

Avaliação de Artigos Científicos em Administração: Critérios e Modelos de Avaliadores Experientes

Helena Belintani Shigaki

Universidade Federal de Minas Gerais - Brasil
belintanihs@gmail.com

Roberto Patrus

Pontifícia Universidade Católica de Minas Gerais - Brasil
robertopatrus@pucminas.br

Resumo

Pretendeu-se, com este estudo, analisar o processo de revisão por pares sob a ótica dos experientes avaliadores de artigos, a fim de se obter respostas quanto aos critérios e possíveis modelos que são utilizados em um processo de avaliação, bem como se há um padrão ou uma grande variabilidade no que é utilizado hoje. A abordagem do estudo foi qualitativa, tendo como procedimento básico a análise de conteúdo. O instrumento de coleta de dados foi a entrevista em profundidade, realizada com onze avaliadores de artigos científicos de periódicos A1 e A2. Os resultados revelaram densa relação com as principais teorias aqui apresentadas, no que tange à (a) relevância da contribuição, (b) escrita/apresentação, (c) design/concepção, (d) referências, (e) metodologia, (f) discussão dos resultados, (g) teoria, (h) critérios pessoais e (i) ética. Também foi possível elucidar três modelos-chave utilizados pelos entrevistados para o processo de avaliação, aqui nomeados de: (a) hipotético-dedutivo, (b) analítico-indutivo e (c) *by the book*. O presente estudo possui contribuição teórica, em lançar luz sobre o trabalho anônimo de pesquisadores de todo o mundo responsáveis pela avaliação de artigos, projetos de pesquisa, cursos ou programas de pós-graduação e; contribuição prática, por apresentar os critérios e modelos de avaliação de avaliadores experientes.

Palavras-chave: critérios de avaliação, periódicos, avaliadores, administração, revisão por pares

Peer Review of Scientific Journal Articles in Management: Criteria and Models of Expert Reviewers

Helena Belintani Shigaki

Universidade Federal de Minas Gerais - Brasil
belintanihs@gmail.com

Roberto Patrus

Pontifícia Universidade Católica de Minas Gerais - Brasil
robertopatrus@pucminas.br

Abstract

This article has sought to analyze the process of peer review of scientific papers under the perspective of expert reviewers, in order to get answers about the criteria and models that are used in an evaluation process, and if there is a pattern or a large variability in what is currently used. The approach of the study was qualitative and the metrological procedure was content analysis. The data collection instrument was the in-depth interview, done by eleven expert A1 and A2 journals reviewers. The results showed dense relationship with the main theories presented here, concerning the (a) importance of the contribution, (b) writing/ presentation, (c) design/ concept, (d) references, (e) methodology, (f) discussion of the results, (g) theory, (h) personal criteria, (i) ethic. It was also possible to elucidate three key-models used for peer review by respondents, here named (a) hypothetico-deductive, (b) analytical-induction, (c) by the book. This study has theoretical contribution, shed some light on the anonymous work of researchers around the world responsible for the evaluation of articles, research projects, courses or graduate programs and; practical contribution, for presenting the criteria and evaluation models of expert reviewers.

Keywords: evaluation criteria, scientific journals, reviewers, management, peer review

A atividade de revisão por pares é aqui compreendida como um processo sistemático de avaliação constituído por dois ou mais revisores os quais fornecem um parecer acerca de um artigo científico (Bloom, 2006, Hames, 2012, Mulligan, Hall, & Raphael, 2013, Smith, 2006, Stumpf, 2008, Ware, 2008). Embora os participantes desse sistema reconheçam sua importância, ultimamente ele tem sido alvo de várias críticas. Além disso, apesar de ser um assunto amplamente debatido, trata-se de um tema pouco explorado empiricamente na comunidade acadêmica, em particular na área de Administração.

A origem da atividade de revisão por pares (*peer review*) é incerta. Há aqueles que acreditam que sua origem está na censura do Estado, no século XVII; ou da realeza, com o intuito de estabelecer práticas de conduta para a sociedade (Fitzpatrick, 2011, Pessanha, 1998), e outros que acreditam que somente após a Segunda Guerra Mundial ela foi institucionalizada (Media, 2008; Mulligan, Hall, & Raphael, 2013). Outra fonte de informação que contribui com essa elucidação histórica está na origem de dois periódicos, com data de 1665 (*Journal des Sçavans*, na França, e o *Philosophical Transactions*, da *Royal Society*, na Inglaterra), que já utilizavam o processo de revisão por pares, formado por uma sociedade de cientistas que avaliavam os projetos submetidos (Davyt & Velho, 2000, Patrus, Dantas, & Shigaki, 2013). E, no século XVIII, na Escócia, as técnicas de avaliação foram institucionalizadas pela *Royal Society of Medicine*, tais como são utilizadas hoje (Davyt & Velho, 2000).

Compreender a revisão por pares sob a ótica do avaliador, que está na base de toda atividade acadêmica, permitirá refletir sobre os rumos da atividade docente e de pesquisa nos dias atuais, em particular na área de Administração. Nesse contexto, pretende-se responder à seguinte questão: Como ocorre o processo de revisão por pares pelos avaliadores de artigos de periódicos científicos? Para responder a essa pergunta, dois objetivos específicos foram traçados: (a) analisar os critérios de revisão de artigos científicos utilizados por avaliadores e (b) descrever os modelos de revisão de artigos científicos utilizados por avaliadores.

Para tanto, a abordagem do estudo foi qualitativa, tendo como procedimento básico a análise de conteúdo. O instrumento de coleta de dados foi a entrevista em profundidade. Os resultados apontaram densa relação com os critérios de avaliação obtidos nas pesquisas de Ali, Young e Ali (1996) e Bornmann, Nast e Daniel (2008), Ward, Hall e Scharamm (1975) a saber: (a) relevância da contribuição, (b) escrita/ apresentação, (c) design/concepção, (d)

referências, (e) metodologia, (f) discussão dos resultados, (g) teoria, (h) critérios pessoais e (i) ética. Também foi possível elucidar três modelos-chave utilizados pelos entrevistados para o processo de avaliação: (a) hipotético-dedutivo, (b) analítico-indutivo e (c) *by the book*.

Foi realizada a revisão de literatura em livros e artigos nacionais e internacionais. Observou-se que o tema tem sido discutido, principalmente, por autores brasileiros, o que justifica o uso, aqui, em sua maioria, de artigos nacionais (56%). Organizamos esse estado da arte a partir das críticas feitas ao sistema, e dos argumentos favoráveis a ele. Em seguida, abordamos os modelos de avaliação de artigos e alguns critérios de avaliação, fruto das escassas pesquisas empíricas encontradas a partir do levantamento bibliográfico.

Revisão por Pares e Critérios de Avaliação de Artigos Científicos

O termo revisão por pares não possui uma definição rigorosa (Pagani, 2002) e diversos autores nacionais e estrangeiros o define de diferentes formas. Assim sendo, revisão por pares pode ser um processo de submissão de manuscritos acadêmicos para avaliação de um ou mais especialistas na mesma área, antes de sua publicação (Bloom, 2006, Hames, 2012, Mulligan, Hall, & Raphael, 2013, Stumpf, 2008, Ware, 2008). Normalmente, são os avaliadores e editores que julgam a qualidade dos trabalhos (Coupé, 2010, Smith, 2006) e mantêm entre eles um processo de responsabilidade mútua (Pagani, 2002). Podendo ser compreendido, também, como medida, avaliação ou julgamento (Botomé, 2011) de pesquisadores cujo propósito “é segregar os trabalhos de baixa qualidade, mantendo elevados os padrões de produção científica” (Serra, Fiates, & Ferreira, 2008, p. 36). Corroborando tal afirmação, para Davyt e Velho (2000), a tradição de que os trabalhos devem ser avaliados apenas por cientistas ajuda a compreender o sentido das palavras. Por “pares”, compreende-se os colegas e competidores, cuja experiência e formação tornaram possível essa emissão de uma opinião confiável.

Em uma pesquisa realizada por Ware (2008), com 3.040 entrevistados, constatou-se que 93% concordam que a revisão por pares seja necessária, 85% concordam que ela ajuda a melhorar a comunicação científica e, ainda, 83% acreditam que, sem a revisão por pares, não haveria controle. Em outro estudo, realizado por Mulligan, Hall e Raphael (2013), com 4.037 entrevistados, 69% diziam estar satisfeitos com o atual sistema e 94% concordam que a crescente melhoria da qualidade dos artigos publicados se deve ao processo de revisão por

pares. As críticas a esse sistema não estão atreladas, necessariamente, à ideia de abandono desse, mas sim visando a sua melhoria (Omote, 2005). Podem-se observar na Figura 1, os argumentos contrários na primeira coluna, por ordem de citações e, nas colunas seguintes, os autores correspondentes às críticas.

