

# Índice h de Hirsch: análise comparativa entre as bases de dados Scopus, *Web of Science* e Google Acadêmico

**Deise Deolindo Silva**

Doutoranda; Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, SP, Brasil;  
deisedeolindo@hotmail.com

**Maria Cláudia Cabrini Grácio**

Doutora; Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, SP, Brasil;  
cabrini@marilia.unesp.br

**Resumo:** Entre os indicadores destinados a avaliar o impacto científico das publicações, destaca-se o índice h de Hirsch, que mede, de forma simultânea, aspectos relacionados à produção e ao impacto da produção científica de um pesquisador, periódico, grupo e centro de pesquisa ou país em uma base de dados. Nesse contexto, esta pesquisa tem por objetivo obter os índices h para 20 pesquisadores representativos na temática “Estudos Métricos da Informação” em três fontes de dados distintas — Scopus, *Web of Science* e Google Acadêmico e identificar e analisar possíveis diferenças entre os valores desses indicadores, a fim de examinar a congruência dos índices nessas fontes de informação. Para isso, levantaram-se, nas plataformas mencionadas, os índices h para cada investigador. Identifica-se, por meio do Teste de Wilcoxon-Mann-Whitney com nível de significância de 5%, que o índice h não difere estatisticamente entre as bases de dados Scopus e *Web of Science*. Por outro lado, esse índice diferiu significativamente dos valores obtidos a partir do Google Acadêmico.

**Palavras-Chave:** Estudos métricos. Índice h. Análise de domínio. Bases de dados.

## 1 Introdução

A socialização do conhecimento produzido é um fenômeno que cresce ao longo das últimas décadas, principalmente com o advento das Tecnologias da Informação e Comunicação (TIC). Observa-se um crescimento significativo do número de resultados de pesquisas científicas publicados em algum meio, especialmente em bases de dados internacionais.

O ato de socializar o conhecimento é retratado por Guimarães (2000), que relaciona esse procedimento a um ciclo espiral de produção e comunicação da

ciência, no qual conhecimento é efeito das relações sociais, a partir de um consenso social, por isso tende a ser ascendente.

Com o aumento da produção científica, a análise desta se tornou um importante objeto de estudo, por ter a capacidade de refletir o grau de desenvolvimento de uma nação e poder proporcionar diretrizes nos âmbitos social, político e econômico. Além desse aspecto, a avaliação dessa produção se tornou um fator imprescindível, assim como a elaboração de critérios que avaliem os diversos domínios científicos.

Observa-se, na literatura, uma crescente preocupação em se elaborar bases de dados confiáveis e mecanismos adequados de avaliação de domínios científicos, assim como de se dispor de instrumentos e indicadores para definição de suas diretrizes, alocação de investimentos e recursos, formulação de programas e avaliação de atividades relacionadas ao desenvolvimento científico e tecnológico, como forma de identificar as elites científicas, garantindo um investimento profícuo das agências de fomento à pesquisa. (MUGNAINI; JANNUZZI; QUONIAM, 2004; THOMAZ; ASSAD; MOREIRA, 2011; VANZ; STUMPF, 2010).

Nesse sentido, as bases de dados dão suporte para a almejada visibilidade da produção científica, pois a partir dessas fontes de informação são encontrados resultados de pesquisas, de pesquisadores e de instituições, nacionais e internacionais. Dentre as plataformas de dados internacionais, pode-se destacar a *Scopus* e a *Web of Science* (WoS). Elas são bases de artigos científicos multidisciplinares que reúnem um amplo conjunto de revistas de diferentes editores. Além disso, possibilitam o acesso global a grande parte da literatura científica publicada internacionalmente; possuem ferramentas adequadas para a realização de uma revisão bibliográfica em um determinado domínio; são abrangentes e incluem contagem de citações por artigo e outros dados adicionais (SILVA, 2013).

A base *Scopus* foi lançada pela editora *Elsevier* em 2004. No ano de 2013, contava com aproximadamente 21.000 títulos de mais de 5.000 editoras internacionais. Com atualização diária, a cobertura se dá desde 1823 e realizam-se as contagens de citações desde 1996. É considerada como a base que tem a maior

cobertura de resumos, citações e textos completos da literatura científica internacional e brasileira. Além disso, foi pioneira na implementação do índice *h* como ferramenta bibliométrica do seu banco de dados (ELSEVIER, 2015; LIMA; VELHO; FARIA, 2012; SILVA, 2013).

