

CITAÇÕES E INDICADORES DE IMPACTO NA AVALIAÇÃO DE REVISTAS

CITATION AND IMPACT INDICATORS FOR JOURNAL EVALUATION

Sônia Elisa Caregnato¹
Samile Andréa de Souza Vanz²

RESUMO

Este artigo apresenta uma revisão de literatura sobre o uso de citações para avaliação de revistas científicas. Considerando que a CAPES, atualmente, está em processo de reformulação das suas políticas e instrumentos de avaliação e que inúmeras críticas ao Qualis têm sido feitas por pesquisadores e gestores, este artigo tem por objetivo discutir a avaliação de periódicos científicos à luz das concepções do papel das citações na atividade científica e dos indicadores bibliométricos delas derivados, especialmente Fator de Impacto, *CiteScore* e Índice h do Google. Apresenta-se o histórico e metodologia utilizada para cálculo dos três indicadores, além da fórmula utilizada pela CAPES para utilização dos mesmos na avaliação das revistas. Conclui que os indicadores de citação usados internacionalmente para avaliação de revistas possuem amparo em teorias e metodologias desenvolvidas há anos, mas que eles precisam ser empregados com cuidado e conhecimento.

Palavras-chave: Avaliação. Revistas científicas. Citações. Indicadores de impacto. Qualis.

ABSTRACT

The paper presents a literature review on the use of citations for journal evaluation. Considering that Brazilian CAPES is currently in the process of reformulating its evaluation policies and instruments and that numerous criticisms of Qualis have been made by researchers and managers over the years, we aim to discuss journal evaluation considering the role of citations in scientific practice and the bibliometric indicators derived from them, especially Impact Factor, CiteScore and h Index from Google. The origins and methodology employed to calculate the three indicators are presented, in addition to the formula used by CAPES for journals evaluation. It concludes that citation indicators for science evaluation are supported by concepts and methodologies about citation practices developed over the years, but that they should be employed carefully and consciously.

Keywords: Journal evaluation. Citation. Impact indicators. Qualis.

Artigo submetido em 26/10/2020 e aceito para publicação em 16/11/2020

1 Docente permanente dos Programas de Pós-Graduação em Comunicação e em Ciência da Informação. Universidade Federal do Rio Grande do Sul, Brasil. Bolsista de produtividade do CNPq. ORCID: <https://orcid.org/0000-0002-5676-2763>. E-mail: sonia.caregnato@ufrgs.br

2 Docente permanente dos Programas de Pós-Graduação em Comunicação e em Ciência da Informação. Universidade Federal do Rio Grande do Sul, Brasil. Bolsista de produtividade do CNPq. ORCID: <https://orcid.org/0000-0003-0549-4567>. E-mail: samilevanz@terra.com.br

1 INTRODUÇÃO

A concepção de que os resultados de uma pesquisa precisam ser examinados pelos demais pesquisadores não é nova. Antes mesmo do surgimento dos primeiros periódicos científicos, quando os resultados de experimentos eram compartilhados por meio de cartas enviadas pelos cientistas aos seus pares, o trabalho de um pesquisador era lido, considerado e criticado pelos colegas a quem as cartas eram destinadas. Segundo Spier (2002), a proposição das ideias empiristas de Bacon motivaram o surgimento, primeiramente, de encontros de estudiosos para debater visões e opiniões sobre a ciência emergente e, posteriormente, de sociedades científicas que, além do debate, passaram a publicar os primeiros periódicos, em que o editor e seus auxiliares selecionavam o que neles era publicado.

Somente no século seguinte à primeira publicação do *Philosophical Transactions* que os manuscritos passaram a ser enviados para inspeção de membros das sociedades, os quais faziam recomendações ao editor antes da publicação (SPIER, 2002). Esse sistema de arbitragem foi se estabelecendo, consistindo no envio de artigos para avaliação de especialistas na mesma área, mas sem a configuração que conhecemos hoje. Para Baldwin (2018), foi só na segunda metade do século 20 que o termo revisão por pares, juntamente com a noção de que a pesquisa precisa ser avaliada para obter legitimidade, tomou forma.

Ao mesmo tempo, a compreensão de que a ciência é fator primordial no desenvolvimento das nações, na melhoria da qualidade de vida das pessoas e no bom funcionamento dos ecossistemas naturais e sociais, que se desenvolveu especialmente no período posterior a Segunda Guerra Mundial, fez com que a maioria dos governos investissem quantias consideráveis de recursos públicos em pesquisa e desenvolvimento. Juntamente com os recursos, vieram também cobranças no sentido de se inspecionar os resultados desses gastos públicos, tanto oriundas das agências financiadoras como da sociedade mais ampla. Tal movimento, entende-se, é positivo pois tem o potencial de estimular um sistema de pesquisa mais aberto, transparente, responsável e em conformidade com as demandas da sociedade.

Moed (2017) aponta que nos últimos dez anos, na maioria dos países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), aumentou a preocupação com o financiamento de pesquisas que sejam efetivas e eficientes. Assim, segundo ele, a fim de otimizar os recursos, orientar a concessão de apoio, racionalizar as ações organizacionais, reestruturar a pesquisa em determinados

campos ou aumentar a produtividade em pesquisa, esses governos passaram a estimular ou impor processos avaliativos.

No Brasil, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) realiza a Avaliação do Sistema Nacional de Pós-Graduação, na forma como foi estabelecida a partir de 1998, por meio da Diretoria de Avaliação e com a participação de consultores *ad-hoc*. Essa avaliação envolve, entre várias outras estratégias, a composição de uma lista estratificada dos periódicos em que os professores dos programas de pós-graduação (PPGs) publicaram artigos, ou seja, o sistema Qualis.

Considerando que a CAPES, atualmente, está em processo de reformulação das suas políticas e instrumentos de avaliação e que inúmeras críticas ao Qualis têm sido feitas por pesquisadores e gestores, este artigo tem por objetivo discutir a avaliação de periódicos científicos à luz das concepções do papel das citações na atividade científica e dos indicadores bibliométricos de impacto delas derivados, considerando a proposta do novo Qualis Referência.