Argumentos contrários	Autores (ano)																
	Meneghini & Fonseca 1990	Kostoff 1997	Pessanha 1998	Campanário 2002	Kruger 2005	Omote 2005	Williamson 2005	Pinho 2005	Thiry-Cherques 2005	Jennings 2006	Smith 2006	Media 2008	Serra, Fiates e Ferreira 2008	Ware 2008	Botomé 2011	Hames 2012	Mulligan, Hall e Raphael 2013
Tempo e atraso na publicação			X	X		X		X		X	X		X	X			X
Abuso de poder				X		X			X		X				X		
Parcialidade		X	X									X			X		
Subjetividade					X	X					X				X		
Preconceito			X										X			X	X
Julgamentos inadequados ou inconsistentes			X			X			X		X					X	
Competição	X					X											
Avaliações negativas		X	X			X											
Sobrecarga de trabalho												X	X				
Dificuldades em detectar erros e fraudes							X				X						
Custo das avaliações		X															
Acesso a informação privilegiada				X													
Desprezo									X								
Falta de sigilo													X				
Injustiça												X					

Figura 1. Argumentos contrários à revisão por pares

Fonte: Elaborado pelos autores, a partir da revisão da literatura.

Uma das mais citadas críticas à revisão por pares está relacionada ao tempo de resposta por parte dos avaliadores e, conseqüentemente, ao atraso da publicação. Já o abuso de poder, ora do avaliador ora do editor (Botomé, 2011, Campanario, 2002, Omote, 2005, Smith, 2006, Thiry-Cherques, 2005), segunda crítica mais citada, pode induzir pareceres inconsistentes ou indevidos (Omote, 2005), seja por resistência a mudanças propostas ou pelos autores contrariarem seu atual ‘sistema de ideias’. As críticas relacionadas à parcialidade (Botomé, 2011, Kostoff, 1996, Media, 2008; Pessanha, 1998) podem ser motivadas pela falta de razões técnicas no processo de avaliação (Kostoff, 1996) e também por fatores que podem invalidar uma análise futura, tais como: citação por apreço ou por obediência às regras editoriais, uso de autores consagrados para realçar o texto, ou citação de colegas da mesma instituição, para prestigiá-los (Freitas, 1998).

A crítica referente à subjetividade (Botomé, 2011, Kruger, 2005, Omote, 2005, Smith, 2006;) varia quanto à condição humana de cada avaliador, a seu modo de vida, seu estilo e suas experiências. Os avaliadores podem, por exemplo, reprovar artigos que apresentam certas inovações e aprovar artigos que não apresentam contribuições consistentes para o meio acadêmico ou empresarial (Stumpf, 2008). Assim, os julgamentos ou pareceres podem

variar em grau, qualidade, profundidade ou quantidade (Botomé, 2011). Já no que tange ao preconceito no processo de avaliação (Hames, 2012, Mulligan, Hall, & Raphael, 2013, Pessanha, 1998, Ware, 2008), ele ocorre quando o avaliador julga o artigo antes de lê-lo com base em poucas informações obtidas, sejam elas relacionadas à nacionalidade do autor, idioma, gênero, referências utilizadas, dentre outros.

Os julgamentos inadequados (Hames, 2012, Pessanha, 1998) ou inconsistentes (Smith, 2006; ThiryCherques, 2005), sejam eles por erro (proposital ou não) ou por competição, também são passíveis de ocorrer. Esses, no caso, não apontam caminhos adequados para os autores (Gondim, 2004). Podem-se ainda considerar, como causas prováveis, a falta de tempo, a falta de interesse do avaliador (Thiry-Cherques, 2005), o abuso de poder, a falta de treinamento e a falta de adequação à área do avaliador (Serra, Fiates, & Ferreira, 2008). Outro critério, a competição entre avaliadores e avaliados (Botomé, 2011, Meneghini & Fonseca, 1990, Omote, 2005), pode ocorrer dependendo da área de cada um, e a probabilidade de haver maior competição ocorre em áreas relativamente pequenas e com poucos recursos, aumentando assim a competição por verbas (Meneghini & Fonseca, 1990) e por espaço de publicação. Outro tipo de competição seria para garantir ou manter o prestígio ou status de um avaliador em seu grupo acadêmico (Miller, 2006).

As avaliações negativas ocorrem por favorecimento próprio (Kostoff, 1996, Omote, 2005; Pessanha, 1998), advindo da competição ou por conflito de interesse, enfrentamentos pessoais ou interesses comerciais (Pessanha, 1998). Outro critério evidenciado foi a sobrecarga de trabalho (Serra, Fiates, & Ferreira, 2008, Ware, 2008) estudada por Ware (2008), que identificou, em seu trabalho, que a grande maioria dos comentários foi realizada por um mesmo grupo de avaliadores e, em alguns casos, podem-se atribuir os problemas de tempo e qualidade dos pareceres à sobrecarga de trabalho (Phelan, Ferreira, & Salvador, 2002). Cabe aqui apontar o conceito de solidariedade, que pode ser compreendido como uma forma de coletividade, contrapondo-se ao individualismo advindo da especialização (Selli & Garrafa, 2006), atuando no sentido de comprometimento (Serva, 1997) e de cooperação (Spink, 2003).

A dificuldade em detectar erros e fraudes nos artigos (Smith, 2006, Willaïsson, 2003) está relacionada à sobrecarga de trabalho e necessidade em encaminhar um rápido parecer ao editor. Hoje há ferramentas gratuitas que detectam plágio nos artigos (por exemplo, CopySpider, Farejador de Plágio, Plagius, Viper, entre outros). O editorial *Three cheers for peers* (três vivas aos pares), do *International Weekly Journal of Science* (Nature, 2006),

afirma que o sistema atual é sem dúvida o melhor que há disponível e que ele não é projetado para detectar fraudes de uma pequena minoria de artigos.

As críticas sobre custos das avaliações (Kostoff, 1996) e o custo do atraso da publicação para a comunidade científica (Jennings, 2006) estão atreladas à primeira e mais recorrente crítica, a do tempo. Também pode estar vinculada à recompensa monetária aos avaliadores, apesar de não ser uma prática comum, principalmente no Brasil.

O acesso às informações privilegiadas (Campanario, 2002) é considerado uma crítica quando ocorre alguma falha na transparência (Lee & Bero, 2006) do sistema e na ética dos avaliadores. Atreladas a isso, a falta de sigilo (Ware, 2008) e a injustiça (Media, 2009) podem ocorrer quando editores e avaliadores permitem “que trabalhos não tão valiosos sejam publicados, enquanto outros que poderiam representar uma inovação são rejeitados. O processo é, portanto, passível de subjetividade” (Stumpf, 2008, p. 22), ou seja, os pareceres podem variar de acordo com cada avaliador (Botomé, 2011). No mesmo sentido, a crítica do desprezo (Thiry-Cherques, 2005) tem por fundamento o aborrecimento do avaliador em realizar uma atividade não remunerada, em um momento de sobrecarga de trabalho, para redação de um parecer consistente.

Os argumentos favoráveis, por outro lado, aparecem em menor número na literatura nacional e estrangeira e, foram organizados na Figura 2.

Autores:	Freitas	Campanário	Gondim	Borba & Murcia	Serra, Fiates & Ferreira	Ware	Fitzpatrick
Data:	1998	2002	2004	2006	2008	2008	2011
Construção e difusão do conhecimento		X		X		X	X
Garantia de qualidade					X	X	
Filtro de publicação						X	
Imparcialidade				X			
Rotatividade dos avaliadores			X				
Controle	X						

Figura 2. Argumentos favoráveis à revisão por pares

Fonte: Elaborado pelos autores, a partir da revisão da literatura.

