

SIGNIFICÂNCIA SEM SIGNIFICADO: UMA ANÁLISE DA FALTA DE INTERPRETAÇÃO DE RESULTADOS NA ÁREA BRASILEIRA DE NEGÓCIOS

GABRIEL BLANGER

UNIVERSIDADE DO VALE DO ITAJAÍ (UNIVALI)

JOSÉ VÍTOR KOHLER

UNIVERSIDADE DO VALE DO ITAJAÍ (UNIVALI)

LUCIANA DA SILVA IMETON

UNIVERSIDADE DO VALE DO ITAJAÍ (UNIVALI)

JEFERSON LANA

UNIVERSIDADE DO VALE DO ITAJAÍ (UNIVALI)

Agradecimento à órgão de fomento:

Este trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), através do Programa de Suporte à Pós-graduação de Instituições de Ensino Comunitárias (PROSUC), de acordo com a Portaria CAPES nº. 149/2017.

SIGNIFICÂNCIA SEM SIGNIFICADO: UMA ANÁLISE DA FALTA DE INTERPRETAÇÃO DE RESULTADOS NA ÁREA BRASILEIRA DE NEGÓCIOS

1. INTRODUÇÃO

A ciência não mora na significância dos dados, mas sim no significado dos efeitos pesquisados para o mundo real. Esta máxima, que deveria guiar toda pesquisa empírica, parece ter sido substituída por análises menos profundas e algum foco inadvertido por valores de significância estatística (Meneghini & Lana, 2024). Na área de administração e negócios, particularmente no contexto brasileiro, tornou-se lugar comum encontrar estudos que se limitam a relatar se os resultados são "estatisticamente significativos", como se essa fosse a única métrica relevante para o avanço do conhecimento. Essa superficialidade nas análises de efeitos dificulta a acumulação de conhecimento e afasta o mundo acadêmico do mundo corporativo de forma geral. Este artigo surge como resposta crítica a essa tendência reducionista, propondo uma mudança de paradigma que coloque o tamanho do efeito e sua interpretação prática no centro da análise quantitativa.

O problema vai além da mera necessidade técnica de reportar medidas de magnitude. Trata-se de uma questão epistemológica fundamental: ao avaliar exclusivamente a significância estatística, uma pesquisa acaba por negligenciar as implicações práticas dos achados empíricos e reduz drasticamente seu potencial de impacto. Como bem destacam Cohen (1994) e Coe (2002), a fixação pelo p-valor representa um obstáculo ao desenvolvimento científico, pois confunde a existência de um efeito (questão estatística) com sua relevância substantiva (questão teórica e prática). Essa confusão é particularmente danosa na administração, campo em que decisões estratégicas demandam não apenas saber se uma intervenção "funciona", mas compreender em que medida e em quais contextos ela produz resultados mensuráveis.

Motivados pela constatação da lacuna na literatura nacional quanto à interpretação da magnitude dos efeitos em pesquisas quantitativas, o presente estudo investigou sistematicamente a adoção e a qualidade dessa prática nos periódicos de maior impacto na área de administração no Brasil. Para tanto, realizamos uma análise de conteúdo de 347 artigos quantitativos publicados em periódicos brasileiros classificados como Qualis-Capes A2 entre os anos de 2022 e 2023. A seleção desse estrato de publicações visou capturar as práticas de pesquisa consideradas de maior rigor e relevância no cenário nacional. A análise centrou-se na identificação da presença de medidas de tamanho do efeito, na forma como foram reportadas e, crucialmente, na profundidade e na contextualização de suas interpretações em relação ao mundo real dos negócios. Os resultados dessa análise revelaram que essa prática ainda é significativamente limitada no cenário nacional: apenas 7% dos estudos analisados articularam adequadamente o cálculo do tamanho do efeito com sua interpretação no contexto empresarial.

Neste artigo, buscamos não apenas diagnosticar essa lacuna na pesquisa brasileira, mas apresentar os riscos dessa prática e oferecer caminhos concretos para sua superação. Nosso objetivo triplo consiste em: (1) apresentar um panorama atualizado das práticas de reporte e interpretação do tamanho do efeito nos periódicos nacionais de administração; (2) demonstrar, por meio de exemplos concretos, como a interpretação contextualizada pode ampliar o impacto e a aplicabilidade dos achados de pesquisa; e (3) propor diretrizes metodológicas para elevar o padrão das publicações brasileiras nessa área.

A relevância deste debate transcende o âmbito acadêmico. Em um contexto organizacional cada vez mais orientado por dados, a capacidade de traduzir resultados estatísticos em insights acionáveis torna-se competência estratégica. Como mostraremos ao longo do artigo,

pesquisas que integram significância estatística e significância prática não apenas produzem conhecimento mais robusto, como também fortalecem a ponte entre academia e mundo corporativo - condição essencial para o desenvolvimento sustentável do campo da administração no Brasil.

2. REFERENCIAL TEÓRICO

A interpretação do tamanho do efeito é um elemento relevante na análise estatística de pesquisas quantitativas, já que complementa a significância do p-valor ao avaliar a magnitude prática dos resultados. Desenvolvido inicialmente por Cohen (1988) e posteriormente ampliado por outros pesquisadores (Lakens, 2013; Cumming, 2014; Fey, et. al. 2023), esse constructo permite quantificar diferenças ou relações entre variáveis, auxiliando na compreensão do impacto real de intervenções ou fenômenos estudados. Nesta seção, são apresentadas as principais métricas, os critérios para sua interpretação e os desafios associados à sua aplicação. Discute-se ainda a necessidade de relatar o tamanho do efeito em vez de depender exclusivamente de testes de hipóteses com base em recomendações recentes que valorizam a clareza metodológica e a robustez das conclusões científicas.

