

Use and Misuse of Bibliometric Measures for Assessment of Academic Performance, Tenure and Publication Support¹

Yaşar Tonta

Hacettepe University Department of Information Management, 06800 Beytepe, Ankara, Turkey
Email: tonta@hacettepe.edu.tr

"Not everything that counts can be counted, and not everything that can be counted counts." -- Albert Einstein

Abstract: Bibliometric methods such as journal impact factor and article influence score based on the number of citations were developed to measure and compare the quality of journals listed in citation indexes. Yet, they are increasingly being used nowadays for research assessment, hiring, tenure and academic promotion, research funding and publication support even though such metrics have not been developed to measure the quality of individual researchers or scientific articles. In this paper, we review the use of journal impact factor, cited half-life, article influence score and h index for academic performance assessment, academic promotion and publication support by Turkish universities and the Turkish Scientific and Technological Research Center (TUBITAK). Examples are provided regarding the consequences of using bibliometric measures beyond what they were originally designed for, and some recommendations are offered.

Key words: research assessment, academic promotions, journal impact factor, cited half-life, article influence score, h index

Introduction

This paper addresses the misuse of bibliometric measures for research assessment, academic promotion and monetary support for academic publications. Although most of the examples in this study come from the Turkish higher education system, such use is quite widespread in other countries as well and thus merits further investigation. For instance, the criteria set for academic promotions by the Turkish Higher Education Council (HEC) and universities for research are almost entirely based on the number of papers published in ISI-indexed journals and the number of citations thereto. The Turkish Scientific and Technological Research Center (TUBITAK) provides monetary support to the authors of such papers on the basis of impact factors of journals in which their papers appeared. Similarly, the Research Council of Thailand provides monetary support to researchers simply by multiplying the number of papers they authored with the impact factors of journals in which they appeared (Arendt, 2010). Therefore, it is useful to look at more carefully as to what bibliometric measures such as journal impact factor, cited half-life, article influence score and h index exactly measure and whether they are suitable to use as criteria to make decisions on tenure, academic promotion and publication support.

It should be mentioned at the outset that bibliometric measures alone are not the sole criteria for research assessment. Rather, such assessment is primarily based on peer review. For example, in the United Kingdom (UK) a panel of experts evaluates the quality of research output of

¹ Paper presented at the Metrics 2014: Workshop on Informetric and Scientometric Research. 77th Annual Meeting of the Association for Information Science and Technology, October 31-November 5, 2014, Seattle, WA. This paper is largely based on an earlier work entitled "An Evaluation of Criteria on Academic Performance, Tenure and Publication Support" (in Turkish). For a more detailed treatment of the topic, see <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/tonta-yukseltme-kriterleri-hakkinda-degerlendirme-11-Temmuz-2014.pdf>.

universities/departments as part of the Research Excellence Framework (REF) and allocates research budgets accordingly. In fact, panel members are not even allowed to use journal impact factors or any journal ranking system in evaluating academic publications (Sgroi & Oswald, 2013, p. F257). Peer review is also used in Turkey but only after the candidates satisfy the requirements of the number of papers/citations in ISI-indexed journals, presumably because bibliometric measures seem more “objective” to HEC authorities than the outcome of peer reviews. However, this approach would undermine the importance of peer review and we would lose faith in our very own expert judgements.

Use of Citation Indexes in Research Evaluation

Journals are screened by Thomson Reuters (formerly ISI) before they get accepted in citation indexes (namely, Science Citation Index, Social Science Citation Index, and Arts & Humanities Index). Once accepted, the impact of papers published therein is measured by several metrics including journal impact factor (JIF). The impact factor of a journal is obtained by calculating the ratio between the number of citations and the number of citable items (e.g., articles) published (Garfield, 1994). Yet, JIF is increasingly being used to measure the quality of a single paper published in a given journal rather than the average quality of the journal itself. In other words, JIF is used to measure how many citations an “average paper” would get in a certain period, usually two years after its publication.