A *Web of Science* é considerada a base de dados mais tradicional; em 2013, contava com mais de 12.000 títulos de pelo menos 3.300 editoras internacionais, os quais são atualizados semanalmente. Sua cobertura e contagem de citações remontam ao ano de 1900. Os critérios para a indexação de uma revista na WoS são extremamente rigorosos; o fato de uma revista ser indexada nessa base é considerado sinônimo da sua qualidade. A *Thomson Reuters* desenvolveu também uma base estatística, *Journal Citation Reports* (JCR), a qual permite avaliar e comparar publicações científicas utilizando dados de citações extraídos de revistas acadêmicas e técnicas e o impacto destas na comunidade científica, indexadas pela coleção principal da WoS (BRASIL, 2009; SILVA, 2013; THOMSON REUTERS, 2015; VIEIRA; WAINER, 2013).

Em novembro de 2004, foi lançado o *Google Acadêmico* (GA), que, diferentemente da *Scopus* e da WoS, é uma ferramenta gratuita de busca de publicações científicas, a qual permite localizar trabalhos acadêmicos de vários tipos (teses, dissertações, artigos de congressos e de periódicos tanto de acesso aberto como fechado) e em múltiplas línguas. Ele busca referências em qualquer documento que esteja disponível na *Web*. Essa base de dados não considera citações feitas por livros, embora as citações recebidas por livros sejam consideradas (BARILAN, 2008; CAREGNATO, 2011; VIEIRA; WAINER, 2013).

Dada a importância dessas plataformas de disseminação do conhecimento científico e a necessidade de se avaliar a produção científica veiculada por elas, os estudos bibliométricos se tornam essenciais ao contribuir para a visualização do desempenho e do comportamento científico de artigos, de autores, instituições, periódicos e países que participam do domínio científico constituído por essas plataformas.

Em concordância, Hjørland e Albrechtsen (1995) ressaltam que para entender a informação é necessário estudar os domínios do conhecimento relacionados às comunidades discursivas, em que se identificam as teorias

análogas de pensamento, a linguagem e o conhecimento. Entre as abordagens para a caracterização e avaliação de um domínio, Hjørland (2002) destaca os estudos bibliométricos, especialmente as análises baseadas nas citações, por constituir abordagem consolidada e objetiva.

As análises bibliométricas têm se mostrado procedimentos tangíveis e confiáveis. As métricas utilizadas nesse contexto são, geralmente, relacionadas à produção, ligação e citação, e tem por intuito explicitar, além da produtividade, a relevância e o impacto de autores, periódicos, instituições, grupos ou países nas diferentes áreas do conhecimento (OLIVEIRA; GRÁCIO, 2011).

Os estudos bibliométricos podem ser considerados como o conjunto de conhecimentos relacionados à avaliação da informação produzida. Referem-se a estudos de natureza teórico-conceitual quando contribuem para o avanço do conhecimento da própria temática, propondo novos conceitos e indicadores ou reflexões e análises relativas à área. São de natureza metodológica quando se propõem a dar sustentação aos trabalhos de caráter teórico da área onde estão aplicados (OLIVEIRA; GRÁCIO, 2011).

Entre os indicadores de impacto, presentes na literatura científica, destaca-se o índice *h*, um parâmetro avaliativo considerado robusto, por avaliar de forma simultânea os aspectos relativos à produção (quantidade de artigos produzidos) e ao impacto (número de citações).

Diversas bases de dados apresentam o cálculo desse indicador, visto que ele é utilizado tanto na avaliação do comportamento da produção científica de pesquisadores como também é um critério avaliativo em diversas agências de fomento. No entanto, esse indicador pode ter valores discordantes se buscado em diferentes fontes de informação.

Considerando o exposto, esta pesquisa teve como objetivo comparar os valores do índice *h* fornecidos pelas fontes de informação *Scopus*, *WoS* e *Google Acadêmico*, com a finalidade de avaliar a congruência entre os valores obtidos por essas três fontes de dados.

Para isso, como universo de pesquisa, tomaram-se 20 pesquisadores, considerados mais produtivos na temática “Estudos Métricos”, conforme estudo

de Silva e Grácio (2015), que atendiam o critério de ter pelo menos 20 artigos publicados na revista *Scientometrics*, em um período de 35 anos (1980 a 2014).

Esta pesquisa se justifica, pois se a congruência dos indicadores for observada entre as plataformas analisadas, o pesquisador e/ou avaliador pode optar em buscar as informações em quaisquer bases de dados. Além disso, trará contribuições para a área dos estudos bibliométricos, principalmente no âmbito metodológico, visto que trabalhar com dados consistentes é um pré-requisito para a obtenção de resultados significativos. Destaca-se, ainda, a carência de estudos comparativos entre as diversas bases de dados na literatura científica.

