The Journal of Academic Librarianship 43 (2017) 487–494

Contents lists available at ScienceDirect

The Journal of Academic Librarianship

journal homepage: www.elsevier.com/locate/jacalib

The *h*-index Debate: An Introduction for Librarians

Cameron Barnes[1]

University of New England, New South Wales 2351, Australia

A B S T R A C T

This article reviews the debate within bibliometrics regarding the *h*-index. Despite its popularity as a decision-making tool within higher education, the *h*-index has become increasingly controversial among specialists. Fundamental questions remain regarding the extent to which the *h*-index actually measures what it sets out to measure. Unfortunately, many aspects of this debate are confined to highly technical discussions in specialised journals. This article explains in simple terms exactly why a growing number of bibliometricians are sceptical that the *h*-index is a useful tool for evaluating researchers. It concludes that librarians should be cautious in their recommendations regarding this metric, at least until better evidence becomes available.

Introduction

University libraries face increasing demands from clients for research support services, particularly in the area of bibliometrics. Part of this trend has involved meeting the requests for quantitative indicators of research impact in connection with applications for tenure, promotion or grant funding. Answering this need is now an essential part of librarians' research support roles at many institutions (Drummond & Wartho, 2009; Corrall, Kennan, & Afzal, 2013; Olmeda-Gómez & de Moya-Anegón, 2016). However, recent research has revealed widespread concerns among librarians regarding gaps in "knowledge, skills and competence" in relation to bibliometrics (Corrall et al., 2013, p. 636). This article assists in some small measure to help close this gap by providing librarians with a guide to the debate within bibliometrics over the *h*-index.

What is the *h*-index?

In 2005, Jorge Hirsch proposed a new measure of research impact (Hirsch, 2005). A researcher's *h*-index is the number of his or her papers (*h*) that have been cited at least (*h*) times. In simple terms, a researcher with 10 published articles, each of which has received at least 10 citations, has a *h*-index of 10. Hirsch expressed a hope that his new index "may provide a useful yardstick to compare different individuals competing for the same resource when an important evaluation criterion is scientific achievement" (2005, p. 16572).

The influence of the *h*-index on bibliometrics

In hindsight, the speed with which Hirsch's proposal attracted support was little short of astonishing. Hirsch posted a draft of his paper as a pre-print on the arXiv server in early August 2005. Within weeks, his idea had been praised in separate articles in *Nature* and *Science* (Ball, 2005; Bhattacharjee, 2005). A revised draft arrived with the editors of the *Proceedings of the National Academy of Sciences* by the middle of the same month. This was rapidly accepted for publication, appearing in the November issue of the journal (Hirsch, 2005).

Hirsch's paper received an immediately favourable response. Within a few years, it was apparent that the *h*-index "has taken the world of research assessment by storm" (Tol, 2008, p. 149). Over the last decade, Hirsch's paper has attracted thousands of citations: currently over four thousand according to Google Scholar (Mingers & Leydesdorff, 2015). Hirsch's article now one of most highly cited papers in the history of bibliometrics (if not the most highly cited), and is recognised as "a veritable citation classic" (Jones, 2016, p. 535).

This outcome is all the more unexpected because Hirsch was not a bibliometrican. Instead, he was a professor of physics at the University of California, San Diego. Hirsch admitted that the success of his idea came as a surprise. In an interview, Hirsch stated that at the time he had "not worked in bibliometrics before and was not totally familiar with the literature on the subject" (Hampton, 2009). For an individual from outside the field to have such an impact with their first paper is unprecedented (Bornmann, 2014).

Many bibliometricians have hailed the *h*-index as a major contribution to the discipline. Some have even gone so far as to divide the history of bibliometrics into pre- and post-Hirsch epochs (Bartneck & Kokkelmans, 2011; Prathap, 2010; Todeschini & Baccini, 2016). Even those who would hesitate to go this far would agree that Hirsch's paper has "resulted in a new research front in bibliometrics" (Bornmann, 2014, p. 749). Citations

E-mail address: cbarnes@une.edu.au.
[1] Postal address: University Library, University of New England, New South Wales 2351, Australia.

http://dx.doi.org/10.1016/j.acalib.2017.08.013
Received 7 August 2016; Received in revised form 16 July 2017; Accepted 24 August 2017
Available online 12 September 2017

to Hirsch paper in bibliometric journals continue to increase, with no end to this trend in sight (Bornmann, 2014).

Although originally intended to compare individuals in terms of their research impact, the *h*-index was soon put to a range of uses. Less than twelve months after Hirsch's paper was published, the first studies appeared extending the use of his metric to subjects as varied the measurement of journals, scientific topics and even chemical compounds (Banks, 2006; Braun, Glänzel, & Schubert, 2006; Saad, 2006). Since then, the *h*-index been employed in hundreds of studies to assess the impact of research groups, university departments, institutions, and even whole nations (e.g. Jacsó, 2009; Lazaridis, 2010; Prathap & Gupta, 2009).

In retrospect, it is easy to understand some of the reasons for the *h*-index's rapid success. The metric arrived at a time of intense academic dissatisfaction with the misuse of the Thomson Reuters (now Clarivate) Journal Impact Factor (JIF) as a guide to the quality of individual papers (Monastersky, 2005; Seglen, 1997; Seglen, 1998). In many countries, the JIF is routinely used by selection committees and granting agencies to evaluate the work of researchers, despite repeated expert warning against the practice, even from one of the metric's creators (e.g. Garfield, 1979; Garfield, 1999: Garfield, 2006). Another reason for the *h*-index's rapid success was its ease of calculation: “the great advantage of the index is that you can get it in about 30 seconds” (Bhattacharjee, 2005, p. 1181).

The *h*-index outside bibliometrics

During the last decade, the *h*-index has enjoyed increasing vogue as a decision-making tool within higher education. The use of a single, easily calculated number to express a researcher's contribution to his or her field has proven extremely attractive to many university administrators. There are claims that the *h*-index has become the “most popular quantitative measure of a researcher's productivity and impact” (Penner, Petersen, Pan, & Fortunato, 2013, p. 8). One recent commentator has even gone so far as to refer to the *h*-index's “rapid, near-universal adoption within the scientific community” (Dienes, 2015, p. 393). What is true is that many universities in Europe and North America now routinely require applicants for academic positions to provide their *h*-indexes. Institutions as diverse as the Universiteit Utrecht and the University of Ohio have gone so far as to set minimum *h*-indexes as criteria for appointment for specific academic ranks (Barnes, 2014).

The growing importance of the *h*-index in different contexts is illustrated by its use by grant administrators and government bodies. Recent examples include:

- Indian government funding of local universities based on their *h*-index scores.
- Employment by the Czech Science Foundation of applicants' *h*-index scores in determining the outcome of funding proposals,
- Use of the *h*-index as an indicator for inclusion in the Italian *Abilitazione scientifica* (the list of academics eligible for promotion to associate or full professor) (Barnes, 2014).