As the San Francisco Declaration on Research Assessment (DORA) clearly states, JIF has been developed to help librarians select journals and it cannot be used to evaluate the quality of a paper (San Francisco, 2012). The disadvantages of using JIF in research evaluation are well documented in the literature (see, for instance, Casadevall & Fang, 2014; Marx & Bornmann, 2013; Seglen, 1997). Among them are: (a) citation distributions are skewed; (b) JIF varies by disciplines and can be manipulated by editorial policies; and (c) the data used to calculate JIF are not transparent and contributions other than research articles are also considered when calculating JIF. Of the 11,500 journals listed in Thomson Reuters’s citation indexes, 43% have JIFs between 0 and 1 (Al & Soydal, 2014). It is likely that considerable number of those low impact journals might have been added after Thomson Reuters’s decision of regional expansion in 2006. Thomson Reuters, too, is against the use of JIF to measure paper quality (Marx & Bornmann, 2013, pp. 62-63). Therefore, taking the average JIF says almost nothing about the quality of an individual paper, let alone predict how many, if any, citations it would get in the coming years.

In spite of several criticisms, Thomson Reuters’s citation indexes are primarily used for academic performance evaluation in many countries including Turkey. In mid-1990s, the Turkish HEC introduced the minimum criteria for tenured professors. One of the criteria has been to have a certain number of refereed papers (usually, 2 to 5 depending on the field and seniority) published in journals that are covered by citation indexes. Universities were free to set their own criteria for research papers provided the number of papers required is not below that determined by HEC. HEC’s decision has certainly increased the number of papers published in indexed journals and considerably improved Turkey’s ranking over the years. Yet, there has been a constant debate since then as to the appropriateness of setting such thresholds as the nature of scholarly communication varies by scientific domains. Furthermore, some universities introduced additional criteria based on JIFs. Papers published in journals with higher JIFs tend to be assessed more favorably and hence are assigned higher scores during the initial evaluation of the portfolios of academics who are up for tenure (e.g., Hacettepe, 2014).

It should also be mentioned that using the products of a commercial company in research evaluation turned out to present some problems, at least in Turkey. Up until about 10 years ago, the then ISI has been the sole publisher of citation indexes. When Elsevier’s Scopus entered the market with more journals indexed, Thomson Reuters decided in 2006 to increase its number of journals through what is called “regional expansion”. This decision was welcomed by many countries having a few journals in Thomson Reuters’s citation indexes. Turkey was no exception, as the number of ISI-indexed journals

published in Turkey has increased from 5 to about 80! Needless to say, this increase cannot simply be explained by an unusual surge in the scientific level of Turkey. Yet, papers that appeared in those journals helped many academicians to satisfy the tenure requirements and enabled their authors to receive monetary support offered by TUBITAK at the same time.

Use of Journal Impact Factors for Publication Support

We pointed out earlier that papers published in journals with higher JIFs are considered of having higher quality even though the majority of them may not necessarily generate the average number of citations as specified by their JIFs. Nonetheless, TUBITAK used JIFs for more than a decade in its support program of international scientific publications and classified journals according to their JIFs as reported in Journal Citation Reports (JCR) published annually by Thomson Reuters. The top quarter of journals with the highest JIFs in a given discipline are assigned “A”, the second quarter “B”, and the rest “C” (and “D”) (UBYT Programı, 2012).² TUBITAK used this classification until 2013 to determine the level of annual monetary support to be dispensed and incentivized the authors to publish more papers in journals covered by citation indexes.

In 2013, TUBITAK has almost doubled the amount of monetary support for individual papers published in journals covered by citation indexes, and, at the same time, changed its algorithm of ranking journals. Apparently, TUBITAK wanted to distinguish the high impact journals further and provide more support to those who published in them. Rather than classifying journals roughly as A, B, C and D and dispensing the same amount of money to authors in each category, TUBITAK decided to rank journals more finely on the basis of its own “journal impact factor” consisting of five-year JIFs and cited half-lives of journals (both provided by JCR). The two are multiplied to come up with TUBITAK JIFs and journals were ranked accordingly. Journals having TUBITAK JIFs 2 standard deviations (SD) above the average would then get the highest monetary support while the ones with 2 SDs below the average would get the minimum support. The amount of support for journals in between ± 2 SDs of average was calculated by means of a linear transformation formula that took the number of journals into account in each JCR discipline. The authors of papers were rewarded on a sliding scale between a maximum of 5,000 Turkish Lira and a minimum 500 Turkish Lira.