## 2 Índices do tipo *h* de Hirsch

Os índices do tipo *h* de Hirsch são indicadores de impacto, destacam-se por serem parâmetros avaliativos robustos, por avaliar de forma simultânea os aspectos relativos à produção e ao impacto.

Hirsch (2005, p. 1) apresenta o *índice h* de um cientista como o valor *h* tal que “[...] *h* de seus  $N_p$  artigos têm pelo menos de *h* citações cada e os outros ( $N_p - h$ ) artigos têm não mais que *h* citações cada”.

O cálculo desse indicador é simples e seu desempenho tem se mostrado melhor do que o de outros indicadores, quando considerados de forma isolada. Além desses aspectos, o índice não é influenciado por um conjunto de poucas citações, nem por artigos altamente citados; permite caracterizar a produtividade científica de um pesquisador com objetividade; pode ter utilidade na tomada de decisões sobre promoções, alocação de verbas e atribuição de prêmios; permite prever o desempenho futuro de um pesquisador com capacidade de previsão superior à de outros índices (EGGHE, 2010; HIRSCH, 2005; HIRSCH, 2007; MARQUES, 2013).

A fim de exemplificar, apresenta-se na Tabela 1 o cálculo do índice *h* para o pesquisador Thelwall. Para esse pesquisador, recuperaram-se na base *Scopus* 137 artigos ( $N_p = 137$ ) e as respectivas quantidades de citações recebidas. Para o cálculo do índice *h*, é necessário colocar em ordem crescente o número de artigos (de 1 a 137) e em ordem decrescente o número de citações e verificar qual é o

valor  $h$ , tal que  $h$  dos  $N_p$  artigos de Thelwall têm pelo menos de  $h$  citações cada e os outros  $(N_p - h)$  artigos têm não mais que  $h$  citações cada. Para esse exemplo, identifica-se que Thelwall tem  $h = 32$ , pois existem 32 artigos com pelo menos 32 citações cada e seus outros 105 artigos não têm mais que 32 citações cada.

**Tabela 1** – Cálculo do índice  $h$  para o pesquisador Thelwall, a partir da base *Scopus*.

Índice $h$	
Número de Artigos ( $N_i$ )	Citações
2	125
3	125
4	115
5	114
6	98
7	96
8	94
9	87
10	80
:	:
30	34
31	32
32	32
33	32
34	31
:	:
49	22
50	22
51	22
52	21
53	21
:	:
133	0
134	0
135	0
136	0
137	0

Fonte: Elaborada pelas autoras.

Egghe (2010) e Marques (2013) comentam que o índice  $h$  não deve ser utilizado para comparar pesquisadores de áreas do conhecimento diferentes; pode

atribuir vantagens para aqueles que têm maior tempo de titulação; pode ser manipulado por meio de autocitações; dá a livros o mesmo peso que dá aos artigos, tornando complicado comparar pesquisadores de áreas em que há a cultura de publicar os resultados de pesquisa em livros, como as humanidades; não considera o contexto das citações; não faz distinção entre um artigo feito por um pesquisador ou um pequeno grupo de colaboradores e um artigo com centenas de autores, cuja participação individual é difícil de avaliar.

Shing e Kumar (2014) apontam que os autores podem aumentar seu índice por meio de um incremento no número de citações recebidas, mesmo que não haja novos artigos publicados.

Devido às deficiências apresentadas, diversas variações desse indicador têm sido desenvolvidas no intuito de minimizar ou dirimir os problemas e questões levantadas. Bornmann, Mutz e Daniel (2008) apresentam diversas variações do índice *h*, que estão sumarizadas no Quadro 1.

**Quadro 1** – Índices do tipo *h* de Hirsch.

Índice	Autor	Definição
<b>Quociente <i>m</i></b>	Hirsch (2005)	$h/y$ em que $h$ = índice <i>h</i> , $y$ = número de anos desde a publicação do primeiro artigo.
<b>Índice <i>g</i></b>	Egghe (2006)	$g$ é o posto mais alto, de modo que os artigos principais têm juntos pelo menos $g^2$ citações.
<b>Índice <i>h(2)</i></b>	Kosmulski (2006)	É o maior número natural tal que seu $h(2)$ trabalhos mais citados receberam cada um pelo menos $[h(2)]^2$ citações.
<b>Índice <i>a</i></b>	Jin (2006)	$\frac{1}{h} \sum_{j=1}^h cit_j$ , em que $h$ = índice <i>h</i> e $cit$ = número de citações
<b>Índice <i>m</i></b>	Bornmann <i>et al.</i> (2008)	É o número mediano de citações recebidas por artigos do núcleo Hirsch.
<b>Índice <i>r</i></b>	Jin, Liang, Rousseau, Egghe (2007)	$\sqrt{\sum_{j=1}^h cit_j}$ em que $h$ = índice <i>h</i> e $cit$ = número de citações
<b>Índice <i>ar</i></b>	Jin, Liang, Rousseau, Egghe (2007)	$\sqrt{\sum_{j=1}^h \frac{cit_j}{a_j}}$ em que $h$ = índice <i>h</i> , $cit$ = número de citações e $a$ = número de anos desde a publicação