The rising influence of the *h*-index has generated growing unease among academics (Burrows, 2012). However, the consensus among many observers seems to be “let's face it: the *h* index is here to stay” (Schreiber, 2007, p. 652).

The growth in *h*-index variants

No citation-based measure of research impact is perfect (Phelan, 1999; Seglen, 1997). The *h*-index is no exception, as Hirsch acknowledged from the beginning (Hirsch, 2005). Some of the best-known shortcomings of the *h*-index include:

- Authors can relatively easily inflate their *h*-index scores through excessive self-citation (Bartneck & Kokkelmans, 2011; Schreiber, 2007; Zhivotovsky & Krutovsky, 2008).
- The *h*-index does not adjust for multiple-authorship. As a result, the metric is said to underestimate the relative contribution of researchers who publish alone or with a few co-authors, and overestimate that of investigators who publish as part of a large team (Burrell, 2007a; Schreiber, 2008).
- Most bibliometricians regularly normalise citation metrics to allow for differences in citation practices between disciplines. However, the *h*-index makes no allowance for this factor (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Batista, Campiteli, & Kinouchi, 2006).
- The *h*-index is not tied to any particular source by definition (such as the JIF). This means that the use of different databases will result in diverse *h*-indexes for the same researcher (Jacsó, 2008).

None of these shortcomings are unique to the *h*-index. However, observers have pointed out one flaw that may be specific to the metric. The *h*-index is said to favour mediocre researchers at the expense of those who publish fewer, but more significant papers (Costas & Bordons, 2007). In particular, the *h*-index is often seen as inherently unfair to early career researchers. The reason is that an individual's *h*-index cannot exceed his or her total number of published papers. Regardless of the number of the citations which their papers attract, even most brilliant early career researchers is likely to have a lower *h*-index than older colleagues (Kelly & Jennions, 2006; Rousseau & Leuven, 2008).

The main effect of these shortcomings has been to encourage an explosion of *h*-index variants. By 2011, there were at least 37 of these (Bornmann, Mutz, Hug, & Daniel, 2011). This total had risen to more than 50 three years later (Bornmann, 2014; Smith, 2015). New specimens continue to be added to the “*h*-index zoo” on a regular basis (e.g. Wan, 2014; Zhang, 2013).

Many of these new metrics are mathematically complex. Others, such as an early proposal to scale the *h*-index to allow comparisons between researchers in different disciplines, are relatively simple (Iglesias & Pecharromán, 2007). None have displaced Hirsch's original metric in terms of popularity. Part of the reason is that most variants are so highly correlated with Hirsch's original index as to be largely redundant (Bornmann, 2012; Bornmann et al., 2011).

Hirsch himself has advanced improvements to his original metric. In his 2005 paper, Hirsch proposed the *m*-quotient, which was the *h*-index divided by the number of years since a researcher's first publication (Hirsch, 2005). This metric allows comparisons to be made between researchers at different stages in their careers. In 2010, Hirsch made another suggestion. This involved weighting an individual's *h*-index to adjust for differences in the number of his or her co-authors (Hirsch, 2010). Neither of these suggestions have been widely adopted.

Aspects of the growth in *h*-index variants have aroused some disquiet. Waltman and van Eck (2012) point out that the creation of new *h*-index variants usually proceeds in a surprisingly ad hoc manner. Typically, the inventor of a new variant begins by identifying one of *h*-index's well-known shortcomings. He or she then proposes a (usually) simple modification of the *h*-index which adjusts for this factor. The new indicator is then justified on the grounds that its use produces results which appear to “be intuitively reasonable” (Waltman & van Eck, 2012, p. 406). In the process, the researcher rarely asks any difficult questions. The underlying validity of the original metric is simply assumed. The same problem is apparent in the many studies which extend the *h*-index beyond its original purpose. The authors of such papers almost never question if Hirsch's metric is fit for its new purpose (Abramo, D'Angelo, & Viel, 2013).

The *h*-index debate

Despite its popularity as a research topic, attitudes toward the *h*-

index among bibliometricians are sharply divided. There is a trend among some bibliometricians to condemn the metric outright. Bornmann has declared that:

> Nowadays, the *h*-index is used primarily by those who are not bibliometrics specialists. It appears therefore that a similar fate to that of the Journal Impact Factor awaits the *h*-index … (2014, p. 750).

Such sweeping assertions that the metric has been rejected by expert bibliometricians must be taken with a grain of salt. Notable "bibliometrics specialists" continue to publish on the *h*-index. Even Bornmann was among their number until recently (Bornmann & Marx, 2011; Bornmann, Marx, Gasparyan, & Kitas, 2012).

For those bibliometricans who regard the metric with increasing scepticism, the issue is not the *h*-index's widely recognised shortcomings. Instead, the problem is much more fundamental. An increasing number of bibliometricians are now convinced that the construction of the *h*-index is inherently arbitrary. For these researchers, the popularity of the *h*-index is evidence that

> Even among actual evaluation practitioners, simplicity and rapidity can prevail over rigor and reliability (Abramo et al., 2013, p. 556).

This issue is rarely raised when the *h*-index is discussed outside the specialist literature. Recent summaries of the *h*-index aimed at librarians do little more than list the metric's well-known shortcomings (e.g. Suiter & Moulaison, 2015). Articles on research support services at university libraries typically give no hint of the extent to which the credibility of the *h*-index is under question (Drummond & Wartho, 2009; Keller, 2015).

What does the *h*-index measure?

The first step in understanding the debate over the *h*-index is to determine exactly what the metric is designed to do. In his original article, Hirsch stated that his metric was designed to measure the "broad impact of an individual's work" or "overall scientific impact" (Hirsch, 2005, p. 16569).

Hirsch returned to this point more recently. He has made clear that the main purpose of the *h*-index is to serve as "an indicator of the impact of a researcher on the development of his or her scientific field" (Hirsch & Buela-Casal, 2014, p. 163).

A widely misunderstood metric?

Despite these statements, the purpose of the *h*-index is often misunderstood. Many observers have assumed that the *h*-index is designed to simultaneously measure both the quality (impact) and quantity (number) of an individual's publications (e.g. Bornmann & Marx, 2014; Egghe, 2010; Gingras, 2014; Kulasegarah & Fenton, 2010; Panaretos & Malesios, 2009). This assumption is based on the observation that an:

> individual's h-index score is determined by the "number of citations" (y) versus "paper number" (x) curve and the y = x line, which leads to an x-shaped graph. (Hirsch & Buela-Casal, 2014, p. 162).

Although understandable, such an interpretation is not at all what Hirsch intended. He has stated that the *h*-index

> originates from the assumption that the number of citations received by a scientist is a better indicator of the relevance of his or her work than the number of papers he or she publishes or the journals where they are published. The fact that every paper published has its own number of citations implies having many numbers for each scientist. I had the idea of developing the h-index as a way of condensing all that information into one single number to facilitate comparisons between scientists (Hirsch & Buela-Casal, 2014, p. 162).