Needless to say, both methods assume that any paper published in these journals is as good as any other one, without taking the individual impact (i.e., number of citations) of each paper. It can be argued that TUBITAK's own JIF measures the individual impact more sensitively as it consists of both the five-year impact factor and cited half-lives of journals. Not quite so. First, the five-year impact factors of journals are skewed, too (most journals having five-year JIFs between 0 and 1). Second, the cited half-life of a journal has nothing to do with its quality: it is the median (in years) of citations to papers published in a given year. If the cited half-life of a journal is, say, 6 years, it simply means that half the citations to these papers would be received within the first 6 years of their publication. Cited half-lives of journals depend on how fast the literatures in certain disciplines obsolesce. For instance, papers in Science, Technology, Engineering and Medicine (STEM) journals obsolesce much faster than that in the Social Sciences. It could be that TUBITAK might have introduced its own JIF to balance the discrepancy between JIFs of journals in Sciences and Social Sciences, as the former are higher but the cited half-lives of them are shorter while the opposite is the case for Social Sciences journals. However, the cited half-life of a journal is simply a measure of the length (in years) of the scientific impact of papers published in journals. It also informs librarians of the duration of usefulness of journals so that they can decide as to how long they should keep the back issues of those journals

² Note that TUBITAK has changed the rules regarding the classification of journals under “C” and “D” at some point. In Social Sciences, the second half of journals were divided into two: 40% of them being labeled as “C” and the last 10% as “D”. Later, TUBITAK stopped supporting the authors of papers publishing in journals under “C” in Sciences (i.e., the last 50% of journals) and “D” in Social Sciences (i.e., the last 10% of journals) (UBYT Uygulama, 2012).

in the collection (Tonta & Ünal, 2008, p. 337). The cited half-life of a journal has nothing to do with the quality of individual papers published in it.

It turns out that TUBITAK's new algorithm did not meet the expectations. For instance, some Archaeology journals receiving the highest monetary support earlier became the least supported ones when the new algorithm was used (Batmaz, 2013). The anecdotal evidence suggested that this was also the case for the top-notch information science journals such as *JASIST* and *Journal of Informetrics*. Consequently, TUBITAK quickly abandoned its new algorithm (based on five-year impact factors and cited half-lives of journals) after using it only once in 2013 and decided to use JCR's article influence score (AIS) in 2014 to rank the journals (TUBITAK, 2013; 2014 Yılı, 2014).³ The AIS of papers published in a journal is calculated by taking into account the five-year JIF along with the whole JCR citation network, and (similar to Google's PageRank algorithm) citations coming from papers in highly cited journals are weighted more heavily. It is suggested that AIS can therefore be used for interdisciplinary comparisons (Arendt, 2010).

It is not possible to compare the relationship between the 2013 and 2014 algorithms as the list of journals supported by TUBITAK in 2013 is not available on the web (but should be known to TUBITAK).⁴ Table 1 provides the amount of monetary support that the 16 Archaeology journals received in 2012, 2013 and 2014. Figure 1 provides the same information for 13 Archaeology journals for which data were available for all three years.