2 AS CITAÇÕES NA AVALIAÇÃO DA CIÊNCIA

A avaliação da ciência em diferentes campos do conhecimento e regiões do mundo é fortemente baseada na aplicação de indicadores cientométricos. Leydesdorff (2005) registra o início dessa aplicação nos anos 1960 e 1970, primeiramente nos Estados Unidos e depois na Europa. Antes disso, explica ele, a avaliação da pesquisa se restringia a duas instâncias: o sistema de avaliação por pares e a avaliação da atividade científica no nível macro das nações por meio da utilização de indicadores econômicos. A informação e as publicações científicas expandiram-se de maneira quase incontrolável após a segunda guerra mundial, o que fez com que instrumentos de organização e recuperação da informação ganhassem força. É neste contexto que surge o *Science Citation Index (SCI)* e em que ele se fortalece como provedor de um padrão objetivo de mensuração da ciência, que considera a frequência média de citações recebida por periódicos científicos, ou seja, o controverso Fator de Impacto.

Indicadores para avaliar as atividades de pesquisa podem ser categorizados em quatro tipos, segundo Moed (2015): input, processo, output (produção) e impacto. Input refere-se a indicadores utilizados para avaliar os recursos financeiros, humanos e de infraestrutura investidos em pesquisa, ou seja, as condições prévias para fomentar a pesquisa; indicadores de processo referem-se à forma como a pesquisa é conduzida, por exemplo, tipos de colaboração estabelecidas, grau de mobilidade de pesquisadores e eficiência, que relaciona input e output; indicadores de output dizem respeito

ao conhecimento resultante da atividade de pesquisa, ou mais especificamente aos seus produtos, corriqueiramente medido por meio de número de publicações e patentes; indicadores de impacto, por sua vez, buscam medir a forma com os conhecimentos produzidos afetam a comunidade científica e a sociedade em geral.

A distinção clara entre impacto científico e impacto social da pesquisa, esse último envolvendo também aspectos culturais, ambientais e econômicos relacionados aos resultados, precisa ser explicitada. Bornmann (2016) mostra que historicamente indicadores cientométricos mediam o impacto do conhecimento científico na academia exclusivamente. O argumento é o de que a sociedade invariavelmente acaba se favorecendo da ciência feita com qualidade, mas essa qualidade deve ser avaliada por outros pesquisadores, com pelo menos o mesmo conhecimento sobre o tema. Desde os anos 1990, segundo o autor, essa concepção vem mudando e agora a ciência precisa mostrar seu valor para além dos limites internos de sua competência.

Uma tentativa de levantar o impacto social da pesquisa científica tem sido na direção de buscar métricas alternativas, ou altmetrics, para sintetizar. Priem e Hemminger (2010) propuseram o uso de indicadores de menções a artigos em serviços da web 2.0, como o twitter, gerenciadores de referência, redes sociais, etc. No entanto, o uso de indicadores altmétricos como forma de medir o impacto social da ciência não está consolidado. Barnes (2015) argumenta que a altmetria mede o consumo imediato de informações nas mídias sociais e não a contribuição de longo prazo de uma pesquisa para a sociedade. Além disso, a menção a artigos científicos nas mídias sociais pode se dar simplesmente por eles apresentarem temas ou títulos inusitados, divertidos ou peculiares e não pelo seu conteúdo. Alguns estudos altmétricos sugerem que a maioria dos artigos circula entre comunidades específicas de pesquisadores, e não entre o público em geral (ALPERIN *et al.*, 2019).

Em decorrência das dificuldades de se medir o efeito que a pesquisa publicamente financiada tem na melhoria da qualidade de vida em sociedade, as práticas avaliativas atuais focam nos fenômenos que acontecem dentro e entre as comunidades científicas. Essas práticas se revelam principalmente nos indicadores de produção e impacto, seguindo a classificação de Moed (2015), que são gerados a partir do escrutínio das publicações e de suas citações. A partir deles, a avaliação institucional infere as ações de pesquisadores, revistas, instituições e países.

Se a prática de citar fontes de informação utilizadas na elaboração de um trabalho científico já está estabelecida como norma nas diferentes áreas do conhecimento, há ainda inúmeras controvérsias acerca do que as citações representam e, portanto, o que se está medindo quando elas são computadas.

Em se tratando de processos avaliativos, o que se busca é identificar a qualidade, um conceito multidimensional que envolve diversas características, como validade, originalidade, valor científico e valor social (ASKNES *et al.*, 2019), portanto muito difícil de ser operacionalizado.

Então, impacto tem sido o termo mais frequentemente empregado na literatura (por exemplo, Aksnes *et al.* (2019), Waltman (2016) e Wouters *et al.* (2015)) para designar a influência de um produto de pesquisa sobre outro, seja no sentido de confirmar ou de refutar os resultados. Wouters e colegas explicam:

A visão comum de que a taxa de citação é uma medida de qualidade científica não é apoiada pela comunidade de especialistas bibliométricos. Os bibliometristas geralmente veem a taxa de citação como uma representação de impacto científico ou de impacto nas comunidades científicas relevantes. Esta é somente uma das dimensões da qualidade científica ou acadêmica. A qualidade é vista como um conceito multidimensional que não pode ser capturado por nenhum indicador (WOUTERS *et al.*, 2015, p.viii, tradução nossa).

Embora em número menor, há também argumentos na defesa de que como as citações certificam o uso/impacto de determinada pesquisa no avanço do conhecimento, em que a publicação é o resultado desse avanço, então impacto de citação é sinônimo de qualidade científica (ABRAMO, 2018). Nesta concepção, as citações negativas e outras anomalias são exceções que se reduzem quando os estudos escalam. De qualquer forma, Abramo (2018) enfatiza que toda sua argumentação diz respeito ao artigo, e mais especificamente ao conhecimento embutido nele, que segundo o autor é a unidade básica de avaliação do impacto, e não ao periódico.