Here Hirsch indicates that the construction of the *h*-index is designed to simplify the process of comparing researchers on the basis of the different citation counts for each of their papers. Exactly how Hirsch intended to achieve this goal is explained later in this paper. The important thing is that he does not believe that the number of papers should be an important factor in measuring a researchers' impact on their field. This point is particularly relevant when examining how the *h*-index behaves in practices.

What is impact?

If the *h*-index measures impact, then what is impact? Unfortunately, bibliometricians rarely define the term. Most researchers in the field are quite happy to use the word impact dozens of times in the same paper without ever venturing even a tentative definition (e.g. Leydesdorff & Bornmann, 2011). Within bibliometrics, impact is best understood as the usefulness of an academic publication in the production of future scientific outputs, which typically take the form of papers.

This point was expressed decades ago by Garfield. Garfield observed that:

> People talk about citation counts as being a measure of the "importance", or "impact" of scientific work, but those who are knowledgeable about the subject use the words in a very pragmatic sense: what they really are talking about is utility (1979, p. 363).

This point has since become axiomatic among most bibliometricians. Since the work of Martin and Irvine (1983), expert opinion has drawn a distinction between a paper's impact (as measured by citations) and its importance or quality. The latter are characteristics which cannot be directly measured through citation data in isolation (Martin & Irvine, 1983).

With these points in mind, it is now time to look more closely at the debate within bibliometrics regarding the *h*-index. The next part of this paper will look at the reasons why a growing number of bibliometricians have questioned whether the *h*-index is in fact a valid measure of research impact. To do so, it is necessary to look first at the evidence for the *h*-index contained in Hirsch's original paper.

Hirsch's case for the *h*-index

In his 2005 paper, Hirsch provided two types of evidence for the *h*-index. The first was purely theoretical. Hirsch proposed a simple model of a researcher's publication history. According to this model, a researcher's total citations increase over time in a linear fashion. As this linear increase continues, the researcher's *h*-index rises in proportion. On this basis of this association, Hirsch argued that the *h*-index could be used to compare different researchers in terms of their cumulative impact (Hirsch, 2005).

Hirsch's argument has been criticised on a number of grounds. His model assumes that: "the researcher publishes *p* papers per year and that each published paper earns *c* new citations per year every subsequent year" (Hirsch, 2005, p. 16570). The problem is that these assumptions are not particularly realistic (Bornmann & Daniel, 2009; Burrell, 2007b; Glänzel, 2006; Jensen, Rouquier, & Croissant, 2009). The issue for Hirsch's critics is not that there are exceptions to each of these assumptions, a point which Hirsch conceded (Hirsch, 2005). Their point is that such exceptions *are* the rule.

Contrary to Hirsch's model, few researchers are constant performers (Bornmann & Daniel, 2009). The typical scientist has a few highly cited papers and many more rarely cited ones (Anderson, Hankin, & Killworth, 2008; Bornmann & Daniel, 2009; Cerchiello & Giudici, 2014; Seglen, 1992). Age-and status-related effects also typically have a marked effect on both publication frequency and research impact. Hirsch's assumption of stability over the course of a researcher's career runs counter to the results of decades of research (Abramo, D'Angelo, & Murgia, 2016; Cole, 1979; Fox, 1983; Hall, Mairesse, & Turner, 2007; Jensen et al., 2009; Levin & Stephan, 1991). In addition, articles do not typically accumulate citations at a constant rate. Instead, the new citations to papers tend to taper off rapidly over time (Burrell, 2007b; Jensen et al., 2009; Kelly & Jennions, 2006).

Partly because of the technical challenges involved, there has been little empirical work on Hirsch's model. However, two studies have examined the idea that the *h*-index increases linearly over time. The first of these looked specifically at the publication careers of eleven physicists mentioned in Hirsch's (2005) paper (Liang, 2006). This analysis showed a range of patterns: only one of the physicists demonstrated the expected linear increase (Liang, 2006). Liang's conclusions were confirmed in a later paper, which examined the *h*-index sequences of Nobel Prize winners in Medicine, Chemistry and Economics. The authors of this article identified five different *h*-index sequences, which they termed convex, concave, S-shaped, IS-shaped and linear. Although the *h*-indexes of some Nobel laureates increased in a linear fashion, as Hirsch's model predicted, these researchers were in the minority (Wu, Lozano, & Helbing, 2011).

The problems with Hirsch's model are rarely discussed. However, they are critical to the question of the metric's validity. If his model does not in fact correspond to reality, then there is no reason to assume a fixed relationship between an individual's *h*-index and their total number of citations. In this case, there would no reason to expect that comparison between researchers in terms of their *h*-index scores would lead to meaningful results.

Hirsch's empirical evidence for the *h*-index

Hirsch's second type of evidence for the *h*-index was empirical. First, he pointed out that many leading physicists had high *h*-indexes. Hirsch noted, for example, that:

> The highest h among physicists appears to be E. Witten's h, which is 110. That is, Witten has written 110 papers with at least 110 citations each (2005, p. 16569).

Here Hirsch makes reference to Witten's popular reputation "as the greatest theoretical physicist in the world" (Lemonick, 2004, p. 100). However, Hirsch recognised that the fact that Witten (and other prominent) physics had high *h*-indexes provided only anecdotal evidence for his metric's usefulness. He therefore proposed a more rigorous test. Hirsch argued that it was possible to show that an individual's *h*-index score was a good indicator of whether a scientist has won honours, such as the Nobel Prize or membership in the National Academy of Science (NAS).

As evidence, Hirsch calculated the *h*-indexes for a group of Nobel Laureates in Physics over the previous twenty years. He found that 84% of this group had an *h*-index of at least 30. Hirsch then looked at the *h*-index scores for the newly elected members of the NAS in Physics and Astronomy. According to his calculations, the median *h*-index of the members of this second sample was 46. On this basis, Hirsch claimed that:

> These examples further indicate that the index h is a stable and consistent estimator of scientific achievement (Hirsch, 2005, p. 1657).

Hirsch's argument has been accepted by a number of commentators (e.g. Aragón, 2013; Panaretos & Malesios, 2009; Woeginger, 2008), but there is a flaw in his logic. It has been observed that:

> One can conclude that it is likely a scientist has a high h-index given the scientist is a Nobel Laureate. But without further information, we know very little about the likelihood someone will become a Nobel Laureate or a Member of the National Academy, given that they have a high h-index. That is the kind of information one wants in order to establish the validity of the h-index (Adler, Ewing, & Taylor, 2008, p. 13).

Moreover, there are other problems with Hirsch's case for his metric's predictive power. The *h*-indices of the Nobel Prize winners in Hirsch's sample ranged from 22 to 79. Not only is this an extremely wide range, there is the additional problem there are very many physicists with no chance of winning the Nobel Prize who have *h*-indexes far higher than 22. The overlap between Hirsch's two sample populations in terms of their *h*-indexes was also very high. In terms of the range, means and standard deviations, the Nobel Prize winners and the members of the NAS were almost identical. However, this is not what we should expect if the *h*-index is a valid measure of impact. Nobel medallists are a far more elite group than members of the NAS. The two groups are worlds apart in terms of their scientific reputation (Barnes, 2014).