Table 1. TUBITAK's support to Archaeology journals (2012-2014)

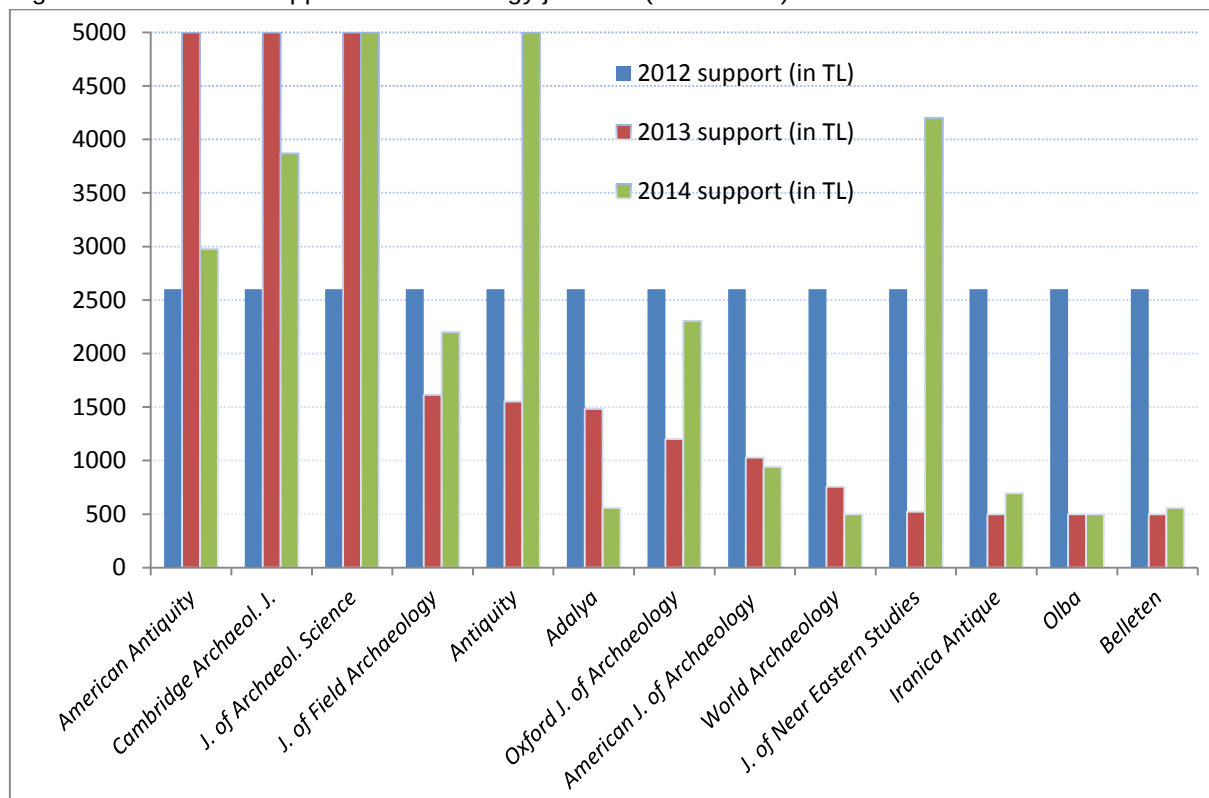
Journal name	Class	2012	Score	2013	Score	2014
		Support (in Turkish Lira)		Support (in Turkish Lira)		Support (in Turkish Lira)
American Antiquity	A	2600	100	5000	74	2976
Cambridge Archaeological Journal	A	2600	100	5000	87	3869
Journal of Archaeological Science	A	2600	100	5000	100	5000
Journal of Field Archaeology	A	2600	50	1613	62	2202
Antiquity	A	2600	48	1553	100	5000
Adalya*	A	2600	47	1484	11	559
Oxford Journal of Archaeology	A	2600	39	1201	63	2304
American Journal of Archaeology	A	2600	34	1028	31	943
World Archaeology	A	2600	24	757	0	500
Archaeological Dialogues	A	2600	10	548	--	--
Journal of Near Eastern Studies	A	2600	7	523	91	4201
Near Eastern Archaeology	--	--	4	506	64	2342
Iranica Antiquae	A	2600	1	500	21	695
Olba*	A	2600	0	500	0	500
Belleten*	A	2600	0	500	11	559
Turkish Academy of Sciences Journal of Archaeology (TUBA-AR)*	--	--	0	500	--	--

Note: Journals with "*" are published in Turkey. The journal list and the support figures are taken from Batmaz (2013). Journals are ranked according to TUBITAK's 2013 algorithm. Scores and support (in Turkish Lira) are rounded to the nearest integer. Figures reflect the amount of support given to journal articles (not case studies, letters to the editor, and so on). 2012 and 2014 data come from http://ulakbim.tubitak.gov.tr/sites/images/Ulakbim/ubyt_2012_dergi_listesi.xls and http://ulakbim.tubitak.gov.tr/sites/images/Ulakbim/ubyt_2014_dergi_listesi.xls, respectively.