Diante do que foi ilustrado na Figura 2, a maioria dos argumentos favoráveis à revisão por pares corrobora a ideia de que ela é uma atividade que auxilia na construção e difusão do conhecimento científico (Borba & Murcia, 2006, Campanario, 2002, Fitzpatrick, 2011, Ware, 2008). Isso se deve ao fato de termos professores avaliadores voluntários comprometidos com a ciência e com a importância do aumento do rigor e a validade do

conhecimento (Campanário, 2002). Seguida pelo segundo argumento mais citado pelos entrevistados, a qualidade de uma avaliação ou parecer decorre da formação e experiência dos avaliadores. Aqueles que participam ou participaram de uma avaliação tendem a produzir artigos com maior rigor acadêmico (Serra, Fiates, & Ferreira, 2008). Outro argumento favorável refere-se ao filtro do trabalho, realizado pelos avaliadores ou editores, e que consiste em selecionar e classificar os artigos recebidos como de baixa, média ou alta qualidade. Isso auxilia no posicionamento e na manutenção da qualidade do periódico e pode também ser utilizado em casos de fraude ou plágio (Ware, 2008).

A imparcialidade com que os avaliadores participam é vista como um resultado positivo diante da avaliação por pares. Especula-se que não são todos os avaliadores que possuem essa característica, isso porque, quando há falta de informação em relação ao nome e à origem dos autores avaliados, o avaliador é menos influenciado por um viés pessoal (Borba & Murcia, 2006). Técnicas de *double blind review* e *total blind review* objetivam o encorajamento de opiniões francas e qualidade de um parecer (Mulligan, Hall, & Raphael, 2013).

A rotatividade dos avaliadores também foi considerada, embora por poucos, como uma característica positiva do sistema. Assim, os grupos de interesse que eventualmente são formados passam a não existir, evitando a característica negativa de favorecimento, em que o ato de avaliar é visto como uma ferramenta de poder (Gondim, 2004). Por último, o controle da atividade científica já está estruturado em agências, cursos superiores e corpos editoriais, por meio de critérios claros e definidos de avaliação. O objetivo do processo de avaliação, em termos de controle, também é o de assegurar que o artigo disponibilizado ao leitor seja de qualidade e protegido de subjetividade e abuso de poder (Mulligan, Hall, & Raphael, 2013). Diante dos argumentos contrários e favoráveis ao processo de revisão por pares, pode-se pensar em como essa avaliação é feita, seus critérios e modelos de avaliação, temas discutidos a seguir.

Embora muito se tenha escrito, sobre a avaliação ou revisão por pares, as pesquisas empíricas sobre os critérios de avaliação de artigos ainda são muito escassas. Ward, Hall, & Scharam (1975), Ali, Young e Ali (1996) e, mais recentemente, Bornmann, Nast, & Daniel (2008) foram as principais referências encontradas. Os dois primeiros se ativeram a reavaliar artigos, já publicados, com os critérios preestabelecidos por eles. O último procurou, entre 46 estudos publicados entre 1967 e 2006, critérios de avaliação de artigos. Foram encontrados 682 critérios que, posteriormente, foram agrupados em 485 e, por

consequente, divididos em nove áreas, a saber: (a) relevância da contribuição, (b) escrita, (c) concepção, (d) método, (e) discussão dos resultados, (f) referências, (g) teoria, (h) afiliação e (i) ética.

A relevância da contribuição do artigo é avaliada pela qualidade e importância (do tema, para o avanço da ciência, para a revista e a partir dos resultados obtidos) e originalidade. A escrita ou apresentação é avaliada pelo estilo de escrita do autor, pela qualidade de seções específicas (como resumo, objetivo, problemática), pela organização, pelas normas de publicação e, pela adequação e utilização das ilustrações. Conforme Calado Dias, Patrus, & Magalhães (2011), deve-se ter atenção com os cuidados formais de um texto. O design ou concepção é avaliado pelo marco teórico, qualidade e adequação do artigo (competência de conteúdo), qualidade da amostragem, generalização e replicabilidade. O método é avaliado a partir de sua adequação, qualidade e validade de operação e medição e também pela apresentação de novas metodologias ou novas formas de uso das usualmente conhecidas. A discussão dos resultados é avaliada pela adequação e objetividade, clareza, amplitude de interpretação e discussão dos resultados, além da importância ao suporte teórico. Esses cinco primeiros critérios foram citados em grande número, tanto por autores de artigos (Ali, Young, & Ali, 1996, Bornmann, Nast, & Daniel, 2008, Botomé, 2011; Campanario, 2002, Hébert, 2007, Kirschbaum & Mascarenhas, 2009, Kruger, 2005, Media, 2008, Mulligan, Hall, & Raphael, 2013 Omote, 2005, Pinto & Andrade, 1999, Rosa & Chachamovich, 2008, Serra, Fiates, & Ferreira, 2008, Ward, Hall, & Scharam, 1975, Ware, 2008) quanto pelo estudo de Bornmann, Nast, e Daniel (2008).

A referência é avaliada por sua relevância, principalmente. Em alguns casos pode também ser avaliada por meio da quantidade de citações e verificação da qualidade, ora pelos autores, ora pela data de publicação. A teoria do artigo é avaliada por meio da apresentação de novidades e interesse dos autores, contribuição teórica e a teoria no geral, sendo correspondente ao critério anterior, referência. Esses dois critérios, embora menos mencionados que os cinco primeiros, ainda possuem destaque na avaliação (Ali, Young, & Ali, 1996, Bornmann, Nest, & Daniel, 2008, Ward, Hall, & Scharam, 1975;).

Por fim, têm-se os dois últimos critérios. A reputação ou afiliação é um critério em que se observa a origem dos autores, também pela quantidade de autores que é colocada em um único artigo. E a ética é avaliada pela ética em si, bem como pela possibilidade de publicação múltipla ou repetida e plágio (Ali, Young, & Ali, 1996, Bornmann, Nest, & Daniel, 2008, Omote, 2005, Ward, Hall, & Scharam, 1975, Ware, 2008). Em geral, com esses princípios,

pode-se pensar em um sistema de avaliação sem discriminação, com garantia de qualidade por meio de políticas e práticas de gestão (Pagani, 2002).

Método

Pretendeu-se, com este estudo, analisar o processo de revisão por pares sob a ótica dos avaliadores experientes de artigos de periódicos científicos. Para isso, foi utilizada uma abordagem de estudo qualitativa, tendo como procedimento básico a análise de conteúdo. A visão qualitativa é a mais adequada para esta pesquisa porque ela engloba as formas simbólicas e se interessa pelo caráter comunicativo das experiências e necessidades sociais, o que possibilita a compreensão de um fenômeno social (Godoy & Balsini, 2004), como é o caso da revisão por pares. A análise de conteúdo é considerada uma técnica híbrida, também utilizada em nível de qualidade e distinção de respostas, antes mesmo de uma quantificação (Bauer, 2002). É capaz de produzir inferências do particular (entrevistas) para o geral (contexto social) de maneira objetivada (Bauer, 2002), além de permitir o acompanhamento da evolução de um fenômeno (Godoy, 1995) em trabalhos futuros.

Para a coleta de dados, o instrumento foi a entrevista em profundidade, como fonte de evidência primária (Eisenhardt, 1989). O objetivo das entrevistas foi recolher percepções sobre o tema de pesquisa a partir da perspectiva do entrevistado (Meyer, 2001) e explorar as diferentes representações acerca do tema estudado (Gaskell, 2002). Além das perguntas demográficas, outras foram feitas, respeitando a diferença entre os termos critérios e modelos. Para critérios, os autores formularam as seguintes perguntas: (a) quais critérios você utiliza para avaliar um artigo?, (b) como eles são organizados hierarquicamente?, (c) há diferença nos critérios utilizados para diferentes periódicos com diferentes classificações?, (d) o que é fundamental para o artigo ser aprovado? e, (e) o que o faz reprovar um artigo? Para métodos, os autores formularam as seguintes perguntas: (a) como você faz a sua avaliação de um artigo? E, (b) há diferença no método utilizado para diferentes periódicos com diferentes classificações?.

Para selecionar os entrevistados, professores avaliadores de periódicos, duas etapas foram realizadas: primeira, com a intenção de localizar quais professores seriam entrevistados, todos os *sites* de periódicos, em Administração, classificados como A2 e A1 conforme critérios Qualis Capes, foram visitados no segundo semestre de 2012, e as informações sobre os avaliadores formaram uma base de dados. O segundo filtro ocorreu

por meio da amostragem não probabilística por julgamento, optando pela diversidade dos entrevistados, no quesito experiência como avaliador de artigos para periódicos, medida em termos de anos como revisor de artigos. Ou seja, para este estudo os avaliadores possuíam em média, 11 anos de experiência em avaliação de trabalhos acadêmicos.

