Check for updates

Researchers' attitudes towards the *h*-index on Twitter 2007–2020: criticism and acceptance

Mike Thelwall[1] · **Kayvan Kousha**[1]

Abstract

The *h*-index is an indicator of the scientific impact of an academic publishing career. Its hybrid publishing/citation nature and inherent bias against younger researchers, women, people in low resourced countries, and those not prioritizing publishing arguably give it little value for most formal and informal research evaluations. Nevertheless, it is well-known by academics, used in some promotion decisions, and is prominent in bibliometric databases, such as Google Scholar. In the context of this apparent conflict, it is important to understand researchers' attitudes towards the *h*-index. This article used public tweets in English to analyse how scholars discuss the *h*-index in public: is it mentioned, are tweets about it positive or negative, and has interest decreased since its shortcomings were exposed? The January 2021 Twitter Academic Research initiative was harnessed to download all English tweets mentioning the *h*-index from the 2006 start of Twitter until the end of 2020. The results showed a constantly increasing number of tweets. Whilst the most popular tweets unapologetically used the *h*-index as an indicator of research performance, 28.5% of tweets were critical of its simplistic nature and others joked about it (8%). The results suggest that interest in the *h*-index is still increasing online despite scientists willing to evaluate the *h*-index in public tending to be critical. Nevertheless, in limited situations it may be effective at succinctly conveying the message that a researcher has had a successful publishing career.

Introduction

The *h*-index, the largest number h such that at least h publications have been cited at least h times each, has since been criticized or dismissed in the bibliometrics literature for conflating publishing with impact, for being incomparable between fields and for being biased

✉ Mike Thelwall
  m.thelwall@wlv.ac.uk

  Kayvan Kousha
  k.kousha@wlv.ac.uk

1 Statistical Cybermetrics Research Group, University of Wolverhampton, Wolverhampton, UK

against younger, female and less publishing-focused researchers (Gingras, 2016). Nevertheless, the *h*-index has been used by research managers and individual scientists, perhaps for the simplicity of its calculation or in the desire for a simple overall indicator of career impact (Hammarfelt & Rushforth, 2017; Leydesdorff et al., 2016). At least one university also constructs tables of *h*-index values expected for promotion in different disciplines (seen by the first author), and others may calculate median *h*-indexes for academic ranks to guide promotion decisions (e.g., Bertocci & Koenig, 2019). In terms of professional support, most academic libraries in Australia provide academics support with obtaining their *h*-index (Haddow & Mamtora, 2017) so there is continuing demand for it. The *h*-index is also still written about in academic publications, with 2880 Scopus journal articles (document type "Journal Article", excluding editorials, reviews and other types) matching the query TITLE-ABS-KEY(*h*-index OR "*h* index") AND (LIMIT-TO(DOCTYPE, "ar")) between 2005 and 2020. About 4% of these articles refer to other types of *h*-index, such as for electroneurography. This interest is still increasing, with an approximately linear growth pattern between 2006 and 2020. Based on their titles and abstracts, articles in 2020 seemed to use the *h*-index (and other indicators) at face value for bibliometric studies (e.g., "The use of *H*-index to assess research priorities in poultry diseases", and "Women's productivity in mental health research in the Gulf Cooperation Council").

Although there are many author-level indicators (Orduña-Malea et al., 2016), the *h*-index seems to be the best known, apart perhaps from total citations. This parallels the situation for the Journal Impact Factor (JIF) and related journal-level indicators that give a simple indicator of the average citation impact of journals and are widely used in formal and informal evaluations, despite extensive criticisms and—in this case—the San Francisco Declaration on Research Assessment (DORA) campaign against their overuse. Perhaps the strength of the *h*-index is its single simple number assigned to academics that correlates moderately with peer judgment of career achievement for academics in some fields (Norris & Oppenheim, 2010), but is not reliant on peer judgement, which has its own biases. Similarly, the strength of the JIF may be its single simple number assigned to journals that correlates positively with journal reputation within a field, and it may be a better impact indicator than article-level citations (Waltman & Traag, 2020). In any case, given the widespread availability and known severe problems with the *h*-index, it is important for bibliometricians to assess how it is viewed by researchers in general and how they talk about it publicly. Both can be investigated to some extent through Twitter.

