baixos níveis de aceitação de elementos relacionados ao rigor e à ética, médios níveis de aceitação para itens relativos à integridade e viabilidade, além da falha no cumprimento da obrigação de submissão de projetos de pesquisa que envolvem seres humanos a um CEP.

O estudo apresenta limitações, seja em decorrência da técnica Delphi escolhida, a qual possui baixa interação entre os especialistas, seja por possíveis vieses por causa do número de painelistas. Outra limitação é a escolha dos critérios gerais, pois pode haver discrepâncias na avaliação das partes interessadas em relação a eles.

Como contribuições do estudo, deve-se destacar que a avaliação da produção científica, tanto na área como institucionalmente em relação aos PPGs, tem se concentrado sobre os resultados, isto é, sobre o produto final. Contudo, tal análise é limitada para evidenciar aos Programas e pesquisadores da área suas possíveis falhas no processo de pesquisa. O estudo pode impactar o campo, ao ajudar pesquisadores a autoavaliarem suas investigações, identificarem deficiências e limitações, resultando em maior aceitação e encurtamento do processo de submissão aos periódicos. Espera-se que o modelo Lógico para avaliar o processo de pesquisa possa contribuir para encorajar o público-alvo a reconsiderar estratégias de pesquisa e reorientar práticas que não estejam consistentes com a boa pesquisa. Também que os PPGs possam utilizar os resultados para melhorar o processo de formação de novos pesquisadores, a fim de mitigar possíveis falhas.

Face às evidências e ao modelo proposto, sugere-se aos Programas intensificarem a capacitação relativa a questões éticas para diminuir as falhas na condução dos estudos que podem comprometer a integridade e a qualidade da produção, além de inviabilizar a publicação dos resultados. Adicionalmente, a adoção de exigência de submissão de projetos que envolvam seres humanos a um CEP como condição para o exame de qualificação e de aprovação pelo CEP como condição para a defesa. Futuros estudos devem ser empreendidos para compreender as percepções dos alunos de pós-graduação sobre os atributos/relações analisadas, a fim de identificar possíveis erros e falhas na formação de novos pesquisadores na área. Além disso, outro inquérito relevante pode ser discutir o papel dos editores e *referees* como responsáveis pela validação da qualidade, em contraste com a sua percepção da qualidade da pesquisa.

Referências

- Antunes, M. T. P., Mendonça, O. R., Neto, Oyadomari, J. C. T., & Okimura, R. T. (2011). Conduta ética dos pesquisadores em contabilidade: diferenças entre a crença e a práxis. *Revista Contabilidade & Finanças*, 22(57), pp. 319–337. DOI: <https://doi.org/10.1590/S1519-70772011000300006>
- Aragão, I. R. B. N., Oliveira, J. R. S., & Lima, G. A. S. F. de. (2014). Resonance of articles and impact factor of Brazilian Accounting Journals. *Revista de Educação E Pesquisa Em Contabilidade (Journal of Education and Research in Accounting)*, 8(1), pp. 5–21. DOI: <https://doi.org/10.17524/repec.v8i1.1011>
- Ballas, A., & Theoharakis, V. (2003). Exploring diversity in accounting through Faculty journal perceptions. *Contemporary Accounting Research*, 20(4), pp. 619–644. DOI: <https://doi.org/10.1506/MLWH-KBTM-ET47-LYKH>
- Beck, C. T. (1993). Qualitative Research: The Evaluation of Its Credibility, Fittingness, and Auditability. *Western Journal of Nursing Research*, 15(2), pp. 263–266. DOI: <https://doi.org/10.1177/019394599301500212>
- Brasil. Resolução 466, de 12 de dezembro de 2012 (2012). Conselho Nacional de Saúde.
- Brasil. Norma Operacional CNS No 001/2013 (2013). Conselho Nacional de Saúde.
- Brinberg, D., & McGrath, J. E. (1985). *Validity and the research process*. Newbury Park: Sage Publications.
- Brinn, T., Jones, M., & Pendlebury, M. (2001). Why do UK Accounting and Finance Academics not Publish in Top US Journals? *The British Acc. Review*, 33(2), pp.223–232. DOI: <https://doi.org/10.1006/bare.2001.0160>
- Brown, L. D., & Gardner, J. C. (1985). Using Citation Analysis to Assess the Impact of Journals and Articles on Contemporary Accounting Research (CAR). *Journal of Accounting Research*, 23(1), pp. 84–109. <http://doi.org/10.2307/2490908>
- Brown, L. D., & Huefner, R. J. (1994). The Familiarity with and Perceived Quality of Accounting Journals: Views of Senior Accounting Faculty in Leading U.S. MBA Programs. *Contemporary Accounting Research*, 11(1), pp. 223–250. DOI: <https://doi.org/10.1111/j.1911-3846.1994.tb00442.x>
- Carmona, S. (2006). Performance Reviews, the Impact of Accounting Research, and the Role of Publication Forms. *Advances in Accounting*, 22(6), pp. 241–267. <https://www.sciencedirect.com/science/article/abs/pii/S088261100622011X?via%3Dihub>
- Chan, K. C., Chang, C. H., Tong, J. Y., & Zhang, F. (2012). An analysis of the accounting and finance research productivity in Australia and New Zealand in 1991-2010. *Accounting and Finance*, 52(1), pp. 249–265. DOI: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-629X.2011.00440.x>
- Clarkson, P. M. (2012). Publishing: art or science? Reflections from an editorial perspective. *Accounting & Finance*, 52(2), pp. 359–376. DOI: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-629X.2012.00472.x>