³ For journals without JCR article influence scores, TUBITAK will continue to use its old algorithm based on five-year impact factors and cited half-lives of journals.

⁴ See the 2014 list of journals at http://ulakbim.tubitak.gov.tr/sites/images/Ulakbim/ubyt_2014_dergi_listesi.xls.

Figure 1. TUBITAK's support to Archaeology journals (2012-2014)



The fluctuation in the amount of support to each of the top Archaeology journals of 2012 can easily be followed in the subsequent years (Table 1 and Fig. 1). In spite of an almost two-fold increase in the amount of support, the average support has actually decreased from 2,600 Turkish Lira to 1,897 Turkish Lira in 2013. In 2014, the average support has increased a little (2,254 Turkish Lira), although it was still below that of 2012. The correlation between the amount of support and the ranks of journals in 2013 and 2014 is not high (Pearson's $r = .58$; Spearson's $\rho = .60$).

One cannot of course use these findings based on a limited number of Archaeology journals to generalize about the quality of TUBITAK's algorithms. However, this finding does not seem to be an isolated incidence because similar results were obtained for a total of 286 Geology journals. More than half (56%) the Geology journals -including one of the most prestigious ones, *Tectonics*, ranked lower than their previous ranks, and almost half the journals (49%) were misranked (Yaltırak, 2014, p. 18).

It should also be pointed out that the two algorithms used in 2013 and 2014 are not that different from each other after all. They are based on JCR's JIFs, journal citation half-lives and article influence scores. Papers published in high impact journals (i.e., high JIFs) usually have high AISs. Arendt (2010) carried out a study based on 5,900 journals listed in JCR Science Edition (2007) to find out if AISs vary by discipline, as is the case for JIFs. She found a statistically significant correlation for all disciplines between the AISs and JIFs (Pearson's $r(172) = 0.896$, $p < 0.001$). Disciplines with higher JIFs also have higher AISs, and AISs vary by discipline, too. For instance, more than 8.5-fold difference was observed between the disciplines with the highest and lowest median article influence scores (the difference was 9.6 fold for JIFs). Arendt (2010) concluded that both metrics were developed to help evaluate the journals and they should not be used formulaically to determine the amount of support to be granted to the departments or to rank the research personnel or individual articles.

TUBITAK's use of JIFs, cited half-lives and article influence scores to measure the quality of individual papers seems unwarranted, and changing the algorithm twice in three years tends to erode the trust

that might have so far built up. The relevant literature should be reviewed to find out the characteristics of different bibliometric measures. Moreover, TUBITAK's support program should be re-evaluated critically. The existence of the support program has been justified over the years on the basis of the positive correlation between the increase in the amount of total support and the number publications. Correlation between the two variables may not necessarily mean causation. After all, there is a positive correlation between the total amount of support provided by TUBITAK and the total number of publications appeared in low impact journals. TUBITAK supported fewer papers in the recent years even though the total amount of support did not change much. But this policy seems to have had a limited effect on authors in terms of encouraging them to publish in more prestigious journals with higher JIFs. The increase in the number of publications may well be due to HEC's earlier requirement of publishing papers in indexed journals or to the increasing number of researchers employed in the newly-established universities. The cause(s) of the increase in the number of publications and the role of the TUBITAK's support program in it needs to be studied carefully before more resources are committed to the program.