Despite the widespread availability of the *h*-index, no previous large-scale survey of academics seems to have elicited researchers' attitudes towards it. For example, interviews with 79 senior faculty at a Canadian university about scholarly metrics did not produce any *h*-index attitudes (Thuna & King, 2017). Another study asked medical and physics postdocs at an Australian university whether they used the *h*-index (most did) but did not ask them what they thought of it (Derrick & Gillespie, 2013). In a partial exception, a survey of 206 tenured non-medical faculty at a US university in 2015 found that an almost equal percentage of academics that had little, moderate or a lot of knowledge about scholarly indicators, such as the *h*-index, JIF and H5 median. Over 70% thought that they played a role in the promotion and tenure processes. There was more knowledge in science and less in the arts and humanities. Few academics thought that they should have high weight in this process and 36% thought they should play little role (DeSanto & Nichols, 2017). In contrast only 30% of mining engineers surveyed believed that the *h*-index was consulted during the promotion process, although 52% did not know (Saydam & Kecojevic, 2014). A New Zealand university faculty survey also found a belief that academic indicators were overrated (Ferrier-Watson, 2019). Another partial exception found the first author's *h*-index

to be almost never helpful for academic literature searches (Lemke et al., 2021). A survey of 471 Virginia Tech faculty about altmetrics and research indicators found that 40% used the h-index, about equally for personal and professional reasons (Miles, et al., 2020). Attitudes towards the h-index should be interpreted in the context of overall attitudes towards using indicators in academic evaluations. In the UK at least, the use of a quantitative indicator is often perceived as encroaching on research freedoms and pressure towards management by metrification (Burrows, 2012).

As argued above, no previous study has investigated academic attitudes towards the h-index, attitudes expressed in public or changes in these over time. This study takes advantage of the January 2021 introduction of free academic access to the complete Twitter archive to partly address this gap. Twitter is suitable since it is used by academics in some countries to disseminate and discuss their research (e.g., 15% in the UK: Zhu, 2014). The following questions drive this study.

- RQ1: Is there decreasing interest in the h-index on Twitter?
- RQ2: Do tweets about the h-index tend to be critical of it?

Methods

The research design was to gather a reasonably comprehensive set of tweets about the h-index and apply a range of descriptive methods to extract information relevant to the research questions.

Data

The tweets were gathered using the Twitter API, academic track, which gives access to the full Twitter archive. This excludes tweets judged to be spam by Twitter algorithms and tweets deleted by their authors. Tweets were searched for with the queries "h-index" and "h index", specifying English as the language. The restriction to English was taken because the vast majority of tweets were in English, so it seemed reasonable to focus on these to give a more homogeneous set to analyse. Although the Twitter academic track allows tweets to be specified by country, most tweets do not have a location associated with them and so focusing on a few countries (e.g., USA, UK) would reduce the amount of texts too much for an analysis of trends. The queries were submitted using the free software Mozdeh in January 2020, obtaining 88,529 tweets.

Data cleaning

A filtering process was applied to remove tweets that were duplicate or near duplicate. For example, a highly retweeted tweet may have many identical copies of itself in the dataset. Tweets were also regarded as duplicates if their texts were identical except for hyperlinks and @usernames because tweets can be shared by forwarding to particular individuals and links can be re-created by new users sending the same message. The filtering process removed these duplicates, with the remaining tweets accounting for 43% of the original set. Not all tweets mentioned the h-index, with the query also matching the text of linked pages or user biographies (e.g., there were 433 unrelated tweets from beichthaus, none of which mentioned the h-index). Tweets not mentioning the h-index were therefore removed

in a final stage. Thus, the final dataset consists of unique tweets in English mentioning "*h*-index" or "*h* index".

Analysis

A monthly time series graph was produced to illustrate any trends in the volume of tweeting about the *h*-index in English. It would have been useful to illustrate the proportion of all English tweets that were about the *h*-index but there is no way to identify the monthly number of English tweets and in any case this proportion is influenced by the share of tweeters that are academics, which has presumably decreased over time.

The top retweets were examined to identify prominent and popular posts about the *h*-index. A highly retweeted tweet may reveal aspects of the *h*-index that are widely agreed with on Twitter or that resonate with tweeters for another reason.

A hundred random tweets were examined through a content analysis (Neuendorf, 2016) to identify their main topics. The classification focused on the role of the term *h*-index in the tweets and whether they were positive, negative or neutral about it. Both coders classified the tweets independently with the same inductive classification scheme developed on the set. A Cohen's kappa (Cohen, 1960) of 0.501 indicates a moderate level of agreement (Landis & Koch, 1977), which is enough to use in practice. The codes for disagreements were revisited and changed when one was judged to be incorrect in retrospect. In cases where both codes were reasonable, such as a joke that could be interpreted as critical of the *h*-index, both codes were retained with half weight.