- Davidson, E. J. (2005). *Evaluation methodology basics: the nuts and bolts of sound evaluation*. (Sage Publications, Ed.) (1st ed.). Thousand Oaks: Sage Publications.
- Denscombe, M. (2010). *The Good Research Guide*. Open University Press.
- Doyle, J., & Arthurs, A. (1995). Judging the quality of research in business schools: the UK as a case study. *Omega*, 23(3), pp. 257–270. DOI: [https://doi.org/10.1016/0305-0483\(95\)00013-E](https://doi.org/10.1016/0305-0483(95)00013-E)
- Dunbar, A. E., & Weber, D. P. (2014). What influences accounting research? A citations-based analysis. *Issues in Accounting Education*, 29(1), pp. 1–60. DOI: <https://doi.org/10.2308/iace-50603>
- Elwyn, G., O'Connor, A., Stacey, D., Volk, R., Edwards, A., Coulter, A., ... Whelan, T. (2006). Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ (Clinical Research Ed.)*, 333(7565), pp. 417. DOI: <https://doi.org/10.1136/bmj.38926.629329.AE>
- European Science Foundation. (2011). *The European Code of Conduct for Research Integrity*. Strasbourg.
- Evans, J. H., Feng, M., Hoffman, V. B., Moser, D. V., & Van der Stede, W. A. (2015). Points to Consider When Self-Assessing Your Empirical Accounting Research. *Contemporary Accounting Research*, 32(3), pp. 1162–1192. <http://doi.org/10.1111/1911-3846.12133>
- Fundação de Amparo à Pesquisa do Estado de São Paulo. (2012). *Code of good scientific practice*. Recuperado 2012 de http://www.fapesp.br/boaspraticas/FAPESP-Codigo_de_Boas_Praticas_Cientificas_2014.pdf
- Jones, M. J., Brinn, T., & Pendlebury, M. (1996). Judging the quality of research in business schools: A comment from accounting. *Omega*, 24(5), pp. 597–602. DOI: [https://doi.org/10.1016/0305-0483\(96\)00036-9](https://doi.org/10.1016/0305-0483(96)00036-9)
- Kelbaugh, B. M. (2003). *Exploration of teamwork in extension: identifying indicators of success using a modified delphi study (Ph.D. dissertation)*. The Ohio State University. https://etd.ohiolink.edu/!etd.send_file?accession=osu1054604099&disposition=inline
- Kellogg Foundation. (2014). W.K. Kellogg Foundation Logic Model Development Guide. Retrieved September 20, 2014, from <https://www.wkcf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide>.
- Ladd, S., & Jernigan, J. (2006). *Evaluation guide: developing and using a logic model*. Retrieved from https://www.cdc.gov/dhdsp/docs/logic_model.pdf.
- LeCompte, M. D., & Goetz, J. P. (1982). Problems of Reliability and Validity in Ethnographic Research. *Review of Educational Research*, 52(1), pp. 31–60. DOI: <https://doi.org/10.3102/00346543052001031>
- Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, 27(8), pp. 775–810. DOI: [https://doi.org/10.1016/S0361-3682\(01\)00011-3](https://doi.org/10.1016/S0361-3682(01)00011-3)
- Lowe, A., & Locke, J. (2005). Perceptions of journal quality and research paradigm: results of a web-based survey of British accounting academics. *Accounting, Organizations and Society*, 30(1), pp. 81–98. <http://doi.org/10.1016/j.aos.2004.05.002>
- Lowensohn, S., & Samelson, D. P. (2006). An Examination of Faculty Perceptions of Academic Journal Quality within Five Specialized Areas of Accounting Research. *Issues in Accounting Education*, 21(3), pp. 219–239. DOI: <https://doi.org/10.2308/iace.2006.21.3.219>
- Martins, O. S., & Lucena, W. G. L. (2014). Produtivismo acadêmico: as práticas dos docentes dos programas de pós-graduação em contabilidade. *Advances in Scientific and Applied Accounting*, 7(1), pp. 66–96. DOI: <http://doi.org/10.14392/asaa.2014070103>
- Mays, N., & Pope, C. (2006). Quality in qualitative health research. In *Qualitative research in health care* (3rd ed., pp. 82–101). London: Blackwell Pub./BMJ Books.
- Mendonça, O. R., de Neto, Riccio, E. L., & Sakata, M. C. G. (2009). Dez anos de pesquisa contábil no Brasil: análise dos trabalhos apresentados nos Enanpads de 1996 a 2005. *Revista de Administração de Empresas*, 49(1), pp. 62–73. DOI: <http://dx.doi.org/10.1590/S0034-75902009000100008>