Índice $h_w$	Egghe e Rousseau (2008)	$\sqrt{\sum_{j=1}^{r_0} cit_j}$ <p>em que <math>cit</math> = número de citações e  <math>r_0</math> = o maior índice de linha <math>j</math> de tal forma que  <math>r_w(j) \leq cit_j</math></p>
--------------	-------------------------	---

Fonte: Elaborado pelas autoras com base em Bornmann, Mutz e Daniel (2008).

Segundo Bornmann, Mutz e Daniel (2008), pode-se dividir os índices do tipo  $h$  de Hirsch em dois grupos. O primeiro composto por: índice  $h$ , quociente  $m$ , índice  $g$  e índice  $h(2)$ , os quais descrevem o núcleo da produtividade de um cientista. Os outros índices (índice  $a$ , índice  $m$ , índice  $r$ , índice  $ar$  e o índice  $h_w$ ) retratam o impacto dos artigos do núcleo.

### 3 Procedimentos metodológicos

O universo de pesquisa é composto pelos 20 autores mais produtivos no periódico *Scientometrics*, no período de 35 anos (1980-2014), com pelo menos 20 artigos científicos publicados, destacados no estudo de Silva e Grácio (2015). Levantou-se o índice  $h$  para cada pesquisador, nas bases *Scopus*, *WoS* e *Google Acadêmico*, no período de 2 a 10 de novembro de 2015.

Para analisar se houve diferença significativa entre os valores obtidos para o índice  $h$  a partir das bases *Scopus*, *WoS* e *Google Acadêmico*, utilizou-se o Teste de *Kruskal-Wallis*.

Conforme Siegel (1975) e Portal Action (2015), esse teste é indicado quando se pretende testar se as distribuições são iguais em posição, isto é, identificar se uma população tende a ter valores maiores do que a outra, ou se elas têm a mesma mediana.

Para isso, consideram-se algumas características relevantes:

- as duas populações  $P_1$  e  $P_2$  das quais não se têm informações a respeito de suas distribuições;
- as amostras são independentes e identicamente distribuídas; e
- os valores amostrais são mutuamente independentes.

A Estatística do Teste é dada por:



$$W = U_m - m \cdot n \quad (1)$$

em que,

- $W$ : estatística do teste;
- $m$ : tamanho da amostra relacionada à primeira população;
- $n$ : tamanho da amostra relacionada à segunda população;
- $S_m$ : soma dos postos relacionados à primeira população.

$$U_m = S_m - \frac{1}{2}m(m + 1) \quad (2)$$

#### 4 Análise e interpretação dos resultados

A Tabela 2 apresenta as medidas estatísticas de tendência e dispersão para o índice  $h$  considerando as bases *Scopus*, *WoS* e *Google Acadêmico*.

Verifica-se que há maior concordância entre os valores obtidos para as fontes de informação *Scopus* e *WoS* e que as estimativas obtidas por meio do *Google Acadêmico (GA)* são superiores para todos os autores analisados.

**Tabela 2** – Medidas de tendência e dispersão para o índice  $h$  dos 20 autores mais produtivos na temática Estudos Métricos (1980-2014).

Pesquisador	SCOPUS	WOS	GA
Média	22,6	22,0	38,6
Desvio Padrão	9,0	8,7	15,4
Mínimo	9	9	15
Máximo	38	38	75
Coefficiente de Variação (CV)	40,0%	39,4%	39,9%

Fonte: Elaborada pelas autoras.

A média do índice  $h$  para os investigadores dessa temática para as diferentes bases foram:  $índice-h_{wos} = 22,0$ ,  $índice-h_{scopus} = 22,6$  e  $índice-h_{GA} = 38,6$ . Verifica-se uma pequena diferença entre os valores médios encontrados para *Scopus* e *WoS*, no entanto, a estimativa para a média do índice  $h$  para o *Google Acadêmico* é superior às demais.