Results

There were 30,681 non-duplicate English-language tweets containing "*h*-index" or "*h* index" between the start of Twitter in 2006 and the end of 2020 (Fig. 1). The first tweet was from December 2007. The tweets with a known location were mainly from the USA (6640, 21.6%), the UK (3809, 12.4%), Australia (1476, 4.8%), and Canada (1388, 4.5%), with other countries contributing less than 2% each and 38.6% being from an unknown location. The most prolific *h*-index tweeters were Nader Ale Ebrahim (233 tweets, 0.5% of all *h*-index tweets), a bibliometrician at Alzahra University, Tehran, and DrugMonkey (167 tweets, 0.5%), an anonymous academic medicine blogger apparently from the USA. Everyone else had 80 or fewer tweets in the set (0.3%).

Changes over time

The overall level of posting about the *h*-index in English has increased over time (Fig. 1). Whilst early increases might be due to increases in the uptake of Twitter (which started in 2006), this does not seem to be a likely explanation for the increased rate of tweeting from 2017 onwards. The biggest spike in the graph was caused by a widely shared September 2012 *Nature* comment with a mathematical formula for a researcher to predict their *h*-index value in five years (Acuna et al., 2012).
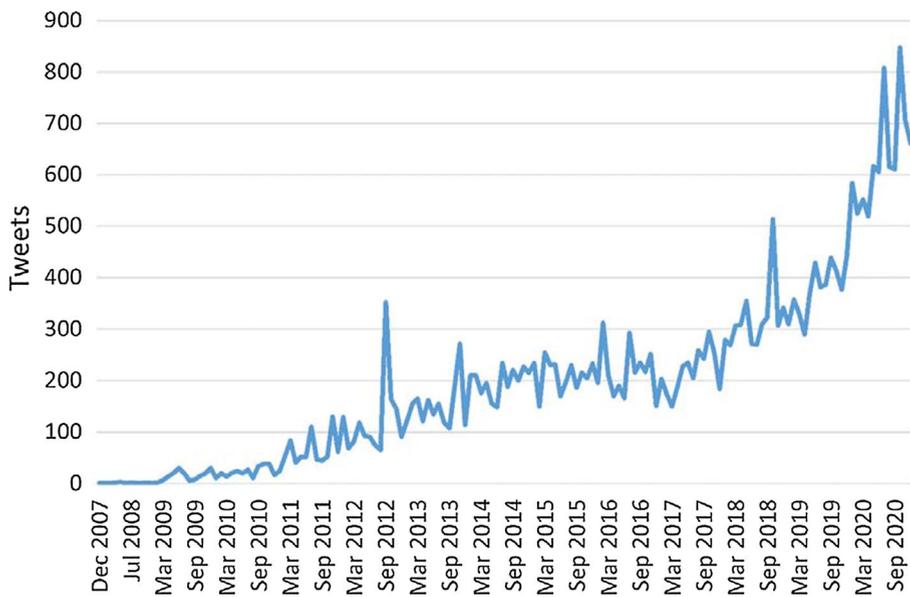
Fig. 1 The monthly number of tweets (excluding duplicates) in English containing h-index or "*h* index"

Content of typical tweets

Over a quarter of the tweets (28.5%) about the *h*-index criticised it or an aspect of it (Table 1, Fig. 2), with 7% being neutral and none explicitly praising it. Thus, the overall tone of evaluative tweets about the *h*-index was substantially negative. Nevertheless, the remaining tweets were non-evaluative and, except for the jokes, could be interpreted as implicitly endorsing the *h*-index by mentioning it without caveats. The jokes (8%)

Table 1 The classification scheme for random tweets. Codes were applied in descending order, allocating the highest matching code