- Meyrick, J. (2006). What is good qualitative research? A first step towards a comprehensive approach to judging rigour/quality. *Journal of Health Psychology, 11*(5), pp. 799–808. DOI: 10.1177/1359105306066643
- Michaelis. (2012). *Michaelis Moderno Dicionário da Língua Portuguesa (versão online)*. Editora Melhoramentos.
- Miranda, G. J., Azevedo, R. F. L., & Martins, G. de A. (2011). Teses das teses em contabilidade na USP. *REPEC Revista de Educação e Pesquisa em Contabilidade, 5*(2, mai/ago), pp. 21–42. DOI: <https://doi.org/10.17524/repec.v5i2.342>
- National Institutes of Health. (1979). The Belmont Report. *The Belmont Report Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, pp. 4–6.
- Northcott, D., & Linacre, S. (2010). Producing spaces for academic discourse: The impact of research assessment exercises and journal quality rankings. *Australian Accounting Review, 20*(1), pp. 38–54. DOI: <http://doi.org/10.1111/j.1835-2561.2010.00079.x>
- Office of the Associate Director for Science (OADS). (2012). *CDC Guidance on Scientific Integrity*. Retrieved from http://www.cdc.gov/od/science/docs/CDCSIGuide_021612.pdf
- Reiter, S. A., & Williams, P. F. (2002). The structure and progressivity of accounting research: the crisis in the academy revisited. *Accounting, Organizations and Society, 27*(6), pp. 575–607. DOI: [https://doi.org/10.1016/S0361-3682\(01\)00050-2](https://doi.org/10.1016/S0361-3682(01)00050-2)
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: a systematic approach* (7th ed.). Thousand Oaks: Sage Publications.
- Schwartzman, S. (1988). Qualidade e relevância da pesquisa universitária: 10 teses para discussão. Retrieved from <http://www.schwartzman.org.br/simon/10teses.htm>
- Shamoo, A., & Resnik, D. (2003). *Responsible conduct of research*. New York: Oxford University Press.
- Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). Quality in Qualitative Evaluation: A framework for assessing research evidence A Quality Framework. *Social Research, 108*(1), p.170. DOI: 10.1093/eurpub/ckv219.
- Taylor, J. (2011). The assessment of research quality in UK universities: Peer review or metrics? *British Journal of Management, 22*(2), pp. 202–217. Available at SSRN: <https://ssrn.com/abstract=1841224> or <http://dx.doi.org/10.1111/j.1467-8551.2010.00722.x>
- Valentine, J. C. (2009). Judging the quality of primary research. In H. Cooper, L. V Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (pp. 129–146). Russell Sage Foundation. <https://psycnet.apa.org/record/2009-05060-007>
- Van der Stede, W. a., Young, S. M., & Chen, C. X. (2005). Assessing the quality of evidence in empirical management accounting research: The case of survey studies. *Accounting, Organizations and Society, 30*(7–8), pp. 655–684. <http://doi.org/10.1016/j.aos.2005.01.003>
- Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), pp. xiii–xxiii. <http://doi.org/10.1.1.104.6570>
- Weiss, C. H. (1997). *Evaluation* (2nd ed.). New Jersey: Prentice Hall.
- Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research, 11*(4), pp. 522–537. DOI: 10.1177/104973201129119299
- Williams, P. F. (2014). The Myth of Rigorous Accounting Research. *Accounting Horizons, 28*(4), pp. 869–887. <http://doi.org/10.2308/acch-50880>
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: a guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks: Sage Publications.