Em relação à quantidade de entrevistados, optou-se pela utilização de dois critérios, mutuamente excludentes: primeiro, em relação à quantidade propriamente dita, e, segundo, a partir da saturação empírica, ou seja, quando não há incidências de novas ideias ou informações por parte dos entrevistados (Gibbs, 2009). Alguns autores argumentam que, para uma pesquisa qualitativa, é suficiente atingir o número de 6 a 12 entrevistados (Paavilainen & Asted, 1997; Partis, 2003). Outros autores apontam entrevistar o maior número de pessoas possível, a fim de obter diversidade nas informações (Godoi, Bandeira-de-Melo, & Silva, 2007), ora obter saturação empírica, justificando a não ampliação da quantidade de entrevistados (Poupart, Deslauriers, Groulx, Laperrière, Mayer, & Pires, 2012). Partindo desses critérios, foram entrevistados, no segundo semestre de 2013, 11 revisores da unidade federativa dos autores, a partir de uma seleção prévia (priorizou-se aqueles com mais experiência em avaliações) e também pela técnica *snowball* (bola de neve), a partir da qual um entrevistado indicou o outro em caráter sucessivo (Aaboen & Frida Lind, 2012). Todos os professores avaliadores que foram convidados a nos ceder uma entrevista, aceitaram participar. Exceto uma professora convidada não participou, devido a indisponibilidade na agenda naquele mês. Eles foram convidados a participar dessa pesquisa por *e-mail*, feito pelos próprios autores. As entrevistas, com duração média de 47 minutos, foram realizadas pessoalmente, com exceção de uma, via Skype (aplicativo de comunicação *online*). Todas foram gravadas, mediante autorização dos entrevistados, transcritas e analisadas. Por ser a transcrição considerada uma mudança de meio (da fala para a escrita), essas foram realizadas pelos próprios autores, mantendo assim a precisão e fidelidade da interpretação (Gibbs, 2009). Quanto ao formato da entrevista, ela iniciou-se com perguntas demográficas estruturadas (por exemplo: há quanto tempo o(a) professor(a) é avaliador(a), quantos artigos são avaliados por ano, qual o tempo efetivo de trabalho) e, posteriormente, com perguntas elaboradas à luz da revisão bibliográfica realizada, composta por perguntas não estruturadas, o que possibilitou ao entrevistado nos contar sobre o seu 'ritual' de revisão (incluindo critérios e modelos).

Foram entrevistados nove homens e duas mulheres. Os entrevistados apresentaram uma experiência como avaliadores, que variou de 5 a 21 anos e se iniciou com um convite da

própria revista em que o avaliador já havia publicado um artigo anteriormente. Todos participam de avaliações de periódico nível A e a grande maioria também dos de nível B. A quantidade de artigos avaliados varia de 4 a 21 por ano. O tempo efetivo de trabalho, em horas, também apresentou variações entre 30 minutos a 20 horas, por artigo.

Para análise do conteúdo das transcrições das entrevistas, procurou-se ressaltar os temas comuns, com o objetivo de identificar similaridades e diferenças entre as falas analisadas, conforme aconselha Godoy (1995). A codificação, ou categorização, foi construída à medida que as análises foram realizadas, e a escolha dessas unidades de registro deve responder de maneira pertinente aos objetivos propostos no estudo (Bardin, 2011). Em momento algum foram utilizados *softwares* (CAQDAS - *Computer assisted qualitative data analysis software*) para transcrição ou até mesmo para análise dos resultados, pois se pretendeu evitar a “tecnicização” e não suscitar uma análise estruturada (Godoi, Bandeira-de-Mello, & Silva, 2007). A análise foi realizada de forma individual, primeiramente, e posteriormente em conjunto com as demais entrevistas realizadas. O que possibilitou a criação da categorização.

Análise e Discussão dos Resultados

Nessa seção, são apresentados os dois principais resultados obtidos. O primeiro resultado refere-se aos critérios de avaliação, divididos em 9 códigos analíticos, a saber: (a) relevância da contribuição, (b) escrita/ apresentação, (c) design/concepção, (d) referências, (e) metodologia, (f) discussão dos resultados, (g) teoria, (h) critérios pessoais e (i) ética. O segundo resultado quanto aos modelos de métodos de avaliação, formulados pelos avaliadores a partir dos critérios de avaliação, podendo ser (a) hipotético-dedutivo, (b) analítico-indutivo, e (c) *by the book* (conforme manual).

Quanto aos Critérios de Avaliação

As avaliações de artigos de periódicos científicos da área de Administração seguem, na maioria das vezes, um padrão exigido pela própria revista, via formulário de avaliação que é

entregue ao avaliador junto ao artigo. No entanto, apesar de seguirem esses critérios, os avaliadores têm consciência de que alguns deles são mais importantes que outros e deveriam ter uma relevância maior. E como não há esse peso de avaliação ou julgamento, essa tarefa fica incumbida ao próprio avaliador. Os códigos analíticos identificados, a partir das entrevistas realizadas, neste estudo são: (a) relevância da contribuição, (b) escrita/apresentação, (c) design/concepção, (d) referências, (e) metodologia, (f) discussão dos resultados, (g) teoria, (h) critérios pessoais e (i) ética.

Percebe-se que dos nove critérios identificados pelo estudo de Bornmann, Nest, e Daniel (2008), somente um não foi mencionado por qualquer um dos entrevistados, a reputação dos autores avaliados em função da avaliação no Brasil ser, em sua maioria, *double blind review*. No entanto houve o surgimento de um novo critério não mencionado anteriormente na pesquisa ora citada: os critérios pessoais. Adiante, a discussão de cada um desses, em ordem decrescente de citações pelos entrevistados.

Relevância da Contribuição.

O código analítico 'relevância da contribuição' possui quatro categorias: (a) relevância do tema para a área; (b) relevância do tema para o avanço da ciência; (c) originalidade; e (d) relevância do tema para a revista. Esse critério foi o único citado por todos os entrevistados, o que lhe confere importância. A relevância do tema para a área engloba também a relevância do tema para a ciência e para a revista, mas possui suas especificidades, a começar pela análise do tema e sua relação com a área de pesquisa, preocupação do avaliador em identificar a contribuição e relevância do artigo, bem como o objetivo e problemática deste, conforme elucidam Bornmann, Nest e Daniel (2008). Em muitos casos, o objetivo não é de fácil identificação ou não possui a relevância adequada, e essa falha logo no início do artigo pode ocasionar a leitura com um viés negativo.

A relevância do tema para o avanço da ciência é avaliada com foco na contribuição e no fato de possuir um assunto claro e instigante. O potencial, a ruptura e o fato de preencher uma lacuna teórica ou empírica são também fatores que contribuem positivamente com essa avaliação, dados que corroboram com os estudos de Bornmann, Nest e Daniel (2008).

Igualmente, a atualidade é um ponto também observado nas avaliações, seja no tema, na proposta e, principalmente, nas referências utilizadas. No entanto é necessário também

que haja um equilíbrio de literatura moderna com os clássicos, criando assim um diálogo entre os autores e possivelmente uma contribuição maior para a área.

A originalidade está relacionada ao grau de ineditismo do artigo como um todo e quanto novo ele é para a área. Em alguns casos, nas entrevistas, foi demonstrado que a originalidade é flexível, ou seja, também é considerado inovador o estudo que faz uma nova aplicação de um método já existente, mas com novas interpretações e conhecimentos, o que corrobora pesquisa de Ali, Young e Ali (1996). Em consonância, as diferentes classificações das revistas exigem níveis diferentes de originalidade.

Por fim, tem-se a relevância do tema para a revista. O avaliador preocupa-se em manter o padrão de qualidade da revista em termos de publicação e até mesmo seu posicionamento diante da comunidade acadêmica. Não cabe aqui pensar nas regras e normas de publicação, mas sim na linha editorial da revista. A relevância da contribuição pode ser considerada como o critério com maior peso em uma avaliação, devido a sua importância relacionada à relevância do assunto, que deve ser claro e oportuno para a área e para o periódico almejado. Já comprovado por Bornmann, Nest e Daniel (2008), 45 dos 46 casos analisados eles levavam este critério em consideração e aqui todos os entrevistados o citaram. Casos em que o estudo não apresenta um avanço no conhecimento ou importância científica são avaliados de forma negativa e conseqüentemente o produto final é considerado como inconclusivo.