Outra forma de interpretar a função que as citações desempenham na avaliação da ciência é distinguir dois aspectos do seu impacto: popularidade e prestígio. Bollen *et al.* (2006) mostram que o apelo popular de uma obra e o reconhecimento que ela recebe de notáveis não são fenômenos equivalentes, assim como a popularidade e o prestígio científico também não o são. No sentido de caracterizar essa diferença nas citações a autores de artigos, Ding e Cronin (2011) definem popularidade como o número de vezes em que um autor foi citado e prestígio como o número de vezes em que o autor foi citado em artigos altamente citados. Em relação a periódicos, Bollen *et al.* (2006) operacionalizaram a popularidade pelo ranking do fator de impacto e o prestígio pela atribuição de pesos aos trabalhos citantes, utilizando-se, para isso, do algoritmo *PageRank* do Google. Ambos os trabalhos mostram a validade de se utilizar uma abordagem que leve em conta as diferenças entre as citações recebidas.

A produção de indicadores de citação mais precisos demanda a busca por evidências empíricas acerca do comportamento de citação no contexto das práticas comunicativas científicas. Nesse sentido, a literatura oferece duas principais perspectivas antagônicas para embasar a interpretação

sobre o fenômeno: a teoria normativa e a teoria socioconstrutivista. A primeira deriva da sociologia da ciência de Merton, em que citações representam o reconhecimento das influências intelectuais sobre o trabalho científico, e a segunda parte de sociólogos da ciência como Knorr-Cetina, para quem o conhecimento é socialmente construído por meio da manipulação e do uso de dispositivos retóricos (BORNMANN; DANIEL, 2008). No extremo, esses dois pontos de vista podem ser tachados como ingênuo e relativista, respectivamente, não sendo compatíveis com os processos de avaliação por meio de citações, por desacreditá-las como representativas da atividade científica.

O fenômeno da citação apresenta muitas mais nuances do que essa polarização poderia descrever. Por exemplo, Davis (2009) mostra que o efeito Mateus na ciência não pode ser explicado por nenhuma das duas teorias sozinhas, as quais ele denomina modelo da recompensa e modelo da persuasão. Portanto, novos modelos interpretativos que integrem as duas perspectivas são necessários. Nesta linha, Silveira e Caregnato (2018) propõem uma concepção que de ênfase na explicitação das influências dos múltiplos contextos nos processos de produção e citação dos pesquisadores. De fato, esses modelos interpretativos necessitam mostrar que o ato de citar textos acadêmicos pode se converter em um instrumento legítimo para ser usado no processo de avaliação da atividade científica e, portanto, também na avaliação de revistas.

Em síntese, delinea-se na literatura a compreensão de que os indicadores métricos não medem a qualidade da pesquisa, nem mesmo seu impactos na sociedade, mas o impacto do conhecimento publicado sobre o corpo de conhecimento científico; que as citações são (por enquanto) a principal forma de avaliar esse impacto, mas que elas próprias apontam para distintos graus de conformação ao corpo de conhecimento; que os modelos avaliativos que se baseiam nesses indicadores precisam considerar a complexidade do fenômeno da citação e esclarecer seus pressupostos. Dessa forma, pode-se avançar no sentido de uma avaliação da pesquisa científica em conformidade com as práticas das diferentes comunidades.

3 AVALIAÇÃO DE REVISTAS CIENTÍFICAS

O crescimento do volume de revistas científicas levou à necessidade de instituir instrumentos para seleção daquelas mais relevantes para diferentes finalidades, que podem ser tanto a aquisição em bibliotecas, como a seleção de veículo para publicação, ou para que pesquisadores acompanhem os desenvolvimentos em suas áreas de conhecimento. O alto custo dos periódicos científicos nos anos

1980 demandou a criação de estratégias de avaliação por parte das bibliotecas (MUELLER, 2006), no que tange o desenvolvimento de coleções. Para algumas áreas a avaliação de revistas representa uma tentativa de assegurar a qualidade ao processo de desenvolvimento e aperfeiçoamento da ciência, garantindo que o que está sendo produzido e veiculado é relevante e confiável (COSTA; YAMAMOTO, 2008).

A avaliação de revistas pode ser feita de forma qualitativa, por meio do julgamento pela própria comunidade científica. No entanto, assim como os métodos qualitativos, a avaliação por pares também recebe críticas, principalmente relacionadas à subjetividade do processo, ao viés dos avaliadores e à falta de transparência, entre outras (WOUTERS *et al.*, 2015). A análise das citações recebidas, por sua vez, também constitui um método que reflete a relevância do artigo (VANZ; CAREGNATO, 2003) e tem, dependendo contexto, maior facilidade de aplicação. Conforme Cronin (1984, p. 25, tradução nossa): “as citações são pegadas congeladas no cenário de realizações acadêmicas; pegadas que testemunham a passagem de ideias”.

Os indicadores de citação ampliaram seus domínios no contexto de grande crescimento das revistas e dificuldade de avaliação qualitativa. Martínez-Ávila (2019) reúne os indicadores de citação mais famosos para revistas científicas: Fator de Impacto, mensurado para revistas incluídas no JCR da Clarivate Analytics; SJR, desenvolvido pela SCImago com dados Scopus; Eigenfactor, desenvolvido por Eigenfactor.org com dados Web of Science; SNIP, desenvolvido pelo Centre for Science and Technology Studies da Universidade de Leiden com dados Scopus; *CiteScore*, desenvolvido pela Elsevier com dados Scopus; e o Índice h, que pode ser calculado a partir de qualquer base de dados de citação, mas comumente está associado ao Google Scholar. Neste artigo, optamos por abordar o Fator de Impacto, o *CiteScore* e o Índice h, que são indicadores bibliométricos a serem utilizados no Qualis Referência.