Apêndice A

Classificação final da Delphi Modificada por características-chave (estágios no processo de pesquisa), atributos/relações, tamanho número de respostas válidas, rodada em que o consenso foi atingido, escore, percentual relativo do escore, posição no ranking, níveis de aceitação e critérios associados.

Caracter.	Atributo/relação	N	Rod.	Score	%	Rank	Nível	Critério
Col. Dados	Participação voluntária dos sujeitos	23	2	222	96.52	1º	Forte	Integridade
Desenho	Objetivo/probl. mostrado precisamente	41	1	395	96.34	2º	Forte	Adequação
Qst. Éticas	Confidencialid. de dados dos participantes	26	2	248	95.38	3º	Forte	Integridade
Reporte	Conclusões x objetivo	41	1	377	91.95	4º	Forte	Valid. interna
Achados	Comparar resultados com outros estudos	41	1	373	90.98	5º	Forte	Valid. externa
Achados	Checar links achados x propósito	41	1	373	90.98	6º	Forte	Valid. interna
Reporte	Revisão da liter. x conceitos principais	41	1	373	90.98	7º	Forte	Valid. interna
Reporte	Teoria para suportar proposições	41	1	371	90.49	8º	Forte	Valid. interna
Desenho	Estratégia útil ao propósito	41	1	369	90.00	9º	Forte	Adequação
Achados	Achados anteriores x hipóteses	41	1	369	90.00	10	Forte	Relevância
Auditabil.	Salvaguardar base de dados p/ checagens	41	1	367	89.51	11	Forte	Confiabilidade
Desenho	Argumentos para escolha da técnica	41	1	366	89.27	12	Forte	Adequação
Achados	Link achados x evidências	41	1	363	88.54	13	Forte	Valid. interna
Col. Dados	Notas de cada passo da pesquisa	41	1	361	88.05	14	Forte	Rigor
Achados	Novas áreas baseadas nos achados	41	1	361	88.05	15	Forte	Contribuição
Achados	Caminho para chegar às conclusões	41	1	359	87.56	16	Forte	Valid. interna
Desenho	Clara visão geral que guia o estudo	41	1	358	87.32	17	Forte	Adequação
Análise	Descrição de ferram. e procedimentos	41	1	358	87.32	18	Forte	Rigor
Sujeitos	Critérios p/ desenho/seleção de sujeitos	41	1	352	85.85	19	Forte	Valid. interna
Desenho	Acesso aos dados	41	1	351	85.61	20	Forte	Viabilidade
Análise	Descrição da natureza & forma dos dados	41	1	351	85.61	21	Forte	Integridade
Análise	Significância dos dados p/atingir objetivo	41	1	351	85.61	22	Forte	Relevância
Achados	Contexto para permitir replicação	41	1	351	85.61	23	Forte	Confiabilidade
Desenho	Visão geral x estratégia	37	2	316	85.41	24	Forte	Adequação
Reporte	Evidenciação das limitações	37	2	315	85.14	25	Forte	Contribuição