The predictive validity of the *h*-index

Two years later, Hirsch presented further empirical evidence for the *h*-index. He argued that the *h*-index has "the highest ability to predict *future* academic achievement" (Hirsch, 2007, p. 19195). As proof, Hirsch looked at the publication history of two readily available groups of physicists. He reported that:

- researchers in his samples with a high *h*-index 12 years after first publication were likely to have a high *h*-index after 24 years; and
- the *h*-index was better than total number of articles and total number of citations to articles in predicting its future cumulative value.

He concluded therefore that his index "can be profitably used in academic appointment processes and to allocate research resources" (Hirsch, 2007, p. 19198).

Again, Hirsch's conclusion has been accepted by a number of commentators (e.g. Alonso et al., 2009; Panaretos & Malesios, 2009). However, there are claims that his argument suffers from a serious flaw. The problem is that "the increase of the *h*-index does not necessarily depend on the factual performance for several years in the future, but is more likely to result from previous, often rather old publications" (Schreiber, 2013a, p. 327). Schreiber has provided examples where the growth of a researcher's *h*-index is the "same, irrespective of whether the investigated researcher had performed as he or she did or whether he (she) had not published any further work" (2013a, p. 325). This conclusion applies even in relation to Hirsch himself:

> If Hirsch had stopped working in 2001, his index would have been unaffected in 2010 and even in 2012 deviate only by one index point. From 2005 onwards no change would have resulted except a deviation of one index point in the year 2009 (Schreiber, 2013a, p. 327).

Another problem is that Hirsch fails to take into account the fact that an individual's *h*-index cannot decline. For that reason, there are claims that Hirsch's test of the predictive power of his metric lacks meaning (García-Pérez, 2013; García-Pérez & Núñez-Antón, 2013). It has been observed that:

> cumulative non-decreasing measures like the h-index contain intrinsic autocorrelation, resulting in significant overestimation of their "predictive power" (Penner et al., 2013, p. 1).

More recent attempts to the show that the *h*-index predicts future scientific achievement (Acuna, Allesina, & Kording, 2012) have been criticised on identical grounds (García-Pérez, 2013; Penner et al., 2013).

The construction of the *h*-index

In bibliometric terms, the construction of the *h*-index is unique. As Egghe has pointed out, the "*h*-index is not an average, not a percentile, not a fraction: it is a totally new way of measuring performance …" (2010, pp. 66–67). One of Hirsch's critics has gone so far as to refer to the "caractère improvise" of the *h*-index (Gingras, 2015, p. 9).

Waltman and van Eck point out that Hirsch's model does not require the *h*-index to be defined as it is. They state that:

> For instance, the *h*-index could equally well have been defined as

follows: A scientist has an *h*-index of *h* if *h* of his publications each have at least 2*h* citations and his remaining publications each have fewer than 2(*h* + 1) citations. Or the following definition could have been proposed: A scientist has an *h*-index of *h* if *h* of his publications each have at least *h*/2 citations and *h* is remaining publications each have fewer than (*h* + 1) / 2 citations (Waltman & van Eck, 2012, p. 408).

Other commentators question Hirsch's decision to graph citations and numbers of papers in this manner in the first place (Schreiber, 2013b). Hirsch's decision to do so depends entirely on his abstract model of the typical researcher's publication history, a model which Hirsch himself describes as the "simplest possible" (Hirsch, 2005, p. 16569). There are also claims that the construction of the *h*-index is simply illogical. The *h*-index is said to assume an

> equality between incommensurable quantities. An author's papers are listed in an order of decreasing citations with paper i having C(i) citations. Hirsch's index is determined by the equality, h = C(h), which posits an equality between two quantities with no evident logical connection (Lehmann, Jackson, & Lautrup, 2008, p. 377).

As a result of such arguments, a growing number of observers have concluded that the construction of the *h*-index is in fact arbitrary (Bornmann, 2014; Franceschini & Maisano, 2010; Gingras, 2014; Glänzel, 2006; Lehmann et al., 2008; Petersen & Succi, 2013; Schreiber, 2013b; Waltman & van Eck, 2012).

How useful is the *h*-index in practice?

Other bibliometricians have found fault in the *h*-index in terms of its low information content. In order to calculate a researcher's *h*-index, it is necessary to discard almost all the evidence for his or her impact. The problem is that:

> An author's *h*-index cannot exceed his/her number of publications and will usually be considerably less. Thus, the vast majority of the hundreds or even thousands of citations that accompany the most highly cited papers effectively contribute zero … Moreover, articles that have received many citations, but which fall just short of the number required to score for *h* … also count for nothing in the sense that *h* is not affected by them (Anderson et al., 2008, p. 578).

In this context, Bornmann points out that:

> For decades, bibliometrics has developed generally accepted methods which identify the significant publications in a publication set (such as those publications, for example, which are among the 10% most cited publications in their subject area and publication year). The benchmark for significant publications should therefore not result from the publication set of a scientist, as is the case with the *h* index; the same benchmark should be used for the publications of all scientists to evaluate whether a publication should be designated significant or not … (2014, p. 749).

The insensitivity of the *h*-index to highly-cited articles has also been grounds for criticism. Hirsch defends this insensitivity on the grounds that citations to such articles:

> may be inflated by as small number of "big hits", which may not be representative of an individual if he or she is a co-author with many other on these papers (Hirsch, 2005, p. 16569).

The insensitivity of the *h*-index to highly cited articles is sometimes described as one of the metric's strengths (Batista et al., 2006). However, it is not entirely clear why this insensitivity is so desirable in a metric intended to compare individuals in terms of their research impact. For decades, highly-cited papers have been seen as the most obvious evidence of an individual's impact on his or her field (Aksnes, 2006; Anderson et al., 2008).

Hirsch's argument seems to assume that "big hits" are an unusual feature in an individual researcher's publication history. However, this seems not to be the case. As pointed out earlier, "publication and citation data have a right-skewed distribution" (Bornmann & Daniel, 2009, p. 3). Most successful researchers, even those who write without co-authors, have a few highly cited papers and many more rarely cited ones (Bornmann & Daniel, 2009; Cerchiello & Giudici, 2014; Seglen, 1992). Hirsch himself is a case in point. He is the sole author of his own "big hit": his 2005 article on the *h*-index.

Some statisticians have gone so far as to state that confidence in the *h*-index involves a willing suspension of disbelief. Hirsch proposed that:

> two individuals with similar *h*s are comparable in terms of their overall scientific impact, even if their total number of papers or their total number of citations is very different (2005, p. 16569).