O Fator de Impacto (FI), um dos indicadores bibliométricos mais antigos, foi proposto por Eugene Garfield em 1955 junto com os índices de citação (GARFIELD, 1955), no âmbito do antigo Institute for Scientific Information (ISI) na Filadélfia. Segundo ele, os índices constituíram-se em um sistema bibliográfico para a literatura científica que poderia eliminar a citação de dados fraudulentos, incompletos ou obsoletos, possibilitando aos pesquisadores conhecer críticas ao seu trabalho por meio do mapeamento da rede de citações na ciência. Nas palavras de Garfield (1955, p. 109, tradução nossa):

[...] o sistema forneceria uma lista completa para a publicação abrangida, de todos os artigos originais que se referiram ao artigo em questão. Isso seria claramente útil na pesquisa histórica, quando se está tentando avaliar a importância de uma obra particular e seu impacto na literatura

e no pensamento do período. Esse “fator de impacto” pode ser muito mais indicativo do que uma contagem absoluta do número de publicações de um cientista.

Os índices de citação de Garfield basearam-se nos índices de Shepard para tribunais e casos jurídicos, e tinham como preocupação, no princípio, o controle bibliográfico. A ideia era diminuir a distância temática e terminológica entre autores e aqueles que buscam informação. A falta de uma cobertura completa não era entendida como problema, poderia ser vista como vantagem, considerando que poderia ser limitada à lista de periódicos cobertos por um dos principais serviços de indexação (GARFIELD, 1955). O lançamento dos índices gerou polêmica pois colocou em destaque uma prática (de citação) considerada até então de menor importância (relevada a segundo plano) pelos pesquisadores. O mapeamento dos links de citação entre autores revelou as escolhas de citações, possibilitando novos horizontes na sociologia da ciência (CRONIN, 1984).

Com o passar dos anos o indicador começou a ser aplicado na avaliação de periódicos e pesquisadores. Atualmente, o Fator de Impacto é operado pela empresa Clarivates, calculado a partir de revistas indexadas pela base de dados Web of Science (WoS) e publicado anualmente no Journal Citation Report (JCR). O FI de um periódico é calculado dividindo-se o número de citações recebidas por itens publicados naquele periódico nos dois anos anteriores pelo número de artigos publicados nos dois anos anteriores (WALTMAN, 2016). Dessa forma um fator de impacto 1,0 indica que os artigos publicados nos últimos dois anos foram citados, em média, uma vez cada. Considera-se possível comparar a importância relativa de um determinado periódico com base em seu FI.

O *CiteScore* é a medida do impacto de citações de periódicos lançada em 2016 pela base de dados Scopus, pertencente à empresa Elsevier. Ele é calculado a partir do número de citações recebidas por um periódico em um ano para documentos publicados nos três anos anteriores, dividido pelo número de documentos publicados nos mesmos três anos (JAMES *et al.*, 2019). Assim como o FI, o cálculo do *CiteScore* baseia-se nos dados compilados a partir dos periódicos indexados em uma base de dados bibliográfica específica, portanto é influenciado pelas decisões editoriais dessa base. Como a Scopus tem uma cobertura maior do que a WoS, ela tem a vantagem de abranger um número maior de periódicos.

O Índice h foi proposto por Hirsch (2005) como um indicador que pondera tanto a produtividade como o impacto científico. Para que um índice H seja igual a X é necessário que N trabalhos tenham pelo menos X citações cada um, e os outros (N-H) artigos tenham $\leq X$ citações cada um (HIRSCH, 2005). Dessa forma, a obtenção de um índice h igual a 8 significa que foram publicados, ao menos,

8 artigos, os quais receberam, ao menos, 8 citações cada um. Se o índice h for zero, o autor pode ter diversos artigos publicados, porém nenhum recebeu citações. Nesse aspecto, o índice h, além de permitir a análise da quantidade de publicações, possibilita a avaliação do seu impacto científico. Pela sua simplicidade, facilidade de uso e por poder ser calculado a partir de qualquer ferramenta que compile citações, o índice h se tornou extremamente popular e vem sendo aplicado para avaliação de pesquisadores, departamentos, revistas e universidades.

Uma das ferramentas que compila citações e, portanto, permite calcular o índice h é o Google Scholar (GS), ou Google Acadêmico, na versão brasileira. Derivado do mecanismo de busca Google, o GS foi lançado em 2005 como forma de recuperação de informações em domínios web de instituições acadêmicas, mas alguns anos depois passou a incluir indicadores bibliométricos de citação para revistas e pesquisadores. Diferentemente das bases de dados WoS e Scopus, que indexam os artigos de periódicos selecionados, o mecanismo do Google Scholar rastreia a web e reconhece os documentos científicos e suas referências, independentemente do tipo de documento, língua ou periódico. Por isso, os indicadores dele derivado são muito mais abrangentes, o que é particularmente útil para a avaliação nas áreas de ciências humanas, sociais e sociais aplicadas, que estão bem menos representadas na WoS e na Scopus do que as ciências da vida, exatas e tecnológicas.

O índice h de um pesquisador, juntamente com a lista de suas publicações e o número de citações para cada uma, é facilmente obtido no GS por meio do seu perfil, se este tiver sido acionado. Para periódicos, o GS oferece a opção “Métricas”, que apresenta o ranking dos 100 principais periódicos em 11 línguas e por área do conhecimento, além de permitir a localização individual de um número muito mais abrangente de títulos. O indicador utilizado por ele para os periódicos é o índice h5, que consiste no índice h do conjunto dos artigos publicados naquele periódico nos últimos 5 anos completos. Portanto, quando se fala do h5 no GS se está falando do índice h de um determinado periódico, calculado a partir das suas publicações dos últimos 5 anos e das citações a essas publicações, recuperadas pelo GS. O número derivado dele, assim como o número total de publicações e citações recuperadas, é geralmente superior ao mesmo índice quando calculado a partir da WoS ou da Scopus, pela seletividade dessas duas bases de dados (MINASNY, 2013). Além da maior seletividade, é importante ressaltar que a WoS faz controle de autocitações e práticas de citações indevidas (VAN NOORDEN, 2013).