Escrita/ Apresentação.

O código analítico ‘escrita/apresentação’ possui quatro categorias-chave: (a) normas de publicação, (b) qualidade de partes específicas, (c) organização e (d) estilo de escrita. Esse critério foi citado por 9 dos 11 entrevistados.

As normas de publicação referem-se aos critérios que são exigidos pela revista. Para o avaliador, eles são a base para a padronização da avaliação e para a minimização da subjetividade. Basicamente, o avaliador segue um roteiro. Alguns entrevistados criticaram a adequação formal desse roteiro, indicando que este atrasa no momento da elaboração do parecer final e, em alguns casos, podem ser preenchidos de forma incorreta e enviados incompletos. Conforme dito por um entrevistado, “geralmente eu sigo um roteiro. Eu leio o roteiro e a partir dali você já fica atento ao que eles pedem [Entrevistado F]”

Na qualidade de partes específicas, preocupa-se, principalmente, com o título, resumo, palavras-chave e objetivos, tanto na qualidade quanto na clareza, delineamento e

contribuição destes. E a competência formal do artigo é avaliada por meio dessa organização e estruturação. Por último, tem-se o estilo de escrita, mais precisamente na linguagem utilizada, nos jargões, na correta utilização da língua portuguesa e, conseqüentemente, no impacto da primeira leitura. Texto com problemas de redação podem fazer o avaliador se preocupar com a forma e negligenciar o conteúdo.

A competência formal de um estudo, nomeada por Bornmann, Nest e Daniel (2008) como critério de escrita e apresentação, também possui sua relevância nos estudos ora realizados. Esse critério de avaliação torna-se importante a partir do momento que prejudica a leitura do avaliador, comprometendo sua análise e também seu tempo. Quando o trabalho está bem redigido, possui qualidade e rigor na proposta e descrições, bem como conformidade com as normas de formatação, o avaliador passa a destinar seu olhar para a competência de conteúdo, conforme afirmam Ali, Young e Ali (1996). Já Calado Dias, Patrus e Magalhães (2011) corroboram com a ideia de que a falha na forma prejudica a adequada compreensão do conteúdo.

Design/ Concepção.

O código analítico ‘design/concepção’ possui três categorias-chave: (a) qualidade e adequação, (b) estrutura e (c) possibilidade de replicação do estudo. Esse critério foi citado por 9 dos 11 entrevistados. Ao contrário do critério ‘escrita/ apresentação’, este se preocupa com a competência de conteúdo e não mais formal.

A qualidade e adequação do artigo é medida, inicialmente, pela orientação da pesquisa para responder à pergunta-problema, verificação e adequação do título, coerência entre os construtos utilizados, consistência interna e conexão das informações utilizadas. A estrutura do artigo, encadeamento de ideias, amarração ou articulação também é avaliada e corrobora o primeiro critério, de qualidade e adequação do conteúdo com o problema proposto inicialmente. Duas citações foram destacadas: “De cara eu vejo o artigo, título e resumo [...] Vou ver ali de acordo com a literatura se ele está perto de cumprir o que está prometendo [...] se tem consistência [...] que não colocou elemento demais [...]. [Entrevistado C]”. E, “O mais importante de um artigo é a questão da coerência [...] às vezes é um autor que está mal citado, um método que poderia ter falado mais [...] ver que o artigo está amarrado é bom. [Entrevistado B]”

A possibilidade de replicação do estudo é fundamental, embora menos mencionada pelos entrevistados. Essa categoria, em termos metodológicos, é utilizada para identificar, de forma clara, como o autor de um trabalho conduziu seu estudo, para que os passos possam ser seguidos por outros pesquisadores e os dados contrapostos. Conforme dito por um entrevistado, “Às vezes a pessoa não coloca questões básicas. Como que o questionário foi aplicado? Como foi isso? Como os resultados foram analisados? [...]. [Entrevistado B]”.

O design ou concepção, ou competência de conteúdo, embora apareça em menor número de códigos, foi encontrado em 43 dos 46 casos por Bornmann, Nest e Daniel (2011). Isso corrobora com o critério anterior, no qual a competência formal possui maior peso que a competência de conteúdo em uma avaliação, uma vez que a primeira estando com qualidade a segunda poderá ser mais bem avaliada. Aqui, a adequação e principalmente a possibilidade de replicar os procedimentos, para se chegar ao mesmo resultado, colaboram como sentido de generalização e validade do estudo.

Referências.

O código analítico ‘referências’ possui uma categoria-chave: (a) relevância na literatura. Esse critério foi citado por 7 dos 11 entrevistados. Verifica-se aqui a qualidade do referencial teórico do artigo, por meio das escolhas literárias do autor, divididas em clássicos ou modernos, nacionais ou internacionais e livros ou artigos, por exemplo. Conforme dito por um entrevistado, “Os autores que eles utilizam [...] utilizar autores clássicos e autores recentes [...] literatura estrangeira [...]. [Entrevistado J]”.

A qualidade das referências utilizadas em um estudo ajuda a construir um apoio teórico consistente e conseqüentemente relevância na literatura. Esse critério foi o quarto mais citado neste estudo, no entanto foi o sexto na classificação de Bornmann, Nest e Daniel (2011), presente em apenas 26 dos 46 casos estudados. Para Ali, Young e Ali (1996), o fator preponderante de medição de qualidade desse critério está na adequação das referências e quantidade de citações, que também estava presente neste estudo por meio das falas dos entrevistados. Pode-se constatar a importância desse critério, no qual avaliadores brasileiros assemelham a garantia de um estudo à qualidade dos autores utilizados e às citações realizadas.

Metodologia.

O código analítico ‘metodologia’ possui duas categorias-chave: (a) qualidade de operacionalização e medição e (b) novidade. Esse critério foi citado por 6 dos 11 entrevistados. Em consonância com os critérios já mencionados, a avaliação da metodologia atua no sentido de atestar a qualidade do estudo, sua relevância e concepção, bem como o encadeamento entre o problema mencionado e a metodologia escolhida. A novidade, assim como nos outros critérios, é medida pelo caráter inovador do método utilizado, seja ele um método novo ou uma nova releitura de um método já existente.

A metodologia, presente em 34 dos 46 estudos de Bornmann, Nest e Daniel (2011), também atestou sua importância neste estudo. O fato de o autor apresentar uma metodologia consistente, clara e completa afeta positivamente a avaliação, conforme afirmam Ali, Young e Ali (1996), uma vez que será possível verificar se houve manipulação incorreta dos dados, provocando a rejeição do artigo, bem como sua adequação ao problema proposto.

Discussão dos Resultados.

O código analítico ‘discussão dos resultados’ possui três categorias: (a) clareza, (b) discussão dos resultados e (c) amplitude de interpretação. Esse critério foi citado por 6 dos 11 entrevistados.

A clareza relaciona-se ao fato de o artigo possuir um modelo ou quadro teórico, bem como a explicitação dos objetivos e problemas, que auxiliam na análise e interpretação dos dados. Na maioria dos critérios, os objetivos e problemas são avaliados, demonstrando que grande parte dos avaliadores se preocupa, de fato, com a competência de conteúdo. A discussão dos resultados pauta-se no cerne da discussão, em identificar os vieses e os pontos mais relevantes para a pesquisa. Na amplitude da interpretação, os avaliadores se preocupam em compreender como a análise foi feita, coadunando com o encadeamento entre problema e metodologia, bem como na relação dos dados encontrados com a teoria existente e na identificação do alcance dos resultados.

A importância de se discutirem os resultados, para um avaliador, está na verificação da capacidade de síntese do autor, na articulação do ponto de vista e na profundidade do diálogo dos resultados obtidos com a teoria existente. As falas dos entrevistados corroboram

Bornmann, Nest e Daniel (2011) no que tange à explicitação dos resultados obtidos e estavam presentes em 31 dos 46 casos analisados.

Teoria.

O código analítico ‘teoria’ possui uma categoria-chave, identificada nas entrevistas pela categoria: (a) teoria. Esse critério foi citado por 4 dos 11 entrevistados. A coerência do referencial teórico, o uso de construtos, o diálogo com os autores citados e a profundidade e atualização da teoria são pontos-chave de avaliação. Apesar de pouco mencionado nas entrevistas, esse item é relevante e aparece nos critérios de concepção e discussão dos resultados. Além disso, pode ter uma flexibilidade na avaliação, dependendo da classificação do periódico.