2.1 A Relevância do Tamanho do Efeito na Pesquisa em Gestão

A pesquisa quantitativa na área de gestão frequentemente utiliza a estatística inferencial para generalizar conclusões de amostras para populações maiores. Nesse contexto, a aplicação e interpretação de testes estatísticos são cruciais para validar hipóteses e construir conhecimento científico. A comunicação eficaz dos resultados estatísticos é essencial não só para a comunidade acadêmica, mas também para fornecer informações valiosas para a tomada de decisões em organizações e políticas.

Tradicionalmente, o p-valor tem sido usado como critério padrão para determinar a significância estatística de um resultado. O p-valor indica a probabilidade de obter resultados tão extremos quanto os observados, assumindo que a hipótese nula é verdadeira, e um limiar de 0,05 é comumente usado para declarar significância (Sun et al., 2010). Contudo, é importante ressaltar que essa abordagem, quando aplicada isoladamente, apresenta limitações (Wasserstein & Lazar, 2016).

No entanto, a ênfase excessiva no p-valor como único indicador de relevância tem sido criticada. Autores como Wasserstein e Lazar (2016) argumentam que o p-valor sozinho não informa sobre a magnitude ou a importância prática de um efeito. A busca por valores de $p < 0,05$ pode levar a interpretações errôneas e à negligência de outros aspectos importantes da análise de dados, como o tamanho do efeito (Sullivan & Feinn, 2012).

Em contrapartida, o tamanho do efeito surge como uma métrica essencial para interpretar os resultados da pesquisa. Enquanto o p-valor indica a probabilidade de um resultado ocorrer por acaso, o tamanho do efeito quantifica a magnitude da relação entre variáveis ou a diferença entre grupos, proporcionando uma medida padronizada que facilita a comparação entre estudos (Ellis, 2010; Lakens, 2013). Métricas como o R^2 , o d de Cohen e o eta quadrado (η^2) são usadas para expressar a magnitude desses efeitos (Fritz et al., 2012).

No âmbito das diretrizes para a apresentação de resultados de pesquisa, a American Psychological Association (APA) tem progressivamente enfatizado a importância da interpretação dos achados para além da mera significância estatística. O Publication Manual da APA (7ª ed., 2020), por exemplo, dedica seções à discussão da magnitude do efeito, intervalos de

confiança e à necessidade de os autores oferecerem uma interpretação dos resultados que considere suas implicações práticas e potencial para replicação em contextos do mundo real (APA, 2020). A norma encoraja os pesquisadores a discutirem a relevância clínica ou prática de seus achados, conectando os resultados estatísticos com o significado substantivo para a área de estudo e para potenciais aplicações (Kline, 2004). Em contraste, as normas brasileiras da Associação Brasileira de Normas Técnicas (ABNT), embora detalhem extensivamente os aspectos formais da apresentação de trabalhos acadêmicos, como citações (NBR 10520, 2023) e formatação (NBR 14724, 2011), não apresentam uma orientação tão explícita e detalhada sobre a interpretação dos resultados com foco em sua aplicação prática e replicabilidade. Essa lacuna nas normas nacionais pode contribuir para a observada prevalência da ênfase na significância estatística em detrimento da avaliação da relevância prática dos achados em pesquisas brasileiras na área de negócios.

A relação entre p-valor e tamanho do efeito é fundamental para uma compreensão completa dos resultados. Um p-valor significativo indica que o efeito provavelmente não ocorreu por acaso, mas não informa sobre sua intensidade. Um tamanho do efeito grande pode ter um p-valor não significativo em amostras pequenas. Portanto, analisar ambos proporciona uma avaliação mais robusta (Sullivan & Feinn, 2012).

Na pesquisa em gestão, é fundamental ir além da significância estatística e considerar a magnitude dos efeitos observados. Cohen (1990) e Sullivan & Feinn (2012) destacam que o principal resultado de uma pesquisa quantitativa deve ser a medida da importância prática dos efeitos, e não apenas o p-valor. Compreender o tamanho do efeito permite avaliar a relevância dos resultados para a gestão e para o desenvolvimento de políticas corporativas. Fey, Hu e Delios (2023) argumentam que um efeito estatisticamente significativo pode ter pouca utilidade prática se o tamanho do efeito for pequeno. Assim, a análise crítica dos resultados deve combinar significância estatística e magnitude dos efeitos, usando métricas apropriadas como d de Cohen ou R^2 , dependendo da análise (Fritz et al., 2012; Lakens, 2013).

A interpretação do tamanho do efeito é essencial para a aplicabilidade prática dos resultados, especialmente em organizações. Efeitos pequenos podem ser relevantes em grande escala, como argumentam Funder e Ozer (2019), e podem ter implicações significativas para a gestão. Aguinis et al. (2023) também enfatizam que até mesmo efeitos estatisticamente pequenos ($d = 0,2$) podem ter implicações estratégicas significativas quando aplicados em larga escala ou em contextos de alto valor agregado.

Além disso, o cálculo do tamanho do efeito é importante para planejar estudos e calcular o poder estatístico, facilitando meta-análises que sintetizam os resultados e permitem a construção cumulativa do conhecimento (Lakens, 2013). No entanto, como observam Fey, Hu e Delios (2023), a apresentação do tamanho do efeito deve incluir uma interpretação crítica de sua relevância prática, especialmente em contextos empresariais, onde pequenos efeitos podem ter um impacto estratégico significativo.