Além de avaliar os diferentes aspectos que compõem cada indicador bibliométricos e suas implicações na avaliação das pesquisas, é também importante considerar que o perfil de publicação e de citação é particular à cada área do conhecimento. Sabe-se que as áreas das ciências exatas e da

vida citam um volume maior de artigos; que as ciências sociais e humanidades citam documentos mais antigos; que as biomédicas citam documentos bem atuais. Portanto, a comparação entre áreas só é possível quando se utilizam indicadores normalizados. Neste artigo, esses indicadores não são abordados, pois o objetivo é aproximar a discussão sobre avaliação de periódicos às diretrizes do Qualis Referência, e não há indícios de que a Capes esteja propondo uma avaliação comparativa entre áreas.

4 QUALIS: da criação à proposta do qualis referência

O crescimento da pós-graduação no Brasil foi significativo e provocou um aumento na produção científica dos pesquisadores envolvidos. A elaboração de um sistema de avaliação de revistas teve como objetivo servir como instrumento para a avaliação dos programas de pós-graduação, através do estabelecimento de parâmetros para avaliação da produção dos professores ligados aos programas de pós-graduação (BARATA, 2016).

O Qualis compreende a qualificação da produção científica que subsidia a avaliação conduzida pela CAPES, alimentada a partir dos relatórios enviados pelos programas de pós-graduação no sistema Sucupira. Até o Qualis 2016, cada área designava uma comissão responsável pela arbitragem dos periódicos utilizados como veículos por docentes e discentes, responsável por elaborar os critérios de avaliação considerando as características da área (BARATA, 2016; COSTA; YAMAMOTO, 2008).

Vogel (2017), em levantamento dos critérios Qualis 2013 e 2016, concluiu que não há consenso na avaliação de periódicos das 49 áreas. Enquanto algumas áreas utilizam somente indicadores bibliométricos como o Fator de Impacto, o índice h e as citações por documentos, há áreas que consideram a presença dos periódicos em bases de dados bibliográficas, como Web of Science, Scopus e Scielo. Outras áreas, por sua vez, focam nos processos editoriais como número de artigos por volume, quantidade de autores doutores, etc. Na classificação 2016, cinco das 49 áreas de avaliação da CAPES não utilizavam nenhum indicador de impacto: Antropologia/Arqueologia; Educação; Arquitetura e Urbanismo; Comunicação e Informação; e Artes/Música. Ao comparar o Qualis 2013 e 2016, a autora observou que os indicadores de impacto ganharam espaço, em uma tentativa de diminuir o nível de subjetividade da avaliação com critérios quantitativos, inclusive em áreas das Humanidades.

Nas áreas que não utilizam ou não há indicadores consolidados baseados em citações, todo o peso da avaliação recai na opinião de especialistas e, mais especificamente, na opinião dos especialistas

que compõem o comitê. Nesse sentido, embora os indicadores baseados em citações não podem e não devem substituir o julgamento de especialistas, não se considera adequada a substituição absoluta de todos os indicadores objetivos pela opinião de uma comissão de especialistas que é juiz e parte no processo de avaliação (MARTINEZ-AVILA, 2019). O juízo de valor é dependente de interesses, já que os integrantes dos comitês Qualis são docentes dos mesmos programas de pós-graduação que editam revistas e são avaliados pelo sistema de avaliação quadrienal.

Em 2019 a CAPES propôs uma nova metodologia para cálculo do Qualis, a partir de critérios mais objetivos e uma fórmula que permitisse a comparação equilibrada entre as áreas de avaliação, atentando para a internacionalização (CAPES, 2019, online). Essa nova metodologia se baseia em quatro princípios: classificação única; classificação por áreas mãe, Qualis Referência e Indicadores bibliométricos.

A metodologia proposta pelo GT Qualis Periódicos da CAPES, instituído pela Portaria nº 150, de 4 de julho de 2018 (CAPES, 2018), parte da premissa de que cada periódico recebe apenas uma classificação, que é atribuída por uma área mãe. Área mãe é aquela que recebeu o maior volume de publicações para aquele periódico em determinado período (os anos 2017 e 2018 no caso do Qualis Referência 2019). Em situações de empate, é considerada área mãe aquela em que o número de publicações no periódico foi mais representativo em relação ao total de produções da área (CAPES, 2019). Nesta lógica, as publicações de pesquisadores de uma determinada área em periódicos que apresentaram maior volume de publicação originário de outras áreas, foram classificadas por essas outras áreas.

A classificação do Qualis Referência é calculada por meio de uma metodologia que considera indicadores objetivos e um modelo matemático. Os indicadores utilizados são o *CiteScore* da Scopus, o Fator de Impacto da WoS e o h5 do Google Scholar, todos eles indicadores de citação. Na seleção dos indicadores por áreas, o Qualis Referência atendeu às particularidades das áreas, identificando o valor do indicador e percentil nas categorias de área das bases de dados (Web of Science e Scopus).

Conforme determinado no Ofício 5 (CAPES, 2019), nos casos em que o periódico possuía *CiteScore* e/ou FI, foi considerado para fins de estratificação o maior valor de percentil entre eles. Nos casos em que o periódico não possuía *CiteScore* e/ou JIF, foi verificado o valor do índice h5 do Google Scholar, disponível no Google Scholar Metrics. Para que houvesse uma correlação entre os indicadores, a área técnica da CAPES utilizou um modelo de regressão que fez a correlação entre valores de h5 e *CiteScore*. Assim, para periódicos que só possuíam h5, foi possível estimar um valor correspondente de percentil, denominado Percentil Imputado.