Use of H Index in Research Evaluation

Proposed by Hirsch (2005) as an alternative to the more traditional JIF, the h index is meant to say something about the life-time scientific success of a researcher by means of productivity and impact. A researcher has an h index of n if s/he published n papers in ISI-indexed journals each of which receiving at least n citations. H index became very popular in a short period of time as it was easy to calculate and has since been used to measure the performance of not only researchers but also of universities, publishers or even single articles (Schubert, 2009). In time, it has also been used for academic performance evaluation.

However, the h index also has some shortcomings. It does not take co-authors into account (Hirsch, 2007, p. 19193). Some researchers think that the h index does not meet some logical requirements and is not a first rate intellectual achievement but, rather, a "clever find" (Rousseau, García-Zorita & Sanz-Casado, 2013, p. 299). Moreover, the h index can be affected from the policy changes of commercial companies offering citation services such as Thomson Reuters. For instance, when Elsevier's Scopus entered the market with broader coverage of journals than that of Thomson Reuters (circa 16,000 journals as opposed to 9,000), Thomson Reuters decided in 2006 to expand its coverage by adding more regional journals and, consequently, by increasing not only the total number of journals indexed but also their h index.

As the h index is based on citation data and the citation rates of papers vary from discipline to discipline, the h indexes of researchers also vary. For example, the top most cited paper in Science published between 2008 and 2012 garnered over 1,000 citations whereas its equivalent in Economics received only 60 citations (Sgroi & Oswald, 2013, p. F256). In fact, the average JIFs tend to vary even in the subfields of Science. There is a considerable difference between average JIFs of Chemistry journals as opposed to that of Mathematics, which is due to the fact that the number of researchers (hence the potential number of researchers who would cite a given paper) and the number of journals in which they can publish are unequal. The average h indexes of researchers in these disciplines clearly reflect this.

Furthermore, the h index is closely related with time: while senior professors continue to increase their h indexes based on citations that they receive for their older publications, junior ones need more time to not only publish more papers but also garner citations thereto in order to boost their h indexes. As h indexes are not normalized according to seniority and research fields, this limits the use of h indexes to measure the performance of researchers and compare them across academic positions (assistant, associate and full professors) and across disciplines.

The h index tends to get used for tenure decisions, too. Using it for tenure decisions with limited time frames (i.e., appointment as associate or full professor) is especially inappropriate. It seems tempting for some universities to devise academic performance criteria based on the h index and specify the h index requirements for assistant, associate and full professors. This may be due to the fact that Hirsch himself suggested certain thresholds of h index coefficients to become a tenured or full professor in top universities as well as to be selected as a member of the National Academy of Sciences in the United States (Hirsch, 2005). (To be fair to Hirsch, he also explicated the caveats of using only the h index for such decisions.) This approach completely ignores the dynamics of publication patterns and citation behaviors in different disciplines.

In fact, h index should not be used in Social Sciences for academic performance evaluation at all. For instance, JIFs of Social Sciences journals are low: almost 60% of journals in Social Sciences have JIFs between 0 and 1, and a further 28% between 1 and 2 (Al & Soydal, 2014). Papers that appear in Social Sciences journals are hardly cited, if at all, in the year that they are published. Hence the immediacy indexes of Social Sciences journals are rather low, too. It takes much longer for papers published in Social Sciences journals to collect citations as the literature obsolesces more slowly in Social Sciences than that in Science, Technology, Engineering and Medicine (STEM). If a paper in Medicine, for example, does not get cited in the first couple of years after its publication, it is less likely that it would get cited at all in the following years. Social Sciences papers on the other hand tend to get cited in relatively longer time periods, thereby prolonging journal citation half-lives. It is not uncommon for Social Sciences journals to have citation half-lives over 5 years (i.e., half the citations for an average paper would accrue within the first 5 years after its publication). Therefore, it is often suggested that the five-year window, rather than the current two-year window, be used to calculate JIFs in Social Sciences.