Ao analisar a dispersão do índice  $h$  desses autores, identifica-se que em todas as plataformas analisadas o comportamento desses pesquisadores difere (CV

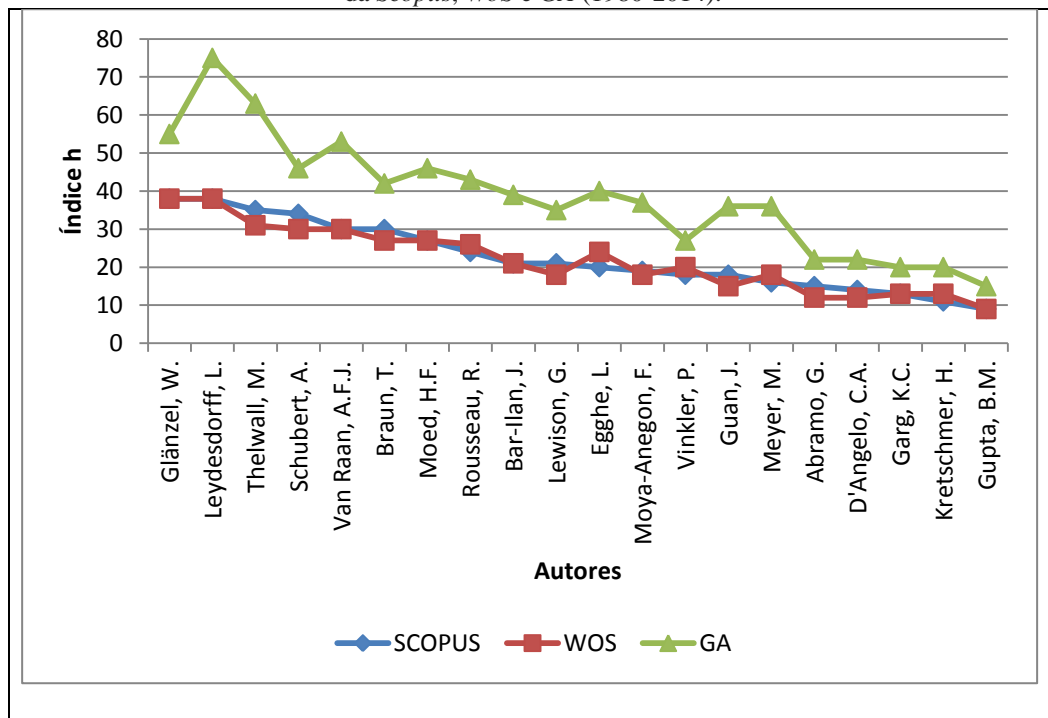
> 30%), ou seja, as publicações dos autores têm comportamentos diferentes. Ao verificar o *Google Acadêmico*, observa-se que o desvio padrão (15,4) é 70% maior que o das outras bases de dados (*Scopus* = 9 e *WoS* = 8,7).

A Figura 1 apresenta os valores encontrados para o índice *h*, obtidos via *Scopus*, *WoS* e *Google Acadêmico*, para os 20 autores mais produtivos em Estudos Métricos.

Considerando as plataformas *Scopus* e *WoS*, observa-se que os autores Glänzel e Leydesdorff apresentam os maiores índices *h*. Além disso, identifica-se uma pequena oscilação entre as estimativas encontradas para os 20 autores. Os investigadores Schubert e Egghe apresentam uma variação maior entre os valores numéricos encontrados nas plataformas discriminadas.

Para a plataforma *Google Acadêmico*, o autor Leydesdorff tem o maior índice *h* ( $h = 75$ ) dentre os pesquisadores analisados, seguido por Thelwall ( $h = 63$ ) e Glänzel ( $h = 55$ ). Gupta apresenta o menor índice *h* em todas as bases analisadas.

**Figura 1** – Índice *h* para os 20 autores mais produtivos em Estudos Métricos, calculados a partir da *Scopus*, *WoS* e *GA* (1980-2014).



Fonte: Elaborada pelas autoras.

Apesar de existir algumas oscilações entre os valores obtidos para o índice *h* dos autores analisados, nas bases de dados *Scopus* e *WoS*, conclui-se que, em

média, a diferença numérica observada se mostrou não significativa pelo Teste de *Kruskal Wallis*, considerando um nível de significância de 5% ( $p\text{-valor} = 0,0001$ ), ou seja, independentemente da plataforma utilizada as estimativas obtidas para os índices *h* não diferem entre si.

Sabe-se que a *Scopus* tem maior quantidade de periódicos indexados em sua base de dados e, apesar disso, o índice *h* dos autores obtidos via *WoS* e *Scopus* não difere significativamente, ou seja, para essa elite científica não faz diferença a indexação de mais periódicos. Os periódicos importantes para essa elite são aqueles reconhecidos/indexados por ambas as bases.