Periódicos com percentis imputados pelo h5 puderam ter trava definida em estrato específico, determinado por cada área de avaliação ou grupos de áreas, a fim de evitar sobrevalorizar periódicos que não estejam nos indexadores internacionais. Algumas áreas observaram este detalhe, outras não. Ao final, o estrato referência foi calculado por intervalos iguais (12,5%) do percentil final, resultando em oito classes com os seguintes recortes:

- a) 87,5 define valor mínimo do 1º estrato (A1)
- b) 75 define valor mínimo do 2º estrato (A2)
- c) 62,5 define valor mínimo do 3º estrato (A3)
- d) 50 define valor mínimo do 4º estrato (A4)
- e) 37,5 define valor mínimo do 5º estrato (B1)
- f) 25 define valor mínimo do 6º estrato (B2)
- g) 12,5 define valor mínimo do 7º estrato (B3)
- h) Valor máximo do 8º estrato inferior a 12,5 (B4)

Quando o algoritmo de busca utilizado pela CAPES não identificou o índice h5 do Google Scholar para determinado periódico, a alternativa foi buscá-lo manualmente por meio do software *Publish or Perish*, que permite a busca por título ou ISSN. Essa tarefa foi desenvolvida pelas comissões Qualis de cada uma das 49 áreas de avaliação que optou pela utilização do h5.

Muito embora tenha havido um esforço considerável para a proposição do Qualis Referência, parece que ainda existem alguns erros conceituais e metodológicos nas decisões tomadas. As citações, apesar de comporem um indicador bibliométrico muito utilizado e que apresenta correspondência com a avaliação por pares, não são abrangentes o suficiente para medir o desempenho de revistas científicas. Revistas novas que vem desenvolvendo trabalho promissor e ganhando credibilidade em sua área, mesmo indexadas na WoS ou na Scopus, nem sempre tem alto volume de citações e portanto, um Índice h equivalente. Além disso, o Índice h é um indicador *size dependent* (WOUTERS *et al.*, 2014), o que beneficia revistas que publicam anualmente maior volume de artigos e que tem mais idade, mesmo que não publique novos artigos (HICKS *et al.*, 2015). A opção pelo Índice h5, no entanto, foi apropriada para dirimir o problema da idade, já que permite a comparação entre periódicos mais novos com os mais velhos. Além disso, distribuição de citações em uma revista é irregular, sendo que cerca de 20% dos artigos recebem 60% das citações, o que pode enviesar as análises baseadas nesses indicadores de citação.

O Qualis Referência também pode ser criticado por reunir três indicadores distintos. Os indicadores de impacto FI e *CiteScore* são calculados a partir de citações feitas por revistas indexadas pelas bases de dados Web of Science e Scopus, o que limita a possibilidade de contagem das citações nas áreas das ciências sociais e humanidades. A WoS indexa principalmente os periódicos com base nos Estados Unidos e em língua inglesa (HICKS *et al.*, 2015). Por outro lado, o Google Acadêmico, por sua capacidade de trabalhar com documentos existentes na web, consegue recuperar todo o tipo de citações, incluindo aquelas que não são propriamente citações feitas em documentos científicos. Revistas terciárias, quais sejam, que reúnem resumos e sumários de outras publicações, são extremamente beneficiadas no cálculo do Índice h do Google Acadêmico.

O uso de indicadores de citação das revistas recebe críticas no momento atual, em que adoção da publicação eletrônica proporcionou às revistas maior visibilidade e grande ampliação de público. O que antes ficava restrito às coleções das bibliotecas que assinavam o periódico, hoje está disponível na internet em acesso aberto, em alguns casos. Isso propicia que os artigos possam ser lidos individualmente e por isso, nada mais justo do que uma medição individual de suas métricas de desempenho, que possam abranger não só citações em documentos científicos, mas também compartilhamentos, comentários, tweets entre outras interações que ocorrem no meio digital. Essas métricas ampliaram as formas tradicionais de medir o uso, a visibilidade e o impacto dos periódicos científicos (DELGADO LÓPEZ CÓZAR, 2020).

O sistema Qualis está imbricado na pós-graduação brasileira e portanto, presente na vida dos pesquisadores. O sistema repercute em muitas atividades que avaliam os pesquisadores, para atribuição de bolsas até contratação de professores. Costa e Yamamoto (2008) já afirmaram há anos que o propósito inicial do Qualis foi extrapolado, que seus resultados vêm sendo utilizados como ferramenta para auxiliar na concessão de financiamentos, para a inclusão dos títulos em bibliotecas e indexadores, para orientar pesquisadores no momento de escolha de revistas para submissão de trabalhos. Segundo Barata (2016), o uso do Qualis Periódico em avaliações de desempenho dos pesquisadores para concessão de bolsa produtividade ou auxílios de pesquisa é incorreto e inadequado, considerando-se que a aplicação do Qualis só faz sentido para análise coletiva da produção de programas de pós-graduação. Áreas como Educação, por exemplo, afirmam que os pesquisadores moldam suas práticas de publicação a partir do Qualis, reiterando a importância do sistema entre a comunidade brasileira (REATEGUI *et al.*, 2020).

Crozariol (2018) alerta que algumas áreas de pesquisa, como a História da Zoologia, correm risco de diminuir o volume de pesquisadores dedicados e sua produção científica, como consequência

direta das dificuldades de avaliação. Nesta área, as opções de publicação se concentram em revistas da História e da Zoologia. No entanto, um Zoólogo que publica em revistas de História pode ser surpreendido por uma avaliação baixa para uma revista de História em sua área (Zoologia ou Biodiversidade), considerando-se que dependendo da área de avaliação a classificação de um mesmo periódico pode ser divergente. Nos Qualis anteriores a CAPES concedia autonomia às áreas para estabelecer seus critérios de avaliação, o que era vantajoso em certos aspectos, mas prejudicial em outros (REATEGUI *et al.*, 2020). A avaliação única para todas as áreas, como proposta pelo Qualis 2019, pode solucionar este problema.

No entanto, autores como Costa e Yamamoto (2008) também reconhecem que o Qualis serve como estímulo aos editores para elevar o padrão de qualidade das revistas. A avaliação promove a visibilidade dos periódicos, incentivando a inserção das mesmas em um fórum internacional. Segundo os autores, a despeito dos aspectos negativos das avaliações, elas provocaram um avanço útil e necessário nas revistas da área da psicologia. Os indexadores, que eram desconhecidos por muitos editores, passaram a ser discutidos pela comunidade a partir do Qualis. A pressão para o desenvolvimento e uso de métricas mais adequadas e baseadas em indicadores sólidos é um incentivo para a melhor avaliação e conseqüente aprimoramento da ciência como um todo (LANE, 2010).