What does all of the above mean for an academic waiting to meet the criterion of h index, say, 2 to be appointed as an associate professor in Social Sciences, for instance? (Hacettepe, 2013). It means that, by definition, she should at least have 2 papers published in ISI-indexed journals. It means that her papers published in Social Sciences journals with, say, JIF 1 would be cited once in 2 years after their publication. It is more likely, however, that it would take longer. If we suppose that the half-lives of journals in which her papers appeared are 5 years, it may take up to 10 years to get a single citation for each of her papers. Even this is not guaranteed. Note that for an h index of 2, each paper should be cited at least twice. This process would be even more difficult, if not impossible, for a candidate who seeks to meet the h index of 3 to be appointed as a full professor. Clearly, such waiting periods to meet the h index criterion with no guaranteed outcome are hardly acceptable.

As indicated earlier, that just because the h indexes of researchers can easily be obtained through Thomson Reuters's Web of Science, Elsevier's Scopus or Google Scholar does not necessarily justify its use in research assessment. H index does not mean much in Social Sciences wherein scholarly communication takes place, in general, through monographs rather than journal articles. Whether in Sciences or Social Sciences, the h index should not be used for comparative purposes. After all, how can one distinguish the quality of scholarship of the two candidates with low h indexes of, say, 2 and 3, during tenure decisions?

There is a tendency to use the citation rates and the h index to predict the future Nobel laureates. For instance, every year since 2002 Thomson Reuters identifies what is called "Citation Laureates" in each category of Nobel prizes (except Literature and Peace) based on the number of citations to their works and tries to predict the winners (Pendlebury, 2009). Since the h index measures the productivity and the cumulative impact of scholars, Hirsch (2005) computed the h indexes of Nobel laureates. He found that the h indexes of winners peaked between 35 and 39. The h index can to a certain extent be used as an indicator of life-time achievement of scholars in some disciplines (i.e., Life Sciences) where h indexes tend to be relatively higher. However, there seems to be no direct correlation between the h index and winning a Nobel Prize (Marques, 2013; Van der Wall, 2011). For instance, the current h index of Peter Higgs, the 2013 Nobel Prize winner in Physics for his discovery of Higgs boson, is 10!

His m index (h index divided by the length of his career in years) would even be below that of “a successful scientist” let alone a “truly unique individual” as classified in Hirsch’s original paper (Hirsch, 2005).⁵ Professor Higgs thinks that academics nowadays are expected to “keep churning out papers” and that he “wouldn’t be productive enough for today’s academic system” (Aitkenhead, 2013).

Conclusions

We quoted Albert Einstein at the beginning of this paper regarding his view of the “frequentist” approach. Goodhart’s Law states that “When a measure becomes a target, it ceases to be a good measure” (http://en.wikipedia.org/wiki/Goodhart%27s_law). The citation counts or the h index as measures seem to be no exception. As academic institutions or research funders introduce new criteria for research assessment and tenure decisions such as JIFs and the h index, journal publishers and researchers try to anticipate what the effect would be on them and adapt their policies accordingly. The measure would then no longer function properly and lose its “information content” (Pendlebury, 2009).

Just because bibliometric measures such as JIFs and h indexes are readily available through Web of Science, Scopus or Google Scholar do not make them ideal measures to use for research assessment. JIFs, for example, seem to become “the poor man’s citation analysis” (Marx & Bornmann, 2013). The “fatal attraction” of bibliometric measures (Van Raan, 2005) that were not developed for research evaluation, tenure and publication support may have adverse effects on academic careers of researchers (Hudson & Laband, 2013, p. F201). We have already referred above to the San Francisco Declaration on Research Assessment (DORA) advising not to use bibliometric measures for research assessment. More recently, the Board of Directors of IEEE, “the world’s largest professional association for the advancement of technology” (ieee.org), adopted the statement that concludes: “. . . **bibliometric performance indicators should be applied only as a collective group (and not individually), and in conjunction with peer review following a clearly stated code of conduct**” (IEEE, 2013, original emphasis). We should pay heed to such recommendations and take the combination of peer review and bibliometrics as the “ideal way of research evaluation” (Bornmann & Leydesdorff, in press).