Ao analisar as estimativas do índice *h* dos autores oriundas da *Scopus* e do *Google Acadêmico* identifica-se que elas apresentam diferenças significativas. Similarmente, verificou-se diferença entre os valores dos índices *h* considerando as bases *WoS* e *Google Acadêmico*, ao nível de significância de 5%.

Devido a tais diferenças, buscou-se analisar o comportamento de alguns autores, em especial, no *Google Acadêmico*. Para isso, analisaram-se os perfis dos pesquisadores Leydesdorff, Glänzel e Kretschmer. O primeiro foi escolhido devido ao fato de ter sido o pesquisador com maior índice *h* no GA; Glänzel obteve o maior índice *h* na *Scopus* e *WoS*; e Kretschmer obteve índices *h* semelhantes nas três plataformas analisadas.

**Quadro 2** – Informações disponíveis nas páginas pessoais do Google Acadêmico dos autores Leydesdorff, Glänzel e Kretschmer.

Autor	Leydesdorff	Glänzel	Kretschmer
<b>Informações</b>			
<b>Ano da publicação mais antiga</b>	1981	1983	1979
<b>Número de publicações</b>	678	393	114
<b>Número de coautores</b>	20	Informação não disponibilizada	9
<b>Número de citações</b>	32179	12612	1499
<b>Área de atuação do autor</b>	Scientometrics Autopoiesis Triple Helix Science studies Sociology of innovation	Scientometrics, Probability theory	Bibliometrics, Informetrics, Information theory, Computer science, Psychology
<b>Link para página pessoal</b>	<a href="http://www.leydesdorff.net/list.htm">http://www.leydesdorff.net/list.htm</a>	Informação não disponibilizada	<a href="http://www.h-kretschmer.de/">http://www.h-kretschmer.de/</a>

<b>Autor</b>	<b>Leydesdorff</b>	<b>Glänzel</b>	<b>Kretschmer</b>
<b>Informações</b>			
<b>Artigo mais citado</b>			
<b>Nome</b>	<u>The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations</u>	<u>National characteristics in international scientific co-authorship relations</u>	<u>The structure of scientific collaboration networks in Scientometrics</u>
<b>Ano</b>	2000	2001	2008
<b>Número de citações</b>	5417	528	179
<b>Revista</b>	Research policy	Scientometrics	Scientometrics

Fonte: Elaborado pelas autoras.

Leydesdorff é pesquisador na Universidade de Amsterdã e tem uma página pessoal, que está disponível em [www.leydesdorff.net/list.htm](http://www.leydesdorff.net/list.htm). Atua em diferentes áreas, a saber: cientometria, autopoiese, Tríplice Hélice (da qual o autor é um dos propositores), estudos da ciência e sociologia da inovação.

No *Google Acadêmico*, são 678 trabalhos indexados, sendo estes artigos de diversas áreas, livros, conferências. Tem trabalhos em coautoria com 20 pesquisadores, os quais foram citados por 32.179 pessoas. O artigo mais antigo, intitulado *Some social-psychological aspects of becoming a physicist*, publicado pela *Scientometrics*, data de 1981 e tem três citações.

O artigo mais citado de Leydesdorff foi publicado no ano de 2000 pela revista *Policy Research*, tendo sido citado por 5.417 autores, na temática Tríplice Hélice (*Triple Helix*).

Glänzel não disponibiliza uma página pessoal e atua nas áreas da cientometria e da teoria da probabilidade. No GA, são 393 trabalhos indexados, os quais foram citados por 12.612 autores. O artigo mais citado foi publicado no ano de 2001, sobre relações de coautoria internacional, pela revista *Scientometrics*, tendo sido citado por 528 pesquisadores.

Kretschmer disponibiliza uma página pessoal, atua nas áreas de bibliometria, informetria, teoria da informação, ciência da informação e psicologia. No GA, são 114 trabalhos indexados, os quais foram citados por 1.499 autores. O artigo mais citado foi publicado no ano de 2008, sobre redes de

colaboração científica, pela revista *Scientometrics*, tendo sido citado por 179 pesquisadores.

Os autores publicaram seus primeiros artigos entre o final da década de 1970 e início dos anos 80. Leydesdorff publicou em torno de 72,5% a mais de artigos quando comparado a Glänzel e, em relação às citações, verifica-se que o primeiro recebeu 2,6 mais citações que o segundo.

No entanto, nota-se que Glänzel atua em áreas mais próximas, a saber: cientometria e teoria da probabilidade, ao passo que Leydesdorff atua em áreas mais diversificadas. Essa característica pode ter influenciado seu índice *h* no GA, pois essa ferramenta busca publicações científicas de vários tipos, tanto de acesso aberto como fechado e em diversas línguas.