5 CONSIDERAÇÕES FINAIS

As revistas devem buscar a publicação de artigos que apresentem resultados inéditos à comunidade científica. Isso é possível com uma equipe editorial atenta, com a participação de pareceristas criteriosos, com processos editoriais ágeis que evitem longo tempo de espera para avaliação e publicação. Bons artigos serão, presumivelmente, lidos pela comunidade científica e citados, levando às revistas a alcançarem bons indicadores de citação.

As principais críticas tecidas à utilização de métricas de impacto se relacionam a cobertura das bases de dados. Neste sentido, a opção pelo uso do Índice h do Google Acadêmico atende a todas as áreas, incluindo as Humanidades, caracterizadas pelo perfil de publicação local. No entanto, misturar indicadores provenientes de duas bases internacionais mais elitistas com o indicador do Google, que indexa todo e qualquer tipo de documento, não soluciona o objetivo principal do Qualis Referência, qual seja, ter uma classificação única. Revistas que estão indexadas na Web of Science e Scopus deveriam ser valorizadas pois alcançar tal patamar demanda um esforço importante, que merece amplo

reconhecimento. No entanto, ao misturar três indicadores diferentes, a fórmula originalmente proposta pela CAPES penalizou tais revistas, supostamente por apresentarem Fator de Impacto e *CiteScore* baixos. Considerando que as coleções de revistas indexadas por estas bases de dados na área de Humanidades é pequena, é natural que o volume de citações gerados por elas seja pequeno também.

Nesse sentido, revistas com baixo FI ou *CiteScore* foram, em geral, mal avaliadas, considerando que estes dois indicadores têm prioridade na fórmula criada pela CAPES para o Qualis Referência. Paralelamente, revistas não indexadas que contam com bom Índice h do Google foram elevadas a extratos mais altos, já que este foi o único indicador utilizado para estas revistas. Nesse sentido, a tão sonhada internacionalização perde terreno, na lógica de que o Qualis molda as práticas de publicação da comunidade científica. Afinal, por que um pesquisador iria se esforçar para escrever um texto em outra língua e submetê-lo a uma revista indexada em bases de dados internacionais, se corre o risco de receber menor pontuação pelo extrato Qualis?

O uso de indicadores de citação para avaliação de revistas científicas não pode ser demonizado pela comunidade científica pois, são indicadores sólidos, construídos com base em teorias e metodologias bem fundamentadas. No contexto de um país latinoamericano não nativo em inglês, é claro, torna-se necessário ponderar o uso de indicadores consolidados, como o Fator de Impacto e o *CiteScore*, com indicadores produzidos localmente. Bases de dados como a SciELO e Redalyc poderiam contribuir com indicadores mais próximos da realidade nacional.

REFERÊNCIAS

ABRAMO, G. Revisiting the scientometric conceptualization of impact and its measurement. **Journal of Informetrics**, v. 12, n.3, p. 590-597, 2018. DOI: 10.1016/j.joi.2018.05.001

AKSNES, D. W.; LANGFELDT, L.; WOUTERS, P. Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. **SAGE Open**, v. 9, n.1, p. 1-17, 2019. DOI : 10.1177/2158244019829575

ALPERIN, J.P; GOMEZ, C. J; HAUSTEIN, S. Identifying diffusion patterns of research articles on Twitter: a case study of online engagement with open access articles. **Public Understanding of Science**, v. 28, n.1, p. 2-18, 2019. DOI:10.1177/0963662518761733

BALDWIN, Melinda. Scientific Autonomy, Public Accountability, and the Rise of “Peer Review” in the Cold War United States. **Isis: a Journal of the History of Science Society**, v. 109, n. 3, p. 538–58, 2018.

BARATA, Rita de Cássia Barradas. Dez coisas que você deveria saber sobre o Qualis. **RBPG**, Brasília, v. 13, n. 30, p. 13 - 40, jan./abr. 2016.

BARNES, Cameron. The Use of Altmetrics as a Tool for Measuring Research Impact. **Australian Academic & Research Libraries**, v. 46, n.2, p. 121-134, 2015. DOI: 0.1080/00048623.2014.1003174

BOLLEN, J.; RODRIQUEZ, M.; VAN DE SOMPEL, H. Journal status. **Scientometrics**, v.69, n.3, p. 669–687, 2006. DOI: 10.1007/s11192-006-0176-z

BORNMANN, Lutz. Scientific Revolution in Scientometrics: The Broadening of Impact from Citation to Societal. In: SUGIMOTO, C. R. (Ed.) **Theories of Informetrics and Scholarly Communication**. Berlin: Degruyter, 2016. p. 357-359.

BORNMANN, L.; DANIEL, H.-D. What do citation counts measure? a review of studies on citing behavior. **Journal of Documentation**, v. 64, n. 1, p. 45-79, 2008.

BUSH, Vannevar. **Science The Endless Frontier**: a Report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development, July 1945. Washington, United States Government Printing Office, 1945. Disponível em: <https://www.nsf.gov/od/lpa/nsf50/vbush1945.htm>

CLARIVATE ANALYTICS. **Clarivate Analytics Impact Factor**. 2020. Disponível em: <https://clarivate.com/essays/impact-factor/>. Acesso em: 25 ago. 2020.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. Ofício nº 5/2019-CGAP/DAV/CAPEs. 9 de maio de 2019. Disponível em: http://uploads.capes.gov.br/files/oficio_qualis_referencia.pdf. Acesso em: 25 de ago. 2020.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. Portaria nº 150, de 4 de julho de 2018. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/28988995/do2-2018-07-06-portaria-n-150-de-4-de-julho-de-2018-28988986 . Acesso em: 25 ago. 2020.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. CAPES melhora ferramentas de avaliação da pós-graduação. Disponível em: <https://www.capes.gov.br/36-noticias/9730-capes-melhora-ferramentas-de-avaliacao-da-pos-%20graduacao> . Acesso em: 25 ago. 2020.