No Brasil, a pesquisa em gestão enfrenta desafios na aplicação dessas práticas. Análises de artigos brasileiros mostram que muitos estudos se concentram na significância estatística, negligenciando o tamanho do efeito e sua interpretação prática. Isso pode levar a uma desconexão entre a pesquisa e a prática gerencial, limitando a aplicação dos resultados acadêmicos. Portanto, é crucial que os pesquisadores brasileiros em gestão comuniquem o tamanho do efeito de forma eficaz, contextualizando seus resultados e destacando as implicações práticas para as organizações. Essa necessidade é ressaltada por diversos autores que defendem uma ciência mais relevante e útil para a sociedade (Wasserstein et al., 2019).

3. ASPECTOS METODOLÓGICOS

A análise de dados foi realizada por meio de uma análise de conteúdo. O objetivo era identificar características dos trabalhos e como os autores descrevem e interpretam o tamanho do efeito em pesquisas quantitativas no Brasil. Essa análise seguiu um processo estruturado para garantir precisão. Este capítulo tem como objetivo detalhar a operacionalização da pesquisa.

3.1 Delineamento e caracterização da pesquisa

O presente estudo se caracteriza por uma abordagem quantitativa, com fins descritivos, buscando analisar de que forma os autores brasileiros descrevem o tamanho do efeito nos resultados de pesquisas quantitativas e em que medida essa descrição considera seu impacto no mundo real. Para atingir este objetivo, foi conduzida uma análise de conteúdo. A análise de conteúdo é uma metodologia frequentemente utilizada em estudos qualitativos e quantitativos que ganhou popularidade entre os acadêmicos que buscam identificar e resumir tendências na literatura (Gaur & Kumar, 2018; Kleinheksel et al., 2020).

Conduzir uma análise de conteúdo abrangente em ciências sociais envolve uma abordagem sistemática que engloba várias etapas principais. Esse método é essencial para analisar várias formas de comunicação, incluindo texto e mídia, e pode ser adaptado para projetos de pesquisa qualitativa e quantitativa. As seções a seguir descrevem as etapas críticas envolvidas nesse processo. As principais etapas da análise de conteúdo consistem em: identificar o problema, selecionar unidades de análise (Mustapha & Ebomoyi, 2019), familiarização de dados (Fuchs, 2023), codificação (Bass & Semetko, 2021), desenvolvimento de temas e relatório final (Fuchs, 2023). Essas etapas não apenas garantem uma compreensão profunda do material analisado, mas também promovem a rigorosidade e a validade dos resultados obtidos durante a pesquisa.

Embora a análise de conteúdo seja um método robusto para extrair insights da comunicação, é importante reconhecer suas limitações, como possíveis vieses na codificação e os desafios de garantir a confiabilidade entre codificadores humanos (Mustapha & Ebomoyi, 2019), (Bass & Semetko, 2021).

Dessa forma, essa pesquisa se propõe a ampliar a discussão sobre as melhores práticas para mitigar esses vieses, propondo um protocolo de codificação mais rigoroso e métodos de validação cruzada que possam aumentar a confiabilidade dos resultados. O próximo capítulo busca detalhar como foi implementado esse protocolo, apresentando os passos seguidos para garantir a consistência e a precisão na análise dos dados coletados.

3.2 Análise de conteúdo

Inicialmente, cada artigo foi classificado de acordo com os seguintes critérios: origem dos dados, geografia dos dados, período dos dados, técnicas de análise de dados, apresentou p-valor ou não, se houve significância e, finalmente, se falou do tamanho do efeito e como foi feita a interpretação no mundo real. Para isso foi utilizada uma planilha compartilhada entre os autores com as opções em formato de lista suspensa para classificar os artigos de acordo com estes critérios.

Após a identificação do problema por parte dos pesquisadores por meio de leituras relevantes na área de ciências sociais, investigando práticas de reporte do tamanho do efeito nos jornais acadêmicos (Sun et al., 2010), o capítulo se concentra em descrever as metodologias

utilizadas para a análise de conteúdo, incluindo critérios de inclusão e exclusão, bem como os instrumentos aplicados para coletar e interpretar os dados de maneira sistemática.

3.2.1 Seleção da unidade de análise

Para a realização da coleta de dados, foram conduzidos levantamentos manuais abrangendo todos os artigos publicados em revistas brasileiras que possuem classificação Qualis A2, correspondendo aos anos de 2022 e 2023. Inicialmente, identificou-se um total de 773 artigos.

Essa diversidade de tipos de pesquisa reflete a riqueza e a complexidade do conhecimento produzido nas áreas abordadas. Na mesma fase do processo, os trabalhos que apresentam uma abordagem mista foram classificados como quantitativos, com a intenção de facilitar uma análise mais aprofundada da metodologia utilizada posteriormente. Essa classificação criteriosa é fundamental para garantir a integridade da análise e a relevância dos dados coletados, permitindo uma compreensão mais abrangente das contribuições científicas presentes nos artigos selecionados.

3.2.2 Familiarização de dados

Após o download de todos os artigos das revistas citadas na íntegra, os autores leram todos os artigos de forma abrangente incluindo título, resumo e metodologia com a finalidade de identificar e fazer uma triagem de artigos quantitativos e não quantitativos. Os artigos não quantitativos incluem uma variedade de formatos, como pesquisas bibliográficas, revisões sistemáticas, estudos de caso voltados para o ensino, ensaios teóricos e investigações qualitativas.

Essa diversidade de tipos de pesquisa reflete a riqueza e a complexidade do conhecimento produzido nas áreas abordadas. Na mesma fase do processo, os trabalhos que apresentaram uma abordagem mista foram classificados como quantitativos, com a intenção de facilitar uma análise mais aprofundada da metodologia utilizada posteriormente. Essa classificação criteriosa é fundamental para garantir a integridade da análise e a relevância dos dados coletados, permitindo uma compreensão mais abrangente das contribuições científicas presentes nos artigos selecionados. A Tabela 1 elenca todas as revistas quantificando os artigos encontrados.