| Code	Description
Own *h*-index value	Tweet about the tweeter's *h*-index, whether reporting a value or discussing it
Other person's *h*-index value	Reporting another person's *h*-index value as evidence of their achievements, including high, low or medium values
Other *h*-index applications	Using the *h*-index for a purpose other than measuring an individual scholar's publishing achievements
Criticism	Criticism of the *h*-index. A tweet that mentioned anything negative about the h-index, such as saying that it is bad
Neutral	Tweets with positive and negative points about the *h*-index
Calculations	Information to help calculate *h*-indexes, or link to a page with this information
Variants	Tweets about variants of the *h*-index, such as i10
Joke	A joke or tweet intended to be funny
Other	Any tweet not matching the above categories
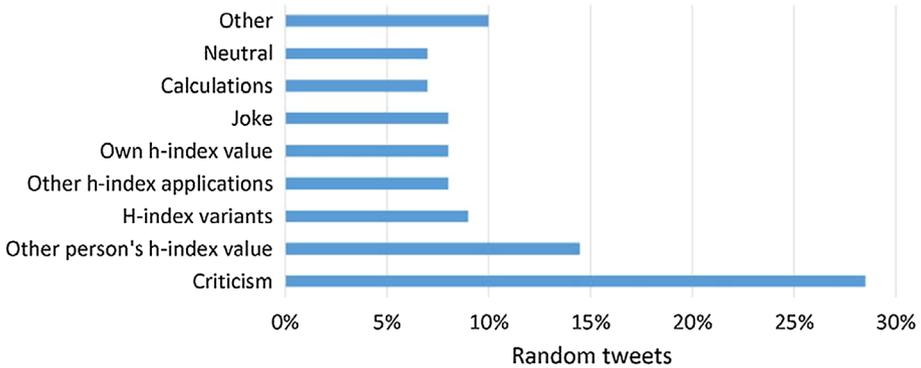
Fig. 2 The main topics of 100 randomly selected h-index tweets

exploited knowledge about the *h*-index within the academic community for humour, such as asking how human hierarchies were organised before the *h*-index, or whether research group photographs were taken by the member with the lowest *h*-index. According to the benign violation theory of humour (McGraw & Warren, 2010), jokes need an element of simultaneous threat and non-threat to work well, so some jokes may reflect a fear of the *h*-index, although other aspects of the jokes might also provide the threat.

The second most common topic of the classified tweets was to report or discuss the *h*-index of a third person (14.5%). These tweets typically reported a high value as evidence of achievement, but some reported a low *h*-index as a criticism. Many tweeters (8%) also reported their own *h*-indexes. In some cases these were low values (e.g., 1, 2, 3) and it was not clear whether the self-report was intended to be self-effacing or a joke (if others in their team had much higher values), or attracting first citations was a proud achievement, or a combination.

A range of tweets focused on technical details, such as proposed *h*-index variants (9%), calculation details (7%), or applications of the *h*-index other than for individual scholars (8%).

Highly retweeted tweets

All English *h*-index tweets with at least 100 retweets were examined and characterized for purpose and attitude towards the h-index (*n* = 31). The following summarises the results.

- *Evidence a third person's achievements* Eight highly retweeted tweets (128–13,752 retweets, total: 18,022) cited a third person's *h*-index as evidence of their achievements. The most retweeted cited Anthony Fauci's *h*-index in this way and another listed his *h*-index amongst other achievements. The succinct nature of the *h*-index seems to fit this use on Twitter.
- *Jokes about the h-index* Four highly retweeted tweets (341–2146 retweets, total: 3719) joked about or with the *h*-index.
- *Criticize the h-index* Thirteen highly retweeted tweets (103–505 retweets, total: 3366) criticized an aspect of the *h*-index, such as through mentioning how it can be gamed, how it is biased, or how it does not reflect non-publishing research achievements.

- *Other face value use* Three highly retweeted tweets (130–170 retweets, total: 450) accepted the *h*-index at face value for a purpose other than to praise a scholar's publishing achievements.

Thus, partly echoing the randomly-selected set above, highly-retweeted tweets reflect a combination of criticism, jokes and face-value uses of the *h*-index.

Limitations

A limitation of the current study is a lack of knowledge about the tweeters. Whilst they seem to be mainly academics, they may include some science journalists, non-research students, non-research librarians and information professionals. Moreover, the results should not be taken as evidence attitudes towards the *h*-index in English speaking countries, since there may be differences between them, some tweets may be from other nations, and only a minority of academics tweet about research. Also, academics that are positive about it may avoid making this public in fear of negative reactions. The trend in the volume of tweeting should also be interpreted cautiously because it is not benchmarked against, for example, percentages of English-speaking academics active on Twitter. Thus, increases in tweeting about the *h*-index do not necessarily correspond to increases in academic interest in it.

Conclusions

The results suggest that on English Twitter there is still an increasing amount of interest in the *h*-index and whilst most evaluative tweets are critical, the majority accept it at face value, such as to report a score. This dual approach is perhaps unsurprising given that busy academics may not wish to learn about bibliometric details and the *h*-index has a surface plausibility in some fields (because recognised academics tend to have high scores, and because publishing more and being cited more are sometimes valued). The situation is perhaps further complicated by the easy availability of the *h*-index in bibliometric databases, its prominence in Google Scholar, and the arguably reasonable use in some high-profile tweets. For example, tweeting Anthony Fauci's high *h*-index is arguably effective at conveying the simple message that he is a successful publishing academic, which might add credibility to his COVID-19 announcements or be of general interest to researchers who want to know more about the public face of science during the pandemic in the USA.