## 5 Considerações finais

Nesta pesquisa, realizou-se uma análise comparativa entre as plataformas de dados *Scopus*, *WoS* e *Google Acadêmico*, objetivando verificar a consistência e a coerência entre os valores encontrados para o índice *h* de Hirsch.

O Teste de *Kruskal Wallis* foi utilizado para verificar se as distribuições dos indicadores *h* de Hirsch são iguais. Os dados evidenciaram que há similaridade entre as estimativas obtidas nas bases de dados internacionais *Scopus* e *WoS*. Desse modo, apesar da diferença numérica observada, essa não é significativa, considerando uma significância de 5%. Esse resultado garante que o uso de quaisquer das duas bases, *Scopus* ou *WoS*, identificará resultados semelhantes para o índice *h* para indivíduos que apresentem as mesmas características do universo estudado — pesquisadores de grande inserção internacional, considerados a elite científica da área.

Esse mesmo teste apontou que as estimativas dos índices *h* para o *Google Acadêmico* são superiores às das outras bases; essa característica pode estar relacionada ao modo de indexação dos trabalhos científicos nessas bases de dados, pois contabiliza artigos e livros. Além disso, como o GA busca referências disponíveis na *Web*, essa informação pode estar duplicada ou ter erros digitação, o que ocasionará uma superestimação da medida.

Como limitação da pesquisa, destaca-se o fato de não se poder estender os resultados para outras áreas do conhecimento e/ou pesquisadores com menor porte científico.

Dessa forma, sugere-se que a pesquisa em questão seja realizada com outros indicadores, bases de dados e se estenda a outros domínios científicos.

Finalizando, considera-se que esta pesquisa traz contribuições metodológicas para a Bibliometria, dada a carência de pesquisas relacionadas a essa temática.

## Referências

BAR-ILAN, J. Which h-index? – A comparison of WoS, Scopus and Google Scholar. **Scientometrics**, Dordrecht, v. 74, n. 2, p. 257–271, 2008.

BORNMANN, L.; MUTZ, R.; DANIEL, H-D. Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. **Journal of the American Society for Information Science and Technology**, New York, v. 59, n. 5, p. 830-837, 2008.

BRASIL. Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Portal de Periódicos. **Acervo: Journal Citation Reports (JCR)**. Brasília: CAPES, 23 out. 2009. Disponível em: <[https://www.periodicos.capes.gov.br/?option=com\\_pcollection&mn=70&smn=79&cid=94](https://www.periodicos.capes.gov.br/?option=com_pcollection&mn=70&smn=79&cid=94)>. Acesso em: 15 set. 2016.

CAREGNATO, S. E. Google acadêmico como ferramenta para os estudos de citações: avaliação da precisão das buscas por autor. **Ponto de Acesso**, Salvador, v. 5, n. 3, p. 72-86, dez. 2011. Disponível em: <[www.brapci.ufpr.br/brapci/v/a/11708](http://www.brapci.ufpr.br/brapci/v/a/11708)>. Acesso em: 20 jan. 2016.

EGGHE, L. The Hirsch index and related impact measures. **Annual review of information science and technology**, White Plains, v. 44, p. 65-114, 2010.

ELSEVIER. **Scopus**: Scopus is the largest abstract and citation database of peer-reviewed literature: scientific journals, books and conference proceedings. Disponível em: <<http://www.elsevier.com/solutions/scopus>>. Acesso em: 25 ago. 2015.

GUIMARÃES, J. A. C. Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de biblioteconomia do Mercosul: uma reflexão. In: ENCUESTRO DE EDIBCIC, 5., 2000, Granada. **La formación de profesionales e investigadores de la información para la sociedad del conocimiento: actas.** Granada: Universidad de Granada, Facultad de Biblioteconomía y Documentación, 2000, p. 206-216.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 102, n. 46, p. 16569-16572, Nov. 2005.

HIRSCH, J. E. Does the h index have predictive power? **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 104, n. 49, p. 19193-19198, Dec. 2007.

HJØRLAND, B. Domain analysis in information science: eleven approaches - traditional as well as innovative. **Journal of Documentation**, London, v. 58, n. 4, p. 422-462, Jan. 2002.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in Information Science: domain-analysis. **Journal of the American Society for Information Science**, Washington, v. 46, n. 6, p. 400-425, Jul. 1995.

LIMA, R. A.; VELHO, L. M. L. S.; FARIA, L. I. L. Bibliometria e “avaliação” da atividade científica: um estudo sobre o Índice h. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 3, p. 3-17, jul./set. 2012.