Tabela 1

Levantamento inicial

Revista	Não quantitativos	Quantitativos	Mistos	Total
BAR	17	30	2	49
RCF	10	58	1	69
REV RAE	52	35	3	90
RBGN	12	56	1	69
REGE	28	20	3	51
REV BBR BRAZILIAN BUSINESS REVIEW	20	50	0	70

REV EBAPE	106	4	2	112
REV RAC	68	18	3	89
REV RAM MACKENZIE	40	33	0	73
REV RAUSP	22	23	1	46
ROS	51	3	1	55
TOTAL	426	330	17	773

Fonte: Dados da pesquisa (2025).

3.2.3 Codificação e desenvolvimento dos temas

Após a triagem, foram selecionados apenas os artigos com abordagem quantitativa e misto da amostra para categorizá-los de acordo com os seguintes critérios: origem dos dados, geografia dos dados, período dos dados, técnicas de análise de dados, apresentou p-valor e significância, falou do tamanho do efeito e fez interpretação no mundo real. As opções de cada categoria estão detalhadas no Apêndice A. Embora existam várias categorias para classificação, o principal foco nesta etapa foi a leitura minuciosa dos resultados com a finalidade de identificar se os autores deixaram claro uma relação do tamanho do efeito com o mundo real.

Por se tratar de uma pesquisa quantitativa, apenas foi feita uma classificação de acordo com os critérios bem estabelecidos. A análise descritiva dos dados coletados foi realizada utilizando Microsoft Excel. Os resultados foram apresentados em forma de gráficos e tabelas para facilitar a interpretação. A discussão dos achados foi baseada na comparação com estudos similares encontrados na literatura científica. Os resultados dessa triagem serão descritos no capítulo 4 desta pesquisa.

3.2.4 Relatório final

A identificação da forma e descrição do tamanho do efeito de acordo com as métricas estatísticas mais comuns como d de Cohen (Cohen, 1988), r de Pearson, coeficiente de determinação (R^2) e eta quadrado (η^2) (Lazaraton, 2005).

Em primeiro momento os artigos foram classificados em:

- a) Apresentaram o tamanho do efeito ou;
- b) Não apresentaram o tamanho de efeito.

Para os artigos que apresentaram o tamanho do efeito, foi analisado como foi realizado apresentado, buscou identificar se o autor fez a interpretação apresentando como esse efeito, embora muitas vezes pequeno, tem real impacto na gestão das organizações, quer seja nos lucros, redução dos custos, retenção de clientes, entre outros. A classificação segundo este critério inclui:

- a) Fez interpretação no mundo real ou;
- b) Não fez interpretação no mundo real.

Em suma, a análise do relatório final dos artigos selecionados concentrou-se em identificar a presença do tamanho do efeito e, sobretudo, em verificar se os autores realizaram uma interpretação contextualizada desse efeito, vinculando-o a implicações práticas relevantes para a gestão das organizações. Essa abordagem metodológica detalhada permitiu categorizar as

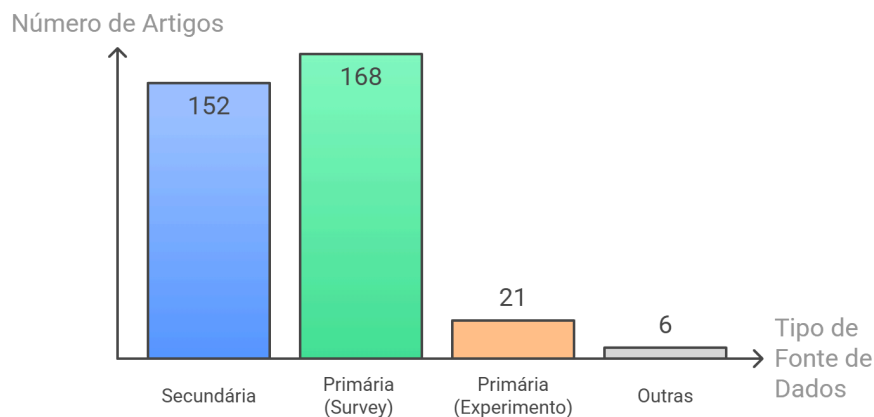
publicações de acordo com o rigor na apresentação e interpretação da magnitude dos resultados, fornecendo a base para as conclusões e as recomendações apresentadas neste estudo.

4. RESULTADOS

A análise dos artigos permitiu traçar um panorama da produção científica brasileira no que se refere ao relato de testes estatísticos. Especificamente, foram examinadas publicações em periódicos com classificação Qualis-Capes A2, majoritariamente de autoria de pesquisadores brasileiros, a fim de identificar como os achados de seus estudos vêm sendo reportados. Para isso, os autores deste estudo preencheram planilhas compartilhadas com as informações extraídas de cada artigo analisado. Os resultados obtidos estão detalhados a seguir.

No quesito origem dos dados do artigo, 152 utilizaram de fonte secundária, 168 fonte primária (por exemplo, *survey*), e 21 primária a partir de experimento e 6 classificados como outros, que podem ser consultados no Apêndice A, conforme Figura 1.