However, it is not at all difficult to think of reasons why this conclusion might not hold in reality. In their report on citation statistics, Adler and his co-authors suggest the following thought experiment:

> Think of two scientists, each with 10 papers with 10 citations, but one with an additional 90 papers with 9 citations each; or suppose one has exactly 10 papers of 10 citations and the other exactly 10 papers of 100 each. Would anyone think them equivalent? (Adler et al., 2008, p. 13).

A number of bibliometricians have argued that the *h*-index is fact highly inconsistent in its ranking of researchers (Bornmann & Daniel, 2009; Gaster & Gaster, 2012; Marchant, 2009; Vinkler, 2007; Waltman & van Eck, 2012). The issue for the metric's critics is not just that the *h*-index cannot "distinguish ground-breaking scientific papers from more conventional scientific studies" (Gaster & Gaster, 2012, p. 630). Instead, the complaint is that the *h*-index cannot even "discriminate among average scientists" (Jin, Liang, Rousseau, & Egghe, 2007, p. 856). As a result, there are claims of an internal debate within bibliometrics "over whether measurement of *h*-index at the individual level is ethical" (Hunter, 2013).

Evidence for convergent validity

From the beginning, one of the criticisms of the *h*-index was that empirical evidence in its support was lacking. Early on, there were claims that there "has been little serious analysis" (Adler et al., 2008, 13). Almost all of the empirical research has been limited to attempts to show convergent validity, the extent to which the *h*-index correlates with other measures of impact. To date, the results of these studies have been inconclusive.

Most of the early studies in this area looked at the association between the *h*-index and citations (Bornmann, Wallon, & Ledin, 2008; Costas & Bordons, 2007; Cronin & Meho, 2006; Saad, 2006; Saad, 2010; van Raan, 2006). These papers revealed a consistent correlation. However, it was soon realised that such correlations were "unremarkable" (Adler et al., 2008, p. 13). It was pointed out that:

> Since the h index combines number of publications and citations counts in one single index, very large correlation coefficients between the measures are not surprising (Bornmann et al., 2008, p. 155).

As a result, the case for the convergent validity of the *h*-index is now seen to rest on the apparent correlation between the *h*-index and peer evaluations (Bornmann & Marx, 2011; Bornmann, Mutz, & Daniel, 2010). Studies have looked at this association for the following groups:

(a) applicants for fellowships through the German Boehringer Ingelheim Fonds (BIF) (Bornmann & Daniel, 2005).
(b) applicants for the European Molecular Biology Organization (EMBO) Long Term Fellowships (LTF) and Young Investigator (YI) programmes (Bornmann et al., 2008).

(c) South African botanists and zoologists rated by the National Research Foundation (NRF) (Lovegrove & Johnson, 2008).
(d) 147 chemistry research groups in Europe (van Raan, 2006).
(e) French researchers within the Centre national de la recherche scientifique (CNRS) (Jensen et al., 2009).

In each case, the authors reported statistically significant correlations between *h*-index scores and peer assessments.

These and similar studies have often been seen as evidence for the convergent validity of the *h*-index (Bornmann et al., 2010; Bornmann & Marx, 2011). In practice, however, the *h*-index appears to be a poor predictor of future career success. Vinkler (2007) points out that the difference in means between successful and unsuccessful applicants for BIF fellowships was extremely small. He observes that

> The difference in means given is, however, not significant in each year at $p < 0.05$ level. For example in 1992: the mean *h*-index for researchers awarded was 2.92 ($n = 13$; SD 2.29) and for researchers rejected: 2.70 ($n = 57$; SD 2.17), significance between the two means was: $p = 0.80$. In 1994: 2.83 ($n = 12$; SD = 1.27) and 2.46 ($n = 52$; SD = 2.11), respectively; $p = 0.56$ (Vinkler, 2007, p. 490).

Similar criticisms can be made of the conclusion drawn by Bormann and his team from their study of EMBO selection committee ratings. In their paper, the correlation coefficient between *h*-index scores and committee ratings were $r = 0.21–0.28$ (Bornmann et al., 2008). However, these figures indicate that *h*-index scores are extremely poor predictors of the final ratings, as they imply that *h*-index scores explained as little as 4–7% of total variability.

Other studies report ambiguous evidence. In the case of the ratings assigned to 147 chemistry research groups in Europe, differences in *h*-indexes explained only 21% ($r^2 = 0.2161$) of the final scores (van Raan, 2006). Differences in *h*-indexes explained 37% ($r^2 = 0.376$) of the variance in ratings assigned to botanists and zoologists by the South African NRF (Lovegrove & Johnson, 2008). In neither of these studies did *h*-index scores explain anywhere near half the total variance, suggesting that the *h*-index has extremely poor predictive power in these contexts.

The best results were reported by the team led by Jensen in their study to test the value of different bibliometric indicators as predictors of career success. Overall, *h*-index data explained 48% of promotion decisions within the French CNRS (Jensen et al., 2009). However, as Jensen and his co-authors admitted, this figure was far from a ringing endorsement of the *h*-index. The metric provided no more than "the least bad correlations" (Jensen et al., 2009, p. 467).

Taken together, these papers provide only weak evidence at best for the convergent validity of the *h*-index. Rather than demonstrate convergent validity, their results are equally explicable by a simple hypothesis. Researchers who publish only a few, rarely cited papers will have low *h*-indices. Such individuals are also unlikely to receive fellowships, be promoted or be highly regarded as researchers by their peers (Barnes, 2014).

Moreover, these studies do not directly address the question of whether the *h*-index is fit for the purpose for which it is designed. Comparisons based on aggregated figures (such as means) are unlikely to demonstrate that the metric is "a useful yardstick to compare different individuals" as Hirsch concluded (Hirsch, 2005, p. 16572). If anything, the weakness of the evidence would suggest reasons for caution.

Does the *h*-index behave as intended?

There are claims that the *h*-index behaves in a manner contrary to its author's intentions. Earlier in this paper, it was observed that Hirsch has stated that his index:

> originates from the assumption that the number of citations received by a scientist is a better indicator of the relevance of his or her work than the number of papers he or she publishes or the journals where they are published (Hirsch & Buela-Casal, 2014, p. 161).

Despite Hirsch's intention, there are claims that "the *h*-index is in fact highly correlated with the number of published papers and thus essentially determined by it" (Gingras, 2014, p. 118). The reason is that a researcher's *h*-index cannot exceed his or her total number of publications, no matter how many citations any individual paper receives. As a result, the effective limit on many researchers' *h*-indexes is their number of published papers, not the citations these papers attract.

This criticism was first raised over a decade ago (Kelly & Jennions, 2006, p. 169). Other researchers have made similar findings (Bornmann et al., 2008; Gaster & Gaster, 2012; Kulasegarah & Fenton, 2010; Van Leeuwen, 2008). The most striking result has come from a recent paper on the publications history of 248 Danish professors in the health sciences. This study found that the correlation between *h*-index and number of papers for these researchers was as high as $r = 0.93$ (Gaster & Gaster, 2012, p. 830).