Uma possibilidade de a teoria não ter apresentado tanta importância para os entrevistados está, provavelmente, relacionada ao fato de os avaliadores já conhecerem os principais autores e as novidades da área. Assim, se preocuparam mais com a estruturação e o uso dos construtos, do que com as citações e os diálogos em si. “O referencial teórico geralmente a gente sendo da área a gente já conhece o referencial teórico, então a gente faz uma leitura mais diagonal, só para ver se está trazendo algum autor mais recente, alguma abordagem diferente. [Entrevistado I]”

A teoria é um critério de avaliação pouco importante, tanto constatado neste estudo quanto nos estudos de Bornmann, Nest e Daniel (2011), em que estava presente em 22 dos 46 casos. A principal preocupação do avaliador está em verificar a robustez do arcabouço teórico. A proposição de novas teorias, por exemplo, não é relevante para os avaliadores brasileiros, embora esteja presente no estudo ora citado.

Critérios Pessoais.

Um critério que não foi mencionado pelos estudos de Ali, Young e Ali (1996), Bornmann, Nest e Daniel (2008) e Ward, Hall e Scharam (1975), foi o que chamamos de critérios pessoais, que varia quanto à condição humana de cada avaliador, seu estilo e suas experiências (Botomé, 2011, Kruger, 2005, Omote, 2005, Smith, 2006). Foi mencionado direta e indiretamente por 4 dos 11 entrevistados.

A questão de ser um artigo empírico, de merecer ser publicado, de possuir uma provocação no título, da possibilidade de o avaliador aprender algo novo, do autor demonstrar ser preguiçoso são os códigos mencionados nas entrevistas e que foram levados em consideração no decorrer da avaliação. “Aí eu faço um critério, normalmente esse merece ser publicado, esse aqui eu acho que não merece ser publicado. E daí eu faço, a partir do momento que eu criei esse critério, justificar a minha opção nos quesitos. [Entrevistado A]”

Os julgamentos variam em grau, quantidade, profundidade e quantidade, e também de acordo com cada avaliador e sua experiência, corroborando Botomé (2011). O que faz um avaliador pensar que um artigo merece ser publicado e o outro não? Quais são os critérios de avaliação por trás da subjetividade? Embora muitos trabalhem com o apoio de um roteiro disponibilizado pela própria revista, esses não possuem pesos e a subjetividade volta a entrar em cena.

Ética.

O código analítico Ética possui duas categorias: (a) publicação múltipla e (b) rede de coautoria. Foi citado apenas por 1 entrevistado, quando mencionado o caso de autoplágio.

Esse trecho da entrevista reafirma o conceito de Leite, Kniess, Rodrigues e Maccari (2011) em que as redes são planejadas e conscientemente escolhidas, podendo-se chegar ao ponto de perda de legitimidade da academia, conforme afirmam Kirschbaum e Mascarenhas (2009). Ora, o mesmo avaliador critica a existência de plágio e autoplágio nas publicações recebidas e atua de forma considerada antiética em suas publicações.

Quanto ao Método de Avaliação

Os entrevistados, primeiramente, relataram seu método, ou modelo, de avaliação de artigos de periódico, que puderam ser descritos em três diferentes métodos de avaliação: (a) hipotético-dedutivo, (b) analítico-indutivo, e (c) *by the book* (conforme manual). Elucida-se, na Figura 3, cada um dos três modelos mencionados.

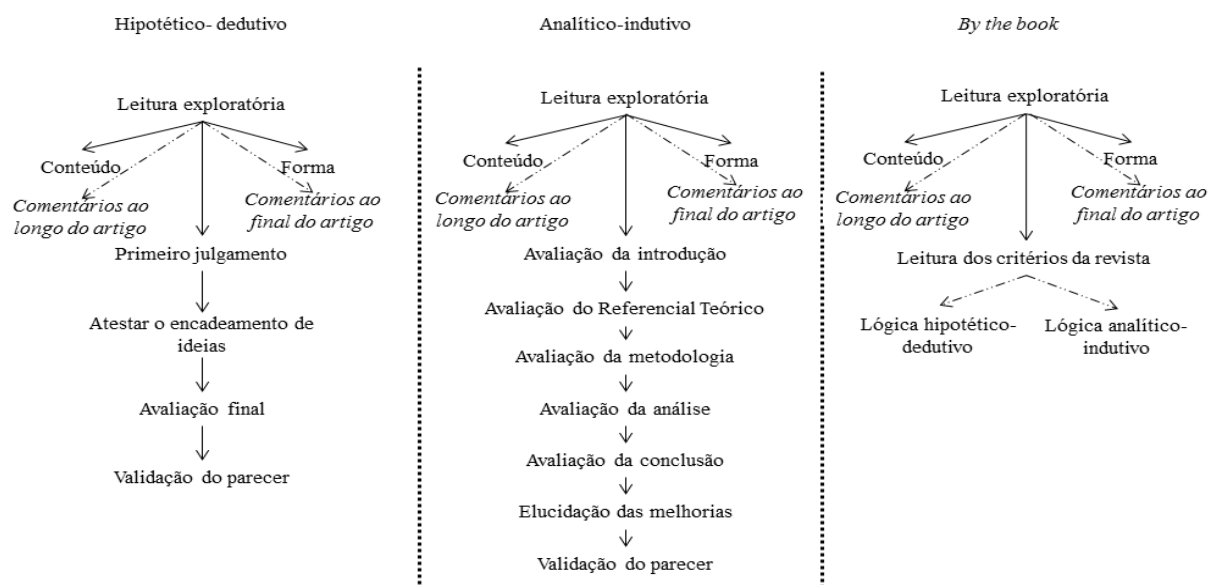


Figura 3. Modelos de métodos de avaliação de artigos

Fonte: Elaborado pelos autores, com dados da pesquisa.

Partindo-se do pressuposto de que o método hipotético-dedutivo é uma modalidade de raciocínio que particulariza uma determinada situação, partindo de um contexto universal ou geral, esse método parte de um problema com uma solução provisória, que fornecerá subsídios para a formulação de hipóteses e, a partir daí, será possível deduzir as consequências (Patrus, 2011). Foram encontrados, nas entrevistas, trechos que se assemelham a esse método.

Os entrevistados que utilizam esse modelo começam a avaliação por meio de uma primeira leitura, exploratória, a partir da qual elaboram um primeiro julgamento do artigo recebido (hipótese). Nessa primeira leitura, o avaliador também se preocupa em perceber o alinhamento entre os principais construtos do texto, a fim de promover o encadeamento das principais ideias. Após esse entendimento geral do artigo, os avaliadores partem de uma tipologia universal e integradora para uma avaliação, que, em muitos casos, é rápida e objetiva. A avaliação, que pode ser realizada no próprio artigo ou em formulário concedido pela revista, obedece inicialmente um critério geral, como: merece ou não ser publicado.

Já o modelo analítico-indutivo parte de informações acumuladas durante a leitura e é realizado passo-a-passo, com cada momento de avaliação bem definido, com início, meio e fim, buscando, também, apontar melhorias para os artigos avaliados. A indução é um método que parte do particular para o geral, que será sempre uma probabilidade (Patrus, 2011). O primeiro passo é realizar a leitura, observando forma e conteúdo, podendo o avaliador providenciar o registro das principais informações e evidências ao longo de sua

leitura ou fazê-lo posterior a essa. Conforme dito por dois entrevistados, “Normalmente eu faço uma leitura mais rápida e vou anotando o que salta aos olhos, problema de redação, problema de articulação das ideias, problema de uma informação mais fora de contexto. [Entrevistado G]” e, “Aí eu faço uma segunda leitura, e vou marcando os pontos fortes e pontos fracos ao longo do próprio texto. [Entrevistado H]”.

Feito isso, os avaliadores aferem todas as seções do artigo: introdução, referencial teórico, metodologia, análise dos dados e conclusões. Busca-se aqui o encadeamento das ideias dos autores. O avaliador, então, elucida possíveis melhorias para o artigo, valida seu parecer e providencia o encaminhamento ao solicitante.

No modelo de avaliação *by the book*, os avaliadores se preocupam em, principalmente, verificar se os artigos atendem aos requisitos exigidos pela revista, por isso o nome *by the book*, ou conforme o manual. No entanto o avaliador pode seguir a lógica hipotético-dedutiva da avaliação ou avaliar o artigo de forma analítico-dedutiva.