Figura 1
Distribuição das fontes de Dados em Artigos

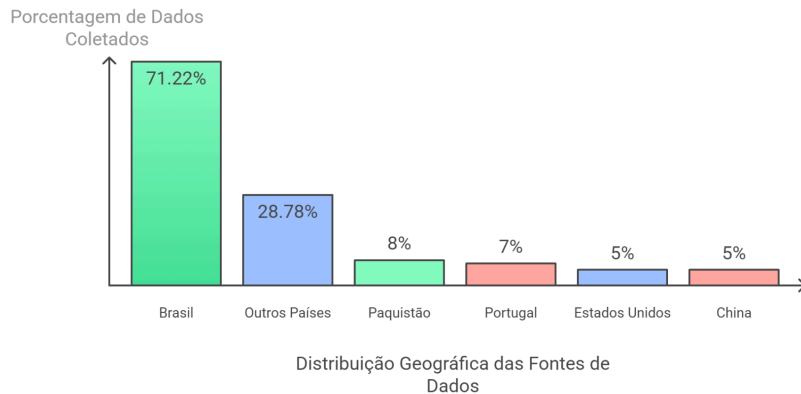


Distribuição de Fontes de Dados em Artigos

Fonte: Dados da pesquisa (2025).

A geografia dos dados indica que do Brasil foi de onde 71,22% dos dados foram coletados e que os 28,78% dos demais locais foram de variados países, destacando-se o Paquistão (8 artigos), seguido de Portugal (7 artigos) e, Estados Unidos e China (5 artigos cada). Os demais 26 locais identificados foram base para 3 artigos ou menos, conforme Figura 2.

Figura 2
Porcentagem dos dados coletados

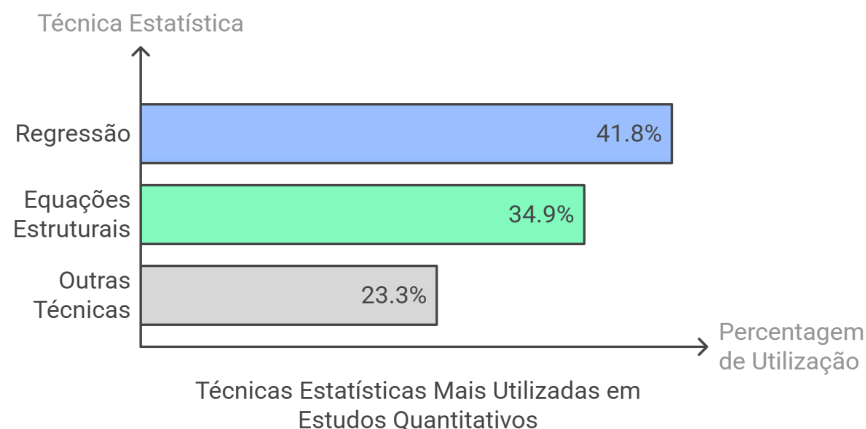


Fonte: Dados da Pesquisa (2025)

Quanto ao período dos dados, observou-se uma predominância de estudos de corte transversal (*cross-sectional*), que representam 65,12% da amostra, ou seja, baseiam-se em dados coletados em um único momento no tempo (Bryman, 2016). Já os estudos longitudinais e em painel, que analisam múltiplas amostras ao longo do tempo (Hair, 2019), corresponderam a 34,88% do total.

Dos artigos quantitativos, as duas técnicas mais utilizadas foram: Regressão com 41,8% e equações estruturais com 34,9% o que corrobora com os estudos de (Brei & Liberali Neto, 2006; Prearo & Gouvêa, 2011; Bido et al., 2012). As demais técnicas todas somadas representam 23,3%, demonstrando a preponderância das duas principais técnicas na publicação nacional como reportado na Figura 3.

Figura 3
Técnicas Estatísticas



Fonte: Dados da pesquisa (2025).

No que se refere ao p-valor, dos 347 artigos analisados 310 o apresentaram, 23 não se aplicavam à sua utilização e, em 14, ele sequer foi mencionado. Cabe ressaltar que esse

quantitativo considerou apenas a presença ou ausência do p-valor nos estudos, sem avaliar os valores específicos reportados. Assim, pode-se concluir que a maioria dos artigos possibilitou a verificação da significância estatística de seus resultados. Em se tratando de significância estatística, 313 artigos apresentaram - mesmo que nem sempre com a especificação exata do p-valor - tanto é que um número maior de artigos descreveu resultado com significância que o total dos que apresentaram p-valor -, o que inclusive supera o número de artigos que efetivamente apresentaram o valor de p. Em apenas oito artigos constatou-se ausência total de significância estatística (nenhum $p < 0,05$), enquanto, em outros 28, essa análise não se aplicava.

Por outro lado, o tamanho do efeito foi reportado em apenas 182 artigos, enquanto 139 não o apresentaram e, em 26 casos, sua demonstração não se aplicava. Esse achado sugere que, apesar da ampla utilização do p-valor, a preocupação com a magnitude dos efeitos ainda não é uma prática consolidada, tanto na produção científica nacional quanto na produção internacional (ver Cumming, 2014). E, ainda, a apresentação dos resultados geralmente finaliza ao constatar a significância ou não significância. É como se o P-valor inferior a 0,05 tenha se tornado o objetivo e não o modo pelo qual os pesquisadores consigam demonstrar que a hipótese foi aceita e que, com os devidos testes, o tamanho do efeito pode ser mensurado (Bakker et al., 2019).

A explicação do tamanho do efeito, ou seja, quanto a variável X influencia na variável Y no mundo real (Ialongo, 2016), não é informada em quase nenhum dos artigos analisados. Ainda que em 182 artigos (52,13%) seja informado o tamanho do efeito (interpretação do beta; r^2 ou “d” de Cohen ou f^2) a interpretação é feita em apenas 24 artigos, somente 7,27% da amostra dos artigos quantitativos que apresentaram p-valor. Os dados são apresentados geralmente em tabelas com siglas e pouco se desenvolve a partir deles. Ao leitor conhecedor de estatística e pesquisa quantitativa a interpretação pode ser sem dificuldade, mas aos leitores iniciantes na pesquisa ou gestores de empresas esse dado não é facilmente interpretado e, na maioria das vezes, é visto sem ser percebido.