Discussion

This paper has sought to unpack the debate over the *h*-index. It has looked closely at the reasoning behind the metric's construction, and the reasons for the growing disquiet among bibliometricians in relation to this point. As part of this investigation, this article has examined both Hirsch's case for his metric and the empirical research carried out over the last ten years. The results have shown that the case for the *h*-index is much less than its current popularity would suggest. There is some weak evidence that this metric might have convergent validity, but in most other respects, firm evidence seems to be lacking or even negative.

Conclusion

As part of their research support role, librarians employed at universities and colleges are expected to be specialists in bibliometrics. To maintain their expertise, research librarians will need to ask questions and become familiar with current controversies. The *h*-index is one area where librarians should be encouraged to avoid complacency. Despite growing criticism from many bibliometrics, the *h*-index may well demonstrate its value in the future. The metric remain a popular topic of ongoing bibliometric research. In the meantime, librarians would be well advised to adopt a cautious attitude toward the *h*-index and its variants in their conversations with academic and professional staff in their institutions.

References

Abramo, G., D'Angelo, C., & Murgia, G. (2016). The combined effects of age and seniority on research performance of full professors. *Science and Public Policy, 43*(3), 301–319. http://dx.doi.org/10.1093/scipol/scv037.

Abramo, G., D'Angelo, C. A., & Viel, F. (2013). The suitability of h and g indexes for measuring the research performance of institutions. *Scientometrics, 97*(3), 555–570. http://dx.doi.org/10.1007/s11192-013-1026-4.

Acuna, D. E., Allesina, S., & Kording, K. P. (2012). Predicting scientific success. *Nature, 489*(7415), 201–202.

Adler, R., Ewing, J., & Taylor, P. (2008). *Citation statistics: A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS).* Berlin: International Mathematical Union.

Aksnes, D. W. (2006). Citation rates and perceptions of scientific contribution. *Journal of the American Society for Information Science and Technology, 57*(2), 169–185. http://dx.doi.org/10.1002/asi.20262.

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). H-index: A review focused on its variants, computation and standardization for different scientific fields. *Journal of Informetrics, 3*(4), 273–289. http://dx.doi.org/10.1016/j.joi.2009.04.001.

Anderson, T. R., Hankin, R. K. S., & Killworth, P. D. (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. *Scientometrics, 76*(3), 577–588. http://dx.doi.org/10.1007/s11192-007-2071-2.

Aragón, A. M. (2013). A measure for the impact of research. *Scientific Reports, 3*, 1649. http://dx.doi.org/10.1038/srep01649.

Ball, P. (2005). Index aims for fair ranking of scientists. *Nature, 436*(7053), 900. http://dx.doi.org/10.1038/436900a.

Banks, M. G. (2006). An extension of the Hirsch index: Indexing scientific topics and compounds. *Scientometrics, 69*(1), 161–168. http://dx.doi.org/10.1007/s11192-006-0146-5.