COSTA, Ana Ludmila Freire; YAMAMOTO, Oswaldo Hajime. Publicação e avaliação de periódicos científicos: paradoxos da avaliação Qualis de Psicologia. **Psicologia em Estudo**, Maringá, v. 13, n. 1, p. 13-24, jan./mar. 2008.

CRONIN, Blaise. **The citation process**: the role and significance of citations in scientific communication. London: Taylor Graham, 1984.

CROZARIOL, Marco Aurélio. Os prejuízos científicos causados pela utilização da classificação Qualis/CAPEs para avaliar o desempenho do pesquisador. **Informativo da Sociedade Brasileira de Zoologia**, jun./jul. 2018, p. 5-6.

DAVIS, P. M. Reward or persuasion?: The battle to define the meaning of a citation. **Learned Publishing**, v. 22, p. 5-11, 2009.

DELGADO LÓPEZ CÓZAR, Emilio. El Ranking de revistas científicas españolas con Sello de calidad Fecyt: un constructo bibliométrico artificioso y obsoleto. **Anuario ThinkEPI**, Barcelona, v. 14, 2020.

DING, Y; CRONIN, B. Popular and/or prestigious? Measures of scholarly esteem. **Information Processing & Management**, v. 47, n. 1, p. 80-96, 2011.

ELSEVIER. **CiteScore**. Disponível em: https://service-elseviercom.ez1.periodicos.capes.gov.br/app/answers/detail/a_id/14880/supporthub/scopus/ . Acesso em: 18 jun. 2019

FREIRE, H. de A.; FREIRE, I. M. Novo Qualis de periódicos da Capes. **Informação & Sociedade**, João Pessoa, v. 29, n. 4, 2019.

GARFIELD, Eugene. Citation Indexes for Science: a new dimension in documentation through association of ideas. **Science**, 1955, v. 122, p. 108-111.

GOOGLE. Google H5 Index. Disponível em: https://scholar.google.com/citations?view_op=top_venues&hl=en&vq=en . Acesso em: 18 jun. 2019.

HICKS, D. *et al.* The Leiden Manifesto for research metrics. **Nature**, v. 520, p. 429-431, 2015. Disponível em: <http://www.leidenmanifesto.org/uploads/4/1/6/0/41603901/leiden-manifesto-portuguese-br-final.pdf> . Acesso em: 24 set. 2020.

JAMES, C. *et al.* CiteScore metrics: Creating journal metrics from the Scopus citation index. **Learned Publishing**, v. 32, n. 4, p. 367–374, 2019. DOI: <https://doi.org/10.1002/leap.1246>.

LANE, Julia. Let's make science metrics more scientific. **Nature**, v. 464, 25 March 2010.

LEYDESDORFF, L. Evaluation of Research and Evolution of Science Indicators. **Current Science**, v. 89, n. 9, p. 1510-517, 2005. Disponível em: <http://www.jstor.org/stable/24110916>. Acesso em: 18 jul. 2020.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, DF, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <<http://www.scielo.br/pdf/ci/v35n2/a04v35n2.pdf>>. Acesso em: 08 maio 2020.

MARTÍNEZ-ÁVILA, D. Qualis Periódicos: el sistema brasileño de evaluación de revistas. **Anuario ThinkEPI**, Barcelona, v. 13, 2019.

MINASNY, Budiman *et al.* Citations and the h index of soil researchers and journals in the Web of Science, Scopus, and Google Scholar. **PEERJ**, v. 1, n. e183, 2013. Disponível em: <https://peerj.com/articles/183.pdf>. Acesso em: 12 jul. 2020.

MOED, Henk F. **Applied Evaluative Informetrics**. Cham (Switzerland): Springer International Publishing, 2017.

PRIEM, J.; HEMMINGER, B.M. Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. **First Monday**, v. 15, n. 7, p. 2010. Disponível em: from: <https://firstmonday.org/article/view/2874/2570>. Acesso em: 11 set. 2020.

REATEGUI, Eliseo *et al.* Evaluation of Brazilian research output in education: confronting international and national contexts. **Scientometrics**, v. 125, 2020, p. 427-444.

SILVEIRA, M. A. da; CAREGNATO, S. E. Demarcações epistemológicas dos estudos de citação: concepção sociocultural das citações. **Perspectivas em Ciência da Informação**, v. 23, n. 1, p. 55-70, 2018. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/3125>. Acesso em: 07 out. 2020.

SPIER, R. The history of the peer-review process. **Trends in Biotechnology**, v. 20, n. 8, p. 357-358, 2002.

VANZ, S. A. S.; CAREGNATO, S. E. Estudos de Citação: uma ferramenta para entender a comunicação científica. **Em Questão**, Porto Alegre, v. 9, n. 2, p. 247-259, 2003. Disponível em: <<http://seer.ufrgs.br/index.php/EmQuestao/article/view/75/35>>. Acesso em: 16 jan. 2018.

VOGEL, Michely Jabala Mamede. Uso de indicadores bibliométricos na avaliação da CAPES: o Qualis periódicos. XVIII Encontro Nacional de Pesquisa em Ciência da Informação 2017. Marília.

VAN NOORDEN, Richard. Brazilian citation scheme ousted: Thomson Reuters suspends journals from its rankings for 'citation stacking'. **Nature**, v. 500, 29 Ago. 2013.

WALTMAN, L. A review of the literature on citation impact indicators. **Journal of Informetrics**, v. 10, p. 365-391, 2016. DOI:10.1016/j.joi.2016.02.007

WOUTERS, P. *et al.* **The metric tide**: Literature review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management). Bristol, HEFCE, 2015. DOI: 10.13140/RG.2.1.5066.3520

WOUTERS, Paul *et al.* **Bibliometrics and the use of citation indicators in research assessment**: a review. 2014.