Faz-se necessário esclarecer, no entanto, que, embora em muitas áreas o tamanho do efeito precise ser grande para ser considerado relevante — como ocorre em diversos estudos clínicos na medicina (Kim et al., 2024; Burrows et al., 2023), nos quais se avalia a eficácia de medicamentos, vacinas ou outras substâncias —, na área de gestão, os efeitos dificilmente são expressivos, uma vez que os fenômenos analisados são multidimensionais. Todavia, mesmo sem apresentar grandes tamanhos de efeito, as pesquisas em gestão possuem grande relevância para as empresas, pois, ainda que uma variável sob seu controle altere a lucratividade em apenas 3%, esse impacto pode representar milhões em ganhos financeiros.

4.1 O uso excessivo do p-valor e a ilusão da significância

Esses achados reforçam críticas amplamente reconhecidas na literatura internacional sobre o uso indiscriminado do p-valor como critério principal de validação estatística. Como apontam (Wasserstein & Lazar, 2016), o p-valor, isoladamente, não fornece informações sobre a magnitude, importância ou direção de um efeito. Trata-se apenas da probabilidade de ocorrência de um resultado extremo, assumindo a hipótese nula como verdadeira. Entretanto, muitos estudos tratam o p-valor como se fosse um marcador absoluto de “verdade científica”, o que reduz a riqueza interpretativa dos dados.

Essa tendência foi chamada por (Cohen, 1994) de “cult of statistical significance”, referindo-se à busca obsessiva por valores inferiores a 0,05 como se fossem comprovações definitivas de validade científica. O presente levantamento ilustra bem essa prática, ao mostrar

que mais artigos relataram resultados significativos do que efetivamente apresentaram o valor do p , indicando uma padronização ritualística da significância. Porém, outras áreas, como por exemplo de medicina e psicologia, também vêm percebendo os erros de interpretação e reporte de p -valor e efeito (Lipsey, 2012; Shadish et al., 2015; Baird & Pane, 2019).

4.2. A negligência com o tamanho do efeito

A baixa frequência de apresentação — e especialmente de interpretação — do tamanho do efeito indica uma falha metodológica e comunicacional relevante. Ainda que muitos estudos incluam métricas como R^2 , d de Cohen ou η^2 , raramente esses valores são discutidos de maneira substantiva. Como argumentam (Sullivan & Feinn, 2012), o tamanho do efeito é a medida que realmente interessa, pois revela o impacto real de uma variável sobre outra. Um efeito pode ser estatisticamente significativo, mas praticamente irrelevante — e o contrário também é verdadeiro.

No campo da gestão, por exemplo, é improvável encontrar efeitos com grande magnitude, dado o caráter multifatorial dos fenômenos organizacionais (Herrmann & Nadkarni, 2014; Klein et al., 2024; Naranjo-Gil & Hartmann, 2007; Qiang et al., 2023; Zhang & Rajagopalan, 2010). No entanto, mesmo efeitos modestos — como uma variação de 2% na lucratividade explicada por uma determinada variável — podem representar milhões de reais em ganhos financeiros. A relevância, portanto, deve ser contextualizada, como já destacava Cohen, (1994) em suas recomendações sobre interpretação de efeitos.

4.3. Implicações para a prática e a ciência brasileira

O foco excessivo no p -valor em detrimento da análise de efeito tem consequências drásticas para a aplicação do conhecimento científico. A ausência de explicações sobre a magnitude dos resultados dificulta a compreensão por parte de gestores, formuladores de políticas públicas e demais públicos que se beneficiariam de interpretações claras e contextualizadas. Como defendem Funder & Ozer, (2019), é responsabilidade dos pesquisadores não apenas relatar resultados estatisticamente válidos, mas também torná-los compreensíveis e úteis para a sociedade.

Além disso, essa prática contribui para a chamada crise de replicação, como demonstrado pela Open Science Collaboration (2015), observada em diversas áreas do conhecimento. Resultados reportados com base apenas na significância estatística, sem avaliação da robustez e da importância prática dos efeitos, tendem a não se replicar em estudos posteriores enfraquecendo a credibilidade da ciência empírica.

Dessa forma, os dados analisados evidenciam a necessidade urgente de mudança na cultura de publicação científica brasileira, incentivando a apresentação e interpretação do tamanho do efeito como prática padrão. Esse movimento já está em curso em outras áreas, como a psicologia internacional, e pode trazer ganhos substanciais para a qualidade e utilidade da produção científica nacional.

5. CONCLUSÃO

Este estudo desvela uma dissonância significativa na pesquisa brasileira em administração. Em contraste com o consenso estabelecido na literatura internacional acerca da centralidade da interpretação contextualizada do tamanho do efeito na análise estatística (Cohen, 1988; Wilkinson et al., 1999), a práxis nacional demonstra um atraso considerável. A análise de

conteúdo de 347 artigos quantitativos publicados nos periódicos Qualis-Capes A2 entre 2022 e 2023 revelou uma adoção incipiente dessa prática, com apenas 7% dos trabalhos estabelecendo uma conexão explícita entre a magnitude do efeito e suas implicações práticas no mundo dos negócios. Surpreendentemente, essa deficiência não parece radicar em um desconhecimento teórico, visto que os pesquisadores brasileiros frequentemente referenciam os autores seminais e as diretrizes da APA. Em vez disso, o achado sugere uma desconexão entre o rigor metodológico preconizado e a efetiva condução e comunicação da pesquisa no contexto nacional.