Barnes, C. (2014). The emperor's new clothes: The *h*-index as a guide to resource allocation in higher education. *Journal of Higher Education Policy and Management, 36*(5), 456–470. http://dx.doi.org/10.1080/1360080X.2014.936087.
Bartneck, C., & Kokkelmans, S. (2011). Detecting *h*-index manipulation through self-citation analysis. *Scientometrics, 87*(1), 85–98. http://dx.doi.org/10.1007/s11192-010-0306-5.
Batista, P. D., Campiteli, M. G., & Kinouchi, O. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics, 68*(1), 179–189. http://dx.doi.org/10.1007/s11192-006-0090-4.
Bhattacharjee, Y. (2005). Impact factor. *Science, 309*(5738), 1181.
Bornmann, L. (2012). Redundancies in *h* index variants and the proposal of the number of top-cited papers as an attractive indicator. *Measurement: Interdisciplinary Research and Perspectives, 10*(3), 149–153. http://dx.doi.org/10.1080/15366367.2012.716255.
Bornmann, L. (2014). *h*-index research in scientometrics: A summary. *Journal of Informetrics, 8*(3), 749–750. http://dx.doi.org/10.1016/j.joi.2014.07.004.
Bornmann, L., & Daniel, H.-D. (2005). Does the h-index for ranking of scientists really work? *Scientometrics, 65*(3), 391–392. http://dx.doi.org/10.1007/s11192-005-0281-4.
Bornmann, L., & Daniel, H.-D. (2009). The state of *h* index research: Is the *h* index the ideal way to measure research performance? *EMBO Reports, 10*(1), 2–6. http://dx.doi.org/10.1038/embor.2008.233.
Bornmann, L., & Marx, W. (2011). The h index as a research performance indicator. *European Science Editing, 37*(3), 77–80.
Bornmann, L., & Marx, W. (2014). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. *Scientometrics, 98*(1), 487–509. http://dx.doi.org/10.1007/s11192-013-1161-y.
Bornmann, L., Marx, W., Gasparyan, A. Y., & Kitas, G. D. (2012). Diversity, value and limitations of the journal impact factor and alternative metrics. *Rheumatology International, 32*(7), 1861–1867. http://dx.doi.org/10.1007/s00296-011-2276-1.
Bornmann, L., Mutz, R., & Daniel, H.-D. (2010). The *h* index research output measurement: Two approaches to enhance its accuracy. *Journal of Informetrics, 4*(3), 407–414. http://dx.doi.org/10.1016/j.joi.2010.03.005.
Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H.-D. (2011). A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *Journal of Informetrics, 5*(3), 346–359. http://dx.doi.org/10.1016/j.joi.2011.01.006.
Bornmann, L., Wallon, G., & Ledin, A. (2008). Is the *h* index related to (standard) bibliometric measures and to the assessments by peers? An investigation of the *h* index by using molecular life sciences data. *Research Evaluation, 17*(2), 149–156. http://dx.doi.org/10.3152/095820208X319166.
Braun, T., Glänzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics, 69*(1), 169–173. http://dx.doi.org/10.1007/s11192-006-0147-4.
Burrell, Q. L. (2007a). Should the h-index be discounted? In I. W. Glanzel, A. Schubert, & B. Schlemmer (Eds.). *The multidimensional world of Tibor Braun: A multidisciplinary encomium for his 75th birthday* (pp. 65–68). Leuven: ISSI.
Burrell, Q. L. (2007b). Hirsch's h-index: A stochastic model. *Journal of Informetrics, 1*(1), 16–25. http://dx.doi.org/10.1016/j.joi.2006.07.001.
Burrows, R. (2012). Living with the h-index? Metric assemblages in the contemporary academy. *The Sociological Review, 60*(2), 355–372. http://dx.doi.org/10.1111/j.1467-954X.2012.02077.x.
Cerchiello, P., & Giudici, P. (2014). On a statistical *h* index. *Scientometrics, 99*(2), 299–312. http://dx.doi.org/10.1007/s11192-013-1194-2.
Cole, S. (1979). Age and scientific performance. *The American Journal of Sociology, 84*(4), 958–977.
Corrall, S., Kennan, M. A., & Afzal, W. (2013). Bibliometrics and research data management services: Emerging trends in library support for research. *Library Trends, 61*, 636–674. http://dx.doi.org/10.1353/lib.2013.0005.
Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics, 1*(3), 193–203. http://dx.doi.org/10.1016/j.joi.2007.02.001.
Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology, 57*(9), 1275–1278. http://dx.doi.org/10.1002/asi.20354.
Dienes, K. R. (2015). Completing *h*. *Journal of Informetrics, 9*(2), 385–397. http://dx.doi.org/10.1016/j.joi.2015.01.003.
Drummond, R., & Wartho, R. (2009). RIMS: The research impact measurement service at the university of new south wales. *Australian Academic and Research Libraries, 40*(2), 76–87. http://dx.doi.org/10.1080/00048623.2009.10721387.
Egghe, L. (2010). The Hirsch index and related impact measures. *Annual Review of Information Science and Technology, 44*(1), 65–114. http://dx.doi.org/10.1002/aris.2010.1440440109.
Fox, M. F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science, 13*(2), 285–305. http://dx.doi.org/10.1177/030631283013002005.
Franceschini, F., & Maisano, D. A. (2010). Analysis of the Hirsch index's operational properties. *European Journal of Operational Research, 203*(2), 494–504. http://dx.doi.org/10.1016/j.ejor.2009.08.001.
García-Pérez, M. A. (2013). Limited validity of equations to predict the future h index. *Scientometrics, 96*(3), 901–909. http://dx.doi.org/10.1007/s11192-013-0979-7.
García-Pérez, M. A., & Núñez-Antón, V. (2013). Correlation between variables subject to an order restriction, with application to scientometric indices. *Journal of Informetrics, 7*(2), 542–554. http://dx.doi.org/10.1016/j.joi.2013.01.010.
Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics, 1*(4), 359–375.
Garfield, E. (1999). Journal impact factor: A brief review. *Canadian Medical Association Journal, 161*(8), 979–980.
Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA, 295*(1), 90–93. http://dx.doi.org/10.1001/jama.295.1.90.
Gaster, N., & Gaster, M. (2012). A critical assessment of the h-index. *BioEssays, 34*(10), 830–832. http://dx.doi.org/10.1002/bies.201200036.
Gingras, Y. (2014). Criteria for evaluating indicators. In B. Cronin, & C. R. Sugimoto (Eds.). *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact* (pp. 109–125). Cambridge, MA: MIT Press.
Gingras, Y. (2015). Dérives et effets pervers de l'évaluation quantitative de la recherche: sur les mauvais usages de la bibliométrie. *Revue Internationale PME, 28*(2), 7–14. http://dx.doi.org/10.7202/1032829ar.
Glänzel, W. (2006). On the h-index — A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics, 67*(2), 315–321. http://dx.doi.org/10.1007/s11192-006-0102-4.
Hall, B. H., Mairesse, J., & Turner, L. (2007). Identifying age, cohort, and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on French physicists. *Economics of Innovation and New Technology, 16*(2), 159–177. http://dx.doi.org/10.1080/10438590600983010.
Hampton, V. (2009). Jorge Hirsch: The man behind the metric. *Research Trends*(14 December 2009), Retrieved from http://www.researchtrends.com/issue14-december-2009/people-focus/.
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572. http://dx.doi.org/10.1073/pnas.0507655102.
Hirsch, J. E. (2007). Does the h index have predictive power? *Proceedings of the National Academy of Sciences, 104*(49), 19193–19198. http://dx.doi.org/10.1073/pnas.0707962104.
Hirsch, J. E. (2010). An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics, 85*(3), 741–754. http://dx.doi.org/10.1007/s11192-010-0193-9.
Hirsch, J. E., & Buela-Casal, G. (2014). The meaning of the h-index. *International Journal of Clinical and Health Psychology, 14*(2), 161–164.
Hunter, P. (2013). The science of progress and the progress of science. *EMBO Reports, 14*(7), 601–604. http://dx.doi.org/10.1038/embor.2013.83.
Iglesias, J., & Pecharromán, C. (2007). Scaling the h-index for different scientific ISI fields. *Scientometrics, 73*(3), 303–320. http://dx.doi.org/10.1007/s11192-007-1805-x.
Jacsó, P. (2008). The plausibility of computing the h-index of scholarly productivity and impact using reference-enhanced databases. *Online Information Review, 32*(2), 266–283. http://dx.doi.org/10.1108/14684520810879872.
Jacsó, P. (2009). The h-index for countries in Web of Science and Scopus. *Online Information Review, 33*(4), 831–837. http://dx.doi.org/10.1108/14684520910985756.
Jensen, P., Rouquier, J.-B., & Croissant, Y. (2009). Testing bibliometric indicators by their prediction of scientists promotions. *Scientometrics, 78*(3), 467–479. http://dx.doi.org/10.1007/s11192-007-2014-3.
Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The *R*-and *AR*-indices: Complementing the *h*-index. *Chinese Science Bulletin, 52*(6), 855–863. http://dx.doi.org/10.1007/s11434-007-0145-9.
Jones, A. W. (2016). Forensic journals: Bibliometrics and journal impact factors. In J. Payne-James, & R. W. Byard (Eds.). *Encyclopedia of forensic and legal medicine* (pp. 528–538). (2nd ed.). Amsterdam: Elsevier.
Keller, A. (2015). Research support in Australian university libraries: An outsider view. *Australian Academic and Research Libraries, 46*(2), 73–85. http://dx.doi.org/10.1080/00048623.2015.1009528.
Kelly, C. D., & Jennions, M. D. (2006). The *h index* and career assessment by numbers. *Trends in Ecology & Evolution, 21*(4), 167–170. http://dx.doi.org/10.1016/j.tree.2006.01.005.
Kulasegarah, J., & Fenton, J. E. (2010). Comparison of the *h* index with standard bibliometric indicators to rank influential otolaryngologists in Europe and North America. *European Archives of Oto-Rhino-Laryngology, 267*(3), 455–458.
Lazaridis, T. (2010). Ranking university departments using the mean h-index. *Scientometrics, 82*(2), 211–216. http://dx.doi.org/10.1007/s11192-009-0048-4.
Lehmann, S., Jackson, A. D., & Lautrup, B. E. (2008). A quantitative analysis of indicators of scientific performance. *Scientometrics, 76*(2), 369–390. http://dx.doi.org/10.1007/s11192-007-1868-8.
Lemonick, M. (2004). Edward Whitten. *Time. 100* 26 April.
Levin, S., & Stephan, P. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review, 81*(1), 114–132.
Leydesdorff, L., & Bornmann, L. (2011). Integrated impact indicators compared with impact factors: An alternative research design with policy implications. *Journal of the American Society for Information Science and Technology, 62*(11), 2133–2146. http://dx.doi.org/10.1002/asi.21609.
Liang, L. (2006). h-index sequence and h-index matrix: Constructions and applications. *Scientometrics, 69*(1), 153–159. http://dx.doi.org/10.1007/s11192-006-0145-6.
Lovegrove, B. G., & Johnson, S. D. (2008). Assessment of research performance in biology: How well do peer review and bibliometry correlate? *Bioscience, 58*(2), 160–164. http://dx.doi.org/10.1641/B580210.
Marchant, T. (2009). An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. *Scientometrics, 80*(2), 327–344. http://dx.doi.org/10.1007/s11192-008-2075-y.
Martin, B. R., & Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy, 12*(2), 61–90. http://dx.doi.org/10.1016/0048-7333(83)90005-7.
Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research, 246*(1), 1–19. http://dx.doi.org/10.1016/j.ejor.2015.04.002.