A presença de critérios pessoais de cada avaliador, no que se refere a seu método de avaliação, impacta o julgamento final. Mediante experiência e outros acontecimentos, o método pode ser alterado sem aviso prévio e o autor jamais saberá como seu trabalho definitivamente será avaliado. E, nesta etapa da entrevista, foi perguntado aos avaliadores se o extrato do periódico influenciaria na adoção de um modelo diferente do que foi relatado. Sete dos entrevistado disseram que sim e, com a finalidade de ilustrar, dois comentários foram selecionados: “não consigo não, não tem jeito, não adianta, não faço não” [Entrevistado C] e, “Se uma revista tem uma classificação maior, eu tendo a exigir um pouco mais do artigo (...) no que ele está contribuindo, a inovação do artigo” [Entrevistado D].

Considerações Finais

O objetivo deste estudo foi alcançado e ao final foi possível analisar, de forma satisfatória, o processo de revisão por pares sob a ótica dos avaliadores de artigos de periódicos científicos. Nesse contexto, foi possível eleger as principais contribuições do estudo: (a) densa relação dos critérios de avaliação obtidos com os estudos de Ali, Young e Ali (1996), Bornmann, Nest e Daniel (2008) e Ward, Hall e Scharam (1975), e (b) elucidação de três modelos-chave para o processo de avaliação.

Os critérios de avaliação mencionados durante as entrevistas estavam de acordo com os estudos apontados neste trabalho, salvo duas exceções: reputação e fatores pessoais. A

reputação não é um critério utilizado pelos avaliadores brasileiros como essencial em uma avaliação, o que contraria a literatura internacional em que o papel de editor e de avaliador é carregado de maior prestígio. Já os fatores pessoais de avaliação foram achados empíricos encontrados a partir das entrevistas, sem correspondentes na literatura estudada.

Os modelos de avaliação encontrados na análise das entrevistas apontam para o estilo de cada avaliador. Alguns preferem partir do geral para o particular, outros somam as observações para induzir uma conclusão enquanto outros seguem o que preconiza o manual. Parece claro que esses modelos não são excludentes, podendo o avaliador se valer deles de forma dialética. Emitir um parecer ou formalizar a avaliação exige, do revisor, clareza a respeito do seu papel no sistema e do assunto avaliado, passando o revisor a ter responsabilidade no trabalho de outro (Pinho, 2005).

Grande parte das críticas ao sistema de avaliação apresentadas na literatura foi encontrada na pesquisa empírica com os entrevistados. As críticas mais comuns estão relacionadas ao tempo, sobrecarga de trabalho, subjetividade das avaliações, falta de reconhecimento do trabalho do avaliador, falta de padrão de qualidade na seleção dos avaliadores, necessidade de ter mais avaliadores para criar a rotatividade, parcialidade dos editores e abuso de poder.

As críticas relacionadas a preconceito, competição e aquelas sobre o custo das avaliações não foram mencionadas pelos entrevistados. O fato de a crítica ao preconceito não aparecer reafirma a ausência do critério de reputação na discussão anterior, ou seja, o avaliador brasileiro não julga um artigo por sua origem, seja da IES, seja do autor. Já quanto ao custo, no Brasil não há a política de remuneração financeira para os avaliadores e em grande parte também não há a cobrança para publicação ao autor. Os elogios ao sigilo e transparência do sistema foram os únicos mencionados e por apenas um entrevistado.

Do ponto de vista prático, é fundamental estudar o sistema de revisão por pares e como ele vem se modificando ao longo do tempo e alterando as atividades dos professores participantes. Como sugestão para pesquisas futuras, recomenda-se (a) o aprofundamento de cada um dos achados dessa pesquisa, em termos de pesquisa descritiva, seja ela qualitativa ou quantitativa, (b) a ampliação do objeto de estudo, por meio da seleção de entrevistados em outros estados brasileiros ou no que se refere à diversificação da amostra (revisores com experiência inferior a 6 anos, revisores de periódicos com estratos inferiores a B4, coordenadores de cursos, editores, entre outros) e, (c) a compreensão se os avaliadores

diferem seus critérios e modelos de avaliação ao avaliar trabalhos acadêmicos de periódicos com estratos diferentes.

Este estudo possui contribuição teórica, ao lançar luz sobre o trabalho anônimo de pesquisadores de todo o mundo responsáveis pela avaliação de artigos, projetos de pesquisa, cursos ou programas de pós-graduação, além de dedicar seu tempo às atividades acadêmicas e administrativas (Shigaki, 2014). E, contribuição prática, pois além da apresentação dos critérios de avaliação, três modelos-chave foram construídos a partir das entrevistas realizadas, que são utilizados pelos entrevistados para a avaliação. Isso porque, de acordo com Botomé (2011), ao emitir um parecer exige-se clareza, critérios, procedimentos e contexto.

Os limites deste estudo estão relacionados às decisões metodológicas dos autores (Calado Dias, Patrus, & Magalhães, 2011). Em relação à seleção dos entrevistados e à estratégia metodológica, eles foram três: (a) a seleção dos entrevistados foi realizada por meio das informações presentes nos sites de periódicos classificados como A pela Capes e muitos deles estavam desatualizados à época, (b) todos os entrevistados são residentes de uma mesma região e (c) possibilidade de julgamento particular e subjetivo dos autores para categorização dos dados (Bauer, 2002), o que foi minimizado com a contribuição e participação de outros pesquisadores. Para estudos futuros, os autores sugerem que a pesquisa seja replicada a avaliadores de outras regiões do Brasil, de outros extratos de periódicos, e também sugerem que o roteiro de entrevista possa ser adaptado a um questionário estruturado a ser enviado a um maior número de avaliadores.

Em síntese, emitir um parecer ou formalizar uma avaliação, para Botomé (2011), exige clareza, critérios, procedimentos e contexto de realização de um trabalho. Ao avaliar um artigo para um periódico, o avaliador passa a ter responsabilidade por sua contribuição no trabalho do outro (Pinho, 2005). Poder-se-ia dizer que, “o que se procura são, na verdade, mecanismos eficientes de controle de qualidade das publicações” (Omote, 2005, p. 331). Em um contexto em que publicar se torna a principal meta de programas de excelência, a reflexão sobre o processo de avaliação desses artigos se torna fundamental. E, muitas vezes, inexplicavelmente esquecida.

Referências

- Aaboen, L., & Frida Lind, A. D. (2012). Capturing processes in longitudinal multiple case studies. *Industrial Marketing Management*, 41(2), 235-246.
- Ali, S. N., Young, H.C., & Ali, N.M. (1996). Determining the quality of publications and research for tenure or promotion decision. *Library review*, 45(1), 39-53.
- Bardin, L. (2011). *Análise de conteúdo*. São Paulo: Edições 70.
- Bauer, M. (2002). Análise de conteúdo clássica: uma revisão. In M. Bauer, & G. Gaskell. *Pesquisa qualitativa com texto, imagem e som*. 2 ed.. Petrópolis: Ed. Vozes.
- Bloom, T. (2006). *Systems: Online frontiers of the peer-reviewed literature*. Recuperado de <http://www.nature.com/nature/peerreview/debate/nature05030.html>
- Borba, J.A., & Murcia, F. (2006). Possibilidades de Inserção da Pesquisa Contábil Brasileira no Cenário Internacional: Uma Proposta de Avaliação dos Periódicos Científicos de Contabilidade e Auditoria. In *Anais do Encontro Nacional de Pós-graduação e Pesquisa em Administração*, Salvador, BA, Brasil.
- Bornmann, L., Nast, I., & Daniel, H-D. (2008). Do editors and referees look for sign of scientific misconduct when reviewing manuscript? A quantitative content analysis of studies that examined review criteria and reasons for accepting and rejecting manuscripts for publication. *Scientometrics*, 77(3), 415-432.
- Botomé, S.P. (2011). Avaliação entre pares na ciência e na academia: aspectos clandestinos de um julgamento nem sempre científico, acadêmico ou de avaliação. *Psicologia USP*, 22(2).
- Calado Dias, S., Patrus, R., & Magalhães, Y.T. (2011). Quem ensina um professor a ser orientador? Proposta de um modelo de orientação de monografias, dissertações e teses. *Administração: ensino e pesquisa*, 12(4), 697-721.
- Campanario, J.M. (2002). El sistema de revisión por expertos (*peer review*): muchos problemas y pocas soluciones. *Revista Española de Documentación Científica*, 25(3), 166-184.