O resultado central desta investigação expõe um paradoxo preocupante, a produção de ciência avaliada como de alto nível (em periódicos com Qualis elevado) frequentemente não cumpre integralmente seu papel de informar e influenciar a prática gerencial. Essa lacuna apresenta evidências de ser um ponto que merece atenção na área de administração, onde efeitos de pequena magnitude estatística podem traduzir-se em impactos financeiros substanciais; um incremento modesto na lucratividade, por exemplo, pode gerar ganhos significativos em grandes organizações. A persistente priorização da significância estatística (valor de p) em detrimento da avaliação da relevância prática dos resultados contribui para a manutenção de uma ciência desconectada das necessidades concretas do ambiente organizacional.

A principal contribuição deste artigo reside na quantificação e na evidenciação empírica dessa lacuna na produção científica brasileira. Os dados apresentados indicam que o desafio premente não se limita à disseminação do conhecimento teórico, mas exige uma transformação cultural na maneira como concebemos, conduzimos e comunicamos nossas pesquisas. Para catalisar essa mudança, identificamos três frentes de ação prioritárias: o fortalecimento do rigor metodológico na interpretação dos resultados, o estabelecimento de critérios editoriais mais exigentes por parte dos periódicos científicos em relação à contextualização prática dos achados, e o investimento na formação de pesquisadores com a capacidade de traduzir efeitos estatísticos em implicações gerenciais tangíveis.

Como qualquer pesquisa, esta possui limitações importantes a serem consideradas. Primeiramente, o foco em periódicos brasileiros classificados como Qualis-Capes A2 entre 2022 e 2023 limita a generalização dos resultados para outros tipos de publicações, áreas de estudo ou períodos. Em segundo lugar, a metodologia de análise de conteúdo, embora útil, pode estar sujeita a vieses na categorização e interpretação dos dados, apesar dos esforços para mitigar esses vieses. Além disso, a definição de "interpretação no mundo real" do tamanho do efeito pode ser um tanto subjetiva, e a pesquisa se concentra principalmente na área de negócios no Brasil, restringindo a aplicabilidade dos achados para outros campos. Por fim, a análise não se aprofunda na qualidade ou profundidade das interpretações do tamanho do efeito, focando principalmente na sua presença ou ausência, e a amostra pode não representar adequadamente as práticas de pesquisa em periódicos de menor classificação.

O caminho a seguir demanda uma reorientação de foco: da mera validação de hipóteses estatísticas para a comunicação do valor aplicado de nossas descobertas. Somente assim, a pesquisa em administração no Brasil poderá cumprir plenamente sua função de estabelecer uma ponte robusta entre a academia e o mundo organizacional, não necessariamente através da formulação de novas teorias, mas pela adoção consistente de práticas metodológicas sólidas e já estabelecidas no cenário internacional. Embora a transformação necessária seja de natureza cultural, os benefícios para a ciência, para as organizações e, em última instância, para a sociedade, serão inequivocamente concretos.

REFERÊNCIAS:

- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.).
- Associação Brasileira de Normas Técnicas. (2011). NBR 14724: Informação e documentação – Trabalhos acadêmicos – Apresentação.
- Associação Brasileira de Normas Técnicas. (2023). NBR 10520: Informação e documentação – Citações em documentos – Apresentação.
- Baird, M. D., & Pane, J. F. (2019). Translating Standardized Effects of Education Programs Into More Interpretable Metrics. *Educational Researcher*, 48(4), 217–228. <https://doi.org/10.3102/0013189X19848729>
- Bakker, A., Cai, J., English, L., Kaiser, G., Mesa, V., & Van Dooren, W. (2019). Beyond small, medium, or large: Points of consideration when interpreting effect sizes. *Educational Studies in Mathematics*, 102(1), 1–8. <https://doi.org/10.1007/s10649-019-09908-4>
- Bass, L., & Semetko, H. A. (2021). *Content Analysis*. <https://www.oxfordpoliticstrove.com/display/10.1093/hepl/9780198850298.001.0001/hepl-9780198850298-chapter-13>
- Bido, D. de S., Souza, C. A. de, Silva, D. da, Godoy, A. S., & Torres, R. R. (2012). Qualidade do Relato dos Procedimentos Metodológicos em Periódicos Nacionais na Área de Administração de Empresas: O caso da modelagem em equações estruturais nos periódicos nacionais entre 2001 e 2010. *Organizações & Sociedade*, 19(60), Artigo 60. <https://periodicos.ufba.br/index.php/revistaoes/article/view/11191>
- Brei, V. A., & Liberali Neto, G. (2006). O Uso da técnica de modelagem em equações estruturais na área de marketing: Um estudo comparativo entre publicações no Brasil e no exterior. *Revista de Administração Contemporânea*, 10, 131–151. <https://doi.org/10.1590/S1415-65552006000400007>
- Bryman, A. (2016). *Social research methods* (Fifth edition). Oxford University Press.
- Burrows, K., Figueroa-Hall, L. K., Alarbi, A. M., Stewart, J. L., Kuplicki, R., Tan, C., Hannafon, B. N., Ramesh, R., Savitz, J., Khalsa, S., Teague, T. K., Risbrough, V. B., & Paulus, M. P. (2023). Association between inflammation, reward processing, and ibuprofen-induced increases of miR-23b in astrocyte-enriched extracellular vesicles: A randomized, placebo-controlled, double-blind, exploratory trial in healthy individuals. *Brain, Behavior, & Immunity - Health*, 27, 100582. <https://doi.org/10.1016/j.bbih.2022.100582>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45(12), 1304–1312. <https://doi.org/10.1177/0956797613504966>

- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49(12), 997–1003. <https://doi.org/10.1037/0003-066X.49.12.997>
- Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- Ellis, P. D. (2010). Effect sizes and the interpretation of research results in international business. *Journal of International Business Studies*, 41(9), 1581–1588.
- Fey, C. F., Hu, T., & Delios, A. (2023). The measurement and communication of effect sizes in management research. *Management and Organization Review*, 19(1), 176-197. <https://trainingmag.com/2023-training-industry-report/>
- Fuchs, K. (2023). A Systematic Guide for Conducting Thematic Analysis in Qualitative Tourism Research. *Journal of Environmental Management and Tourism*, 14(6), Artigo 6. [https://doi.org/10.14505/jemt.v14.6\(70\).17](https://doi.org/10.14505/jemt.v14.6(70).17)
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156–168. <https://doi.org/10.1177/2515245919847202>
- Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2–18. <https://doi.org/10.1016/j.jwb.2017.11.003>
- Gaur, A., & Kumar, M. (2018). A systematic approach to conducting review studies: An assessment of content analysis in 25 years of IB research. *Journal of World Business*, 53(2), 280–289. <https://doi.org/10.1016/j.jwb.2017.11.003>
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8 <https://doi.org/10.4097/kja.24089>
- Hair, J. F. (2019). *Multivariate data analysis* (Eighth edition). Cengage.
- Herrmann, P., & Nadkarni, S. (2014). Managing strategic change: The duality of CEO personality. *Strategic Management Journal*, 35(9), 1318–1342. <https://doi.org/10.1002/smj.2156>
- Ialongo, C. (2016). Understanding the effect size and its measures. *Biochemia Medica*, 150–163. <https://doi.org/10.11613/BM.2016.015>
- Kim, S. H., Kang, H., Jun, I.-J., Park, H. W., Yoo, B. H., Lim, Y.-H., & Kim, K.-M. (2024). Effect of perioperative intravenous ibuprofen versus acetaminophen on postoperative opioid consumption and pain after general anesthesia: A systematic review and meta-analysis with trial sequential analysis of randomized controlled trials. *Korean Journal of Anesthesiology*, 77(4), 455–467. <https://doi.org/10.4097/kja.24089>
- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. American Psychological Association

- Kirk, R. E. (2003). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Sage Publications.
- <https://doi.org/10.5688/ajpe7113>
- Klein, S. P., Spieth, P., Ricart, J. E., & Noble, C. H. (2024). How media coverage elicits strategic change: The moderating role of the business model design. *Strategic Management Journal*, smj.3688. <https://doi.org/10.1002/smj.3688>
- Kleinheksel, A. J., Rockich-Winston, N., Tawfik, H., & Wyatt, T. R. (2020). Demystifying Content Analysis. *American Journal of Pharmaceutical Education*, 84(1), 7113. <https://doi.org/10.5688/ajpe7113>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Lazaraton, A. (2005). *Quantitative Research Methods*. Em *Handbook of Research in Second Language Teaching and Learning*. Routledge.
- Leland, W. (1999). Statistical methods in psychology journals: guidelines and explanations. *American psychologist*, 54(8), 594-604.
- Lipsey, M. W. (2012). *Translating the Statistical Representation of the Effects of Education Interventions Into More Readily Interpretable Forms*.
- Meneghini, E. M. P., & Lana, J. (2024). Pensata: Escolhendo entre efeitos fixos e aleatórios na análise de dados em painel. *Revista Eletrônica de Negócios Internacionais (Internext)*, 19(1), 16–23. <https://www.redalyc.org/journal/5575/557578949001/557578949001.pdf>
- Mustapha, A. I., & Ebomoyi, I. (2019). Methodological Usefulness of Content Analysis in Social Science Research. *African Research Review*, 13(1), Artigo 1. <https://doi.org/10.4314/afrev.v13i1.9>
- Naranjo-Gil, D., & Hartmann, F. (2007). Management accounting systems, top management team heterogeneity and strategic change. *Accounting, Organizations and Society*, 32(7–8), 735–756. <https://doi.org/10.1016/j.aos.2006.08.003>
- Prearo, L. C., & Gouvêa, M. A. (2011). Avaliação do emprego da técnica de modelagem de equações estruturais em teses e dissertações de universidades públicas de alta performance. *Revista da FAE*, 14(2), Artigo 2.
- Qiang, W., Wong, S., Koh, K., & Tong, Y. H. (2023). Does board turnover enhance firm performance? A contingency approach. *Corporate Governance: An International Review*, 31(3), 405–424. <https://doi.org/10.1111/corg.12455>
- Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). The Role of Between-Case Effect Size in Conducting, Interpreting, and Summarizing Single-Case Research.
- Shaver, J. M. (2019). Interpreting interactions in linear fixed-effect regression models: When fixed-effect estimates are no longer within-effects. *Strategic Management Journal*, 40(9),

1394–1430.

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—Or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, 4(3), 279–282. <https://doi.org/10.4300/JGME-D-12-00156.1>

Sun, S., Pan, W., & Wang, L. L. (2010). A comprehensive review of effect size reporting and interpreting practices in academic journals in education and psychology. *Journal of Educational Psychology*, 102(4), 989–1004. <https://doi.org/10.1037/a0019507>

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129–133. <https://doi.org/10.1080/00031305.2016.1154108>

Zhang, Y., & Rajagopalan, N. (2010). Once an outsider, always an outsider? CEO origin, strategic change, and firm performance. *Strategic Management Journal*, 31(3), 334–346. <https://doi.org/10.1002/smj.812>