Monastersky, R. (2005). The number that's devouring science. *Chronicle of Higher Education, 52*(8), A.12–A.17.

Olmeda-Gómez, C., & de Moya-Anegón, F. (2016). Publishing trends in library and information sciences across European countries and institutions. *The Journal of Academic Librarianship, 42*(1), 27–37. http://dx.doi.org/10.1016/j.acalib.2015.10.005.

Panaretos, J., & Malesios, C. (2009). Assessing scientific research performance and impact with single indices. *Scientometrics, 81*(3), 635–670. http://dx.doi.org/10.1007/s11192-008-2174-9.

Penner, O., Petersen, A. M., Pan, R. K., & Fortunato, S. (2013). Commentary: The case for caution in predicting scientists' future impact. *Physics Today, 66*(4), 8–9. http://dx.doi.org/10.1063/PT.3.1928.

Petersen, A. M., & Succi, S. (2013). The Z-index: A geometric representation of productivity and impact which accounts for information in the entire rank-citation profile. *Journal of Informetrics, 7*(4), 823–832. http://dx.doi.org/10.1016/j.joi.2013.07.003.

Phelan, T. J. (1999). A compendium of issues for citation analysis. *Scientometrics, 45*(1), 117–136. http://dx.doi.org/10.1007/BF02458472.

Prathap, G. (2010). Is there a place for a mock *h*-index? *Scientometrics, 84*(1), 153–165. http://dx.doi.org/10.1007/s11192-009-0066-2.

Prathap, G., & Gupta, B. M. (2009). Ranking of Indian universities for their research output and quality using a new performance index. *Current Science, 97*(6), 751–752. Retrieved from http://www.jaduniv.edu.in/upload_files/ach_data/1267097167-1.pdf.

van Raan, A. F. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics, 67*(3), 491–502. http://dx.doi.org/10.1556/Scient.67.2006.3.10.

Rousseau, R., & Leuven, K. U. (2008). Reflections on recent developments of the h-index and h-type indices. *COLLNET Journal of Scientometrics and Information Management, 2*(1), 1–8.

Saad, G. (2006). Exploring the h-index at the author and journal levels using bibliometric data of productive consumer scholars and business-related journals respectively. *Scientometrics, 69*(1), 117–120. http://dx.doi.org/10.1007/s11192-006-0142-9.

Saad, G. (2010). Applying the *h*-index in exploring bibliometric properties of elite marketing scholars. *Scientometrics, 83*(2), 423–433. http://dx.doi.org/10.1007/s11192-009-0069-z.

Schreiber, M. (2007). A case study of the Hirsch index for 26 non-prominent physicists. *Annalen der Physik, 16*(9), 640–652. http://dx.doi.org/10.1002/andp.200710252.

Schreiber, M. (2008). To share the fame in a fair way, h_m modifies *h* for multi-authored manuscripts. *New Journal of Physics, 10*(4), 1–9. http://dx.doi.org/10.1088/1367-2630/10/4/040201.

Schreiber, M. (2013a). How relevant is the predictive power of the h-index? A case study of the time-dependent Hirsch index. *Journal of Informetrics, 7*(2), 325–329. http://dx.doi.org/10.1016/j.joi.2013.01.001.

Schreiber, M. (2013b). A case study of the arbitrariness of the h-index and the highly-cited-publications indicator. *Journal of Informetrics, 7*(2), 379–387. http://dx.doi.org/10.1016/j.joi.2012.12.006.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43*(9), 628–638.

Seglen, P. O. (1997). Citations and journal impact factors: Questionable indicators of research quality. *Allergy, 52*(11), 1050–1056. http://dx.doi.org/10.1111/j.1398-9995.1997.tb00175.x.

Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica, 69*(3), 224–229. http://dx.doi.org/10.3109/17453679809000920.

Smith, D. R. (2015). "Platinum h": Refining the h-index to more realistically assess career trajectory and scientific publications. *Archives of Environmental & Occupational Health, 70*(2), 67–69. http://dx.doi.org/10.1080/19338244.2015.1016833.

Suiter, A. M., & Moulaison, H. L. (2015). Supporting scholars: An analysis of academic library websites' documentation on metrics and impact. *The Journal of Academic Librarianship, 41*(6), 814–820. http://dx.doi.org/10.1016/j.acalib.2015.09.004.

Todeschini, R., & Baccini, A. (2016). *A handbook of bibliometric indicators: Quantitative tools for studying and evaluating research.* Weinheim: Wiley-VCH.

Tol, R. S. (2008). A rational, successive *g*-index applied to economics departments in Ireland. *Journal of Informetrics, 2*(2), 149–155. http://dx.doi.org/10.1016/j.joi.2008.01.001.

Van Leeuwen, T. (2008). Testing the validity of the Hirsch-index for research assessment purposes. *Research Evaluation, 17*(2), 157–160. http://dx.doi.org/10.3152/095820208X319175.

Vinkler, P. (2007). Eminence of scientists in the light of the h-index and other scientometric indicators. *Journal of Information Science, 33*(4), 481–491. http://dx.doi.org/10.1177/0165551506072165.

Waltman, L., & van Eck, N. J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology, 63*(2), 406–415. http://dx.doi.org/10.1002/asi.21678.

Wan, X. (2014). X-index: A fantastic new indicator for quantifying a scientist's scientific impact. Retrieved from http://arxiv.org/abs/1405.0641.

Woeginger, G. J. (2008). An axiomatic characterization of the Hirsch-index. *Mathematical Social Sciences, 56*(2), 224–232. http://dx.doi.org/10.1016/j.mathsocsci.2008.03.001.

Wu, J., Lozano, S., & Helbing, D. (2011). Empirical study of the growth dynamics in real career h-index sequences. *Journal of Informetrics, 5*(4), 489–497. http://dx.doi.org/10.1016/j.joi.2011.02.003.

Zhang, C.-T. (2013). The h′-index, effectively improving the h-index based on the citation distribution. *PloS One, 8*(4), e59912. http://dx.doi.org/10.1371/journal.pone.0059912.

Zhivotovsky, L. A., & Krutovsky, K. V. (2008). Self-citation can inflate *h*-index. *Scientometrics, 77*(2), 373–375. http://dx.doi.org/10.1007/s11192-006-1716-2.