Desenho	Visão geral x teoria	41	1	349	85.12	26	Forte	Adequação
Análise	Links implícit/explicit - achados x objetiv.	41	1	349	85.12	27	Forte	Rigor
Col. Dados	Anotações sobre eventos divergentes	28	2	237	84.64	28	Forte	Rigor
Reporte	Possibilidade explícita de generalização	41	1	345	84.15	29	Forte	Valid. externa
Auditabil.	Registros de mudanças no desenho	29	2	242	83.45	30	Forte	Confiabilidade
Análise	Contexto x impacto p/ análise de dados	41	1	341	83.17	31	Forte	Confiabilidade
Achados	Insights para pensar a área de conheç.	37	2	302	81.62	32	Forte	Contribuição
Qst. Éticas	Estratégia para mitigar possíveis danos	26	2	212	81.54	33	Forte	Integridade
Reporte	Discussão do impacto p/ o conhecimento	37	2	301	81.35	34	Forte	Impacto
Qst. Éticas	CEP para registrar os estudos	26	2	207	79.62	35	Forte	Rigor
Reporte	Fronteiras do estudo	37	2	294	79.46	36	Forte	Contribuição
Qst. Éticas	Respeito formal aos sujeitos humanos	25	2	195	78.00	37	Forte	Integridade
Auditabil.	Guarda de documentos p/ reduzir riscos	25	2	194	77.60	38	Forte	Confiabilidade
Sujeitos	Representatividade dos sujeitos	37	2	287	77.57	39	Forte	Valid. interna
Qst. Éticas	Uso dos princípios de um código de ética	26	2	187	71.92	1º	Moderado	Integridade
Refl./Neutr.	Como lidar com erros e vieses	37	2	266	71.89	2º	Moderado	Integridade

Caracter.	Atributo/relação	N	Rod.	Score	%	Rank	Nível	Critério
Desenho	Estratégia de pesquisa x público-alvo	37	2	264	71.35	3º	Moderado	Viabilidade
Desenho	Restrições de tempo	37	2	261	70.54	4º	Moderado	Viabilidade
Qst. Éticas	Regras para registro no Comitê (CEP)	25	2	176	70.40	5º	Moderado	Rigor
Refl./Neutr.	Impacto da participação da equipe	37	2	249	67.30	6º	Moderado	Integridade
Reporte	Apreciação de drafts por colegas	37	2	229	61.89	1º	Baixo	Valid. interna
Qst. Éticas	Submissão formal ao CEP	25	2	144	57.60	2º	Baixo	Rigor
Qst. Éticas	Submissão compulsória ao CEP	25	2	142	56.80	3º	Baixo	Rigor
Auditabil.	Registros de razões para mudanças	29	2	237	81.72	1º	Ñ formado	Confiabilidade
Col. Dados	Concordância formal dos sujeitos	26	2	212	81.54	2º	Ñ formado	Integridade
Reporte	Razões para as limitações	37	2	293	79.19	3º	Ñ formado	Integridade
Achados	Impacto da natureza das divergências	37	2	289	78.11	4º	Ñ formado	Valid. interna
Análise	Outras visões para conhecer o contexto	37	2	278	75.14	5º	Ñ formado	Confiabilidade