A practical conclusion from the results is that bibliometricians and librarians will need to continue helping researchers and research managers to understand the major limitations of the *h*-index and recommend avoiding it in most circumstances. This does not seem to be a difficult task since there does not seem to be a strong case in support of its use, at least on Twitter, but it will need to continue with each new generation of academics and research managers.

Finally, this article also demonstrates that it is now possible to investigate interest in academic issues expressed on Twitter with the new academic research access to historical tweets. This has two advantages: a time series analysis of trends, and the possibility to gather enough tweets to analyse topics that are rarely discussed. In the past, long term monitoring of Twitter would have been needed to capture relevant tweets (e.g., Thelwall et al., 2021), but now it is quick and straightforward to capture them retrospectively.

References

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

Bertocci, G. & Koenig, S. (2019). Pathway to successful achievement of speed faculty promotion and tenure. https://engineering.louisville.edu/wp-content/uploads/2019/05/Promotion-and-Tenure-Process.pdf.

Burrows, R. (2012). Living with the *h*-index? Metric assemblages in the contemporary academy. *The Sociological Review, 60*(2), 355–372.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Derrick, G. E., & Gillespie, J. (2013). A number you just can't get away from: Characteristics of adoption and the social construction of metric use by researchers. https://eprints.lancs.ac.uk/id/eprint/88786/.

DeSanto, D., & Nichols, A. (2017). Scholarly metrics baseline: A survey of faculty knowledge, use, and opinion about scholarly metrics. *College and Research Libraries, 78*(2), 150–170.

Ferrier-Watson, A. (2019). Traditional metrics, altmetrics and researcher profiles: A survey of faculty perceptions and use. http://hdl.handle.net/10063/8362.

Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. MIT Press.

Haddow, G., & Mamtora, J. (2017). Research support in Australian academic libraries: Services, resources, and relationships. *New Review of Academic Librarianship, 23*(2–3), 89–109.

Hammarfelt, B., & Rushforth, A. D. (2017). Indicators as judgment devices: An empirical study of citizen bibliometrics in research evaluation. *Research Evaluation, 26*(3), 169–180.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lemke, S., Mazarakis, A., & Peters, I. (2021). Conjoint analysis of researchers' hidden preferences for bibliometrics, altmetrics, and usage metrics. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24445.

Leydesdorff, L., Wouters, P., & Bornmann, L. (2016). Professional and citizen bibliometrics: Complementarities and ambivalences in the development and use of indicators—a state-of-the-art report. *Scientometrics, 109*(3), 2129–2150.

McGraw, A. P., & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological Science, 21*(8), 1141–1149.

Miles, R., Pannabecker, V., & Kuypers, J. A. (2020). Faculty perceptions of research assessment at Virginia Tech. *Journal of Altmetrics*. https://doi.org/10.29024/joa.24.

Neuendorf, K. A. (2016). *The content analysis guidebook*. (2nd ed.). SAGE.

Norris, M., & Oppenheim, C. (2010). Peer review and the *h*-index: Two studies. *Journal of Informetrics, 4*(3), 221–232.

Orduña-Malea, E., Martín-Martín, A., & Delgado-López-Cózar, E. (2016). The next bibliometrics: ALMetrics (Author Level Metrics) and the multiple faces of author impact. *El Profesional de la Información, 25*(3), 485–496.

Saydam, S., & Kecojevic, V. (2014). Publication strategies for academic career development in mining engineering. *Mining Technology, 123*(1), 46–55.

Thelwall, M., Makita, M., Mas-Bleda, A., & Stuart, E. (2021). "My ADHD hellbrain": A Twitter data science perspective on a behavioural disorder. *Journal of Data and Information Science*. https://doi.org/10.2478/jdis-2021-0007.

Thuna, M., & King, P. (2017). Research impact metrics: A faculty perspective. *Partnership: The Canadian Journal of Library and Information Practice and Research*. https://doi.org/10.21083/partnership.v12i1.3906.

Waltman, L., & Traag, V. A. (2020). Use of the journal impact factor for assessing individual articles need not be statistically wrong. *F1000Research, 9*, 366. https://doi.org/10.12688/f1000research.23418.1.

Zhu, Y. (2014). Seeking and sharing research information on social media: A 2013 survey of scholarly communication. In: *Proceedings of European Conference on Social Media ECSM* (pp. 705–712).