MARQUES, Fabrício. Os limites do índice-h: supervalorização do indicador que combina qualidade e quantidade da produção científica gera controvérsia. **Pesquisa Fapesp**, São Paulo, n. 207, p. 35-39, maio 2013.

MUGNAINI, R.; JANNUZZI, P. M.; QUONIAM, L. Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal. **Revista Ciência e Informação**, Brasília, v. 33, n. 2, p. 123-131, maio/ago. 2004.

OLIVEIRA, E. F. T; GRÁCIO, M. C. C. Indicadores bibliométricos em ciência da informação: análise dos pesquisadores mais produtivos no tema estudos métricos na base Scopus. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 16, n. 4, p. 16-28, out./dez. 2011.

PORTAL ACTION. **Manual da ferramenta Action Stat para o teste de Wilcoxon – Amostras Independentes**. Disponível em: <<http://www.portalaction.com.br/manual-nao-parametricos/teste-wilcoxon-mann-whitney-exemplo-amostras-independentes>>. Acesso em: 19 jan. 2016.

SHING, M.; KUMAR, S. Application of h and g indices to quantify scientific productivity of physicists at JNU, India. **Library Philosophy and Practice**, [Lincoln], Paper 1028, 21 p., Jan. 2014. Disponível em: <<http://digitalcommons.unl.edu/libphilprac/1028>>. Acesso em: 30 jan 2015.

SIEGEL, S. **Estatística não-paramétrica para as ciências do comportamento**. São Paulo: McGraw-Hill do Brasil, 1975. 350 p.

SILVA, D. Scopus e Web of Science – comparação. **A Biblioteca Informa**: boletim das bibliotecas da Universidade de Aveiro, Aveiro, n. 28, 11 mar. 2013. Disponível em: <<http://blogs.ua.pt/bibliotecainforma/?p=1859>>. Acesso em: 26 ago. 2015.

SILVA, D. D.; GRÁCIO, M. C. C. Indicadores bibliométricos de impacto: estudo comparativo entre os índices h e g aplicados aos pesquisadores mais produtivos na temática Estudos Métricos. In: ENCUESTRO IBÉRICO EDICIC, 7., 2015, Madrid. **Desafios y oportunidades de las Ciencias de la Información y Documentación en la era digital**. Madrid: Universidad Complutense, 2015. 10 p. Disponível em: <[http://edicic2015.org.es/ucmdocs/actas/art/230-Deolindo\\_indicadores-G-H.pdf](http://edicic2015.org.es/ucmdocs/actas/art/230-Deolindo_indicadores-G-H.pdf)>. Acesso em: 26 ago. 2015.

THOMAZ, P. G; ASSAD, R. S.; MOREIRA, L. F. P. **Uso do fator de impacto e do índice h para avaliar pesquisadores e publicações**. Arquivo Brasileiro de Cardiologia, São Paulo, v. 26, n. 2, p. 90-93, fev. 2011.

THOMSON REUTERS. **Web of science**: the most comprehensive and versatile research platform available. Disponível em: <<http://thomsonreuters.com/en/products-services/scholarly-scientific-research/scholarly-search-and-discovery/web-of-science.html>>. Acesso em: 25 ago. 2015.

VANZ, S. A. S.; STUMPF, I. R. C. Procedimentos e ferramentas aplicados aos estudos bibliométricos. **Informação e Sociedade**: estudos, João Pessoa, v. 20, n. 2, p. 67-75, maio/ago. 2010.

VIEIRA, P. V. M.; WAINER, J. Correlações entre a contagem de citações de pesquisadores brasileiros, usando o Web of Science, Scopus e Scholar. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 18, n. 3, p. 45-60, jul./set. 2013.

## Hirsch's h-index: comparative analysis between Scopus, Web of Science and Google Scholar databases

**Abstract:** Among the indicators to evaluate the scientific impact of publications, there is the Hirsch's h-index that, simultaneously, measures aspects related to the



production and the impact of the scientific production of a researcher, periodical, group, research center or country on a database. In this context, this research aims to get the indices h to 20 researchers considered representative in the theme “Metric Studies of Information”, on three distinct data sources (Scopus, Web of Science and Google Scholar), in order to identify and to analyze possible differences between the values of these indicators, examining the consistency of the indices in these information sources. To this end, the h-index for each investigator in the mentioned platforms was surveyed. By means of the Wilcoxon-Mann-Whitney test, using a 5% significance level, it was identified that the h-indices did not differ statistically between the Scopus and Web of Science databases. On the other hand, this index differed significantly in relation to the values obtained from Google Scholar.

**Keywords:** Information metric studies. H-index. Domain analysis. Databases.

Recebido em: 16/09/2016

Aceito em: 08/11/2016