- Coupé, T. (2010). Peer review versus citations – an analysis of best paper prizes. *Kyiv School of Economics*, 42(1), 295-301.
- Davyt, A., & Velho, L. (2000). A avaliação da ciência e a avaliação por pares: passado e presente. Como será o futuro? *História, Ciência e saúde*, 7(1), 93-116.
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, 14(4), 532-550.
- Fitzpatrick, K. (2011). *Planned Obsolescence: publishing, technology, and the future of the academy*. New York: NYU Press Short.
- Freitas, M. (1998). Avaliação da produção científica: considerações sobre alguns critérios. *Psicologia Escolar e Educacional*, 2(3), 211-228.
- Gaskell, G. (2002). Entrevistas individuais e Grupais. In M. Bauer, & G. Gaskell. *Pesquisa qualitativa com texto, imagem e som*. 2 ed.. Petrópolis: Ed. Vozes.
- Gibbs, G. (2009). *Análise de dados qualitativos*. Porto Alegre: Artmed.
- Godoi, C. K., & Balsini, C. P. V. (2004). A metodologia qualitativa em estudos organizacionais. In: *Anais de Encontros de Estudos Organizacionais*. Atibaia, RJ, Brasil.
- Godoi, C. K., Bandeira-de-Mello, R., & Silva, A.B. (2007). *Pesquisa qualitativa em estudos organizacionais: paradigmas, estratégias e métodos*. São Paulo: Saraiva.
- Godoy, A. R. (1995). A pesquisa qualitativa e sua utilização em administração de empresas. *Revista de Administração de Empresas*, 35(4), 65-71.
- Gondim, S. (2004). A face oculta do parecerista: discussões éticas sobre o processo de avaliação de mérito de trabalhos científicos. *Organização & Sociedade*, 11(31), 195-199.
- Hames, I. (2012). *O estado atual de revisão por pares: críticas, desafios e inovações*. Ata de Reunião da Associação Brasileira de Editores Científicos.
- Hébert, P. C. (2007). Examen par les pairs ou lobby politique? Le choix est Clair. *CMAJ*, 176(10), 1.391.

- Jennings, C. G. (2006). *Quality and value: The true purpose of peer review?* Nature. International weekly journal of science. London, UK: Nature Publishing Group.
- Kirschbaum, C., & Mascarenhas, A. O. (2009). Nos limites da autonomia: reflexões sobre práticas de blind review e editoria de revistas científicas em administração no Brasil. *Revista de Administração de Empresas-eletrônica*, 8(1).
- Kostoff, R. N. (1996). Performance measures for government-sponsored research: overview and background. *Scientometrics*, 36(3), 281-292.
- Kruger, H. (2005). Avaliação de trabalhos científicos. *Organização & Sociedade*, 12(33), 179-182.
- Lee, K., & Bero, L. (2006). What authors, editors and reviewers should do to improve peer review. Nature. *International weekly journal of science*. London, UK: Nature Publishing Group.
- Leite, N., Kniess, C., Rodrigues, A., & Maccari, E. (2011). A ética na produção, orientação, submissão, avaliação e publicação científica: quem assume a responsabilidade? In *Anais do Encontro de Ensino e Pesquisa em Administração e Contabilidade*, João Pessoa, Paraíba, Brasil.
- Media, M. M. (2008). Peer review in scholarly journals: Perspective of the scholarly community – an international study. *Publishing research consortium*. Recuperado de <http://www.publishingresearch.org.uk/PeerReview.htm>
- Meneghini, R., & Fonseca, L. (1990). Índices alternativos de avaliação da produção científica em bioquímica no Brasil. *Ciência e Cultura*, 42(9), 629-646.
- Meyer, C. B. (2001). A Case in Case Study Methodology. *Field Methods*, 3(4), 329–352.
- Miller, C. (2006). Peer review in the organizational and management sciences: prevalence and effects of reviewer hostility, bias, and dissensus. *Academy of Management Journal*, 49 (3), 425-431.
- Mulligan, A., Hall, L., & Raphael, E. (2013). Peer Review in a Changing World: An International Study Measuring the Attitudes of Researchers. *Journal of the American society for information science and technology*. 64(1).

- Nature. (2006). *Three cheers for peers* [439, Editorial].
- Omote, S. (2005). Revisão por pares na revista brasileira de educação especial. *Revista Brasileira de Educação Especial*, 11(3), 323-324.
- Paavilainen, R., & Asted, P. (1997). Self-reported family health and well being after early discharge from maternity hospital: a phenomenological study. *Journal of Advanced Nursing*, 26(2), 266-272.
- Pagani, F. (2002). L'examen par les pairs: un instrument de cooperation et de changement: *Analyse d'une méthode de travail de l'OCDE*.
- Partis, M. (2003). Hope in homeless people: a phenomenological study. *Primary Health Care Research and Development*, 4(1), 9-19.
- Patrus, R. (2011). Uma Exegese de Alguns Métodos do Pensamento Filosófico e Científico: Propondo uma Metodologia Existencial. *Administração: Ensino e Pesquisa*, Rio de Janeiro, 12(2), 191-218.
- Patrus, R., Dantas, D., & Shigaki, H. (2013). Pesquisar é Preciso, Publicar não é Preciso: História e Controvérsias sobre a Avaliação Por Pares. In *Anais do Encontro de Ensino e Pesquisa em Administração e Contabilidade*, Brasília, DF, Brasil.
- Phelan, S., Ferreira, M., & Salvador, R. (2002). The first twenty years of the Strategic Management Journal: 1980-1999. *Strategic Management Journal*, 23, 1161-1168.
- Pinho, J. A. G. (2005). Brevíssimo manual do editor: considerações obre submissão e avaliação de artigos, o papel dos pareceristas e do editor de revistas científicas. *Organizações & Sociedade*, 12(34), 169-173.
- Pinto, A. C., & Andrade, J. B. (1999). Fator de impacto de revistas científicas: qual o significado deste parâmetro. *Química nova*, 22(3), 448-453.
- Pessanha, C. (1998). Critérios editoriais de avaliação científica: notas para discussão. *Revista Ciências da Informação*, 27(2), 226-229.

- Poupart, J., Deslauriers, J. P., Groulx, L-H., Laperrière, A., Mayer, R., & Pires, A.P. (2002). *A pesquisa qualitativa: enfoque epistemológicos e metodológicos*. 3 ed.. Petrópolis: Vozes.
- Rosa, A. M., & Chachamovich, J. (2008). Editorial a convite: O que faz a excelência de uma revista científica. *Revista de Psiquiatria*, 25(2). Recuperado de www.scielo.br/pdf/rprs/v25n2aO2.pdf
- Selli, L., & Garrafa, V. (2006). Solidariedade crítica e voluntariado orgânico: outra possibilidade de intervenção societária. *História, ciência e saúde*, 13(2), 239-251.
- Serra, F. A. R., Fiates, G. G., & Ferreira, M. P. (2008). Publicar é difícil ou faltam competências? O desafio de pesquisar e publicar em revistas científicas na visão de editores e revisores internacionais. *Revista de Administração Mackenzie*, 9(4), 32-55.
- Serva, M. (1997). A Racionalidade Substantiva Demonstrada na Prática Administrativa. *Revista de Administração de Empresas*, 37(2), 18-30.
- Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. *Journal of the royal society of medicine*, 99(4), 178-182.
- Spink, P. K. (2003). A perda, a descoberta e a transformação de uma tradição de trabalho: a teoria sociotécnica nos dias de hoje. *Organizações & Sociedade*, 10(28), 337-343.
- Stumpf, I. (2008). Avaliação pelos pares nas revistas de comunicação: visão dos editores, autores e avaliadores. *Perspectivas em Ciência da Informação*, 13(1), 18-32.
- Thiry-Cherques, H. R. (2005). Comentando: “A face oculta do parecerista: discussões éticas sobre o processo de avaliação de mérito de trabalhos científicos”. *Organizações & Sociedade*, 12(32), 169-172.
- Ward, A. W., Hall, B. W., & Scharam, C. F. (1975). Evaluation of published educational research: a national survey. *American educational research journal*, 12(2), 109-128.
- Ware, M. (2008). *Peer review: benefits, perceptions and alternatives*. London, UK: Publishing research consortium. 2008.

Submissão: 08/08/2016

Segunda versão: 04/11/2016

Aceite: 04/11/2016