References

- 2014 yılı UBYT Programı teşvik miktarları hesaplama yöntemine dair bilgi notu. (2014). Retrieved, November 2, 2014, from http://ulakbim.tubitak.gov.tr/sites/images/Ulakbim/ubyt_2014_hesap.pdf.
- Aitkenhead, D. (2013, December 6). Peter Higgs: I wouldn’t be productive enough for today’s academic system. *The Guardian*, Retrieved, October 26, 2014, from <http://www.theguardian.com/science/2013/dec/06/peter-higgs-boson-academic-system>.
- Al, U. & Soydaş, İ. (2014). Akademik atıf dizinleri ile savaşı. *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, 31(1), 23-42. Retrieved, November 4, 2014, from <http://yunus.hacettepe.edu.tr/~umutal/publications/war.pdf>.
- Arendt, J. (2010). Are article influence scores comparable across scientific fields? *Issues in Science and Technology Librarianship*, No. 60. Retrieved, June 25, 2014, from <http://www.istl.org/10-winter/refereed2.html>.
- Batmaz, A. (2013, June 14). Türkiye’de bilim üretimi ve arkeoloji. *Cumhuriyet Bilim ve Teknoloji*, (1369), 18. Retrieved, November 2, 2014, from <http://www.arkeolojikh Haber.com/?p=2569>.
- Bornmann, L. & Leydesdorff, L. (in press). Scientometrics in a changing science landscape. *EMBO Reports*.
- Casadevall, A. & Fang, F.C. (2014). Causes for the persistence of impact factor mania. *mBio*, 5(2). Retrieved, June 25, 2014, from <http://mbio.asm.org/content/5/2/e00064-14.full.pdf>.
- Doğan, M.H. & Soylak, M. (2014, October 10). Hangi bilim insanlarımızın başarıları hızla artıyor? *Cumhuriyet Bilim ve Teknoloji Dergisi*, No. 1438, pp. 10-11.
- Garfield, E. (1994, June 20). The impact factor. *Current Contents*, (25), 3-7. Retrieved, October 30, 2014, from <http://wokinfo.com/essays/impact-factor/>

⁵ It is sometimes implicitly assumed that there is a threshold of h index of about 80 to win a Nobel prize in Life Sciences (e.g., Doğan & Soylak, 2014). However, this is not the case.

- Hacettepe Üniversitesi Öğretim Üyeliğine Yükseltme ve Atama Kriterleri Taslağı Genel (GSF ve ADK Hariç) – (08.10.2014). (2014). Retrieved, November 2, 2014, from <https://www.hacettepe.edu.tr/duyuru/rekduy/GENELKRITERLERTASLAK081014.doc>.
- Hacettepe Üniversitesi Öğretim Üyeliğine Yükseltme ve Atama Kriterleri Taslağı Sosyal Bilimler Alanları. (2013). 25 Haziran 2013. 30 Ocak 2014 tarihinde <http://www.hacettepe.edu.tr/akademik/taslak/GENEL.pdf> adresinden erişildi.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *PNAS*, 102(46), 16569-16572.
- Hirsch, J.E. (2007). Does the h index have predictive power? *PNAS*, 104(49), 19193-19198.
- Hudson, J. & Laband, D.N. (2013). Using and interpreting journal rankings: Introduction. *The Economic Journal*, 123(570), F199–F201.
- IEEE. (2013, September 9). Appropriate use of bibliometric indicators for the assessment of journals, research proposals, and individuals. Retrieved, November 4, 2014, from http://www.ieee.org/publications_standards/publications/rights/ieee_bibliometric_statement_sept_2013.pdf.
- Marques, F. (May 2013). The limits of the h-index. *Revista Pesquisa FAPESP*, Edition 207. Retrieved, October 26, 2014, from <http://revistapesquisa.fapesp.br/en/2013/06/25/the-limits-of-the-h-index/>.
- Marx, W. & Bornmann, L. (2013). Journal Impact Factor: “the poor man's citation analysis” and alternative approaches. *European Science Editing*, 39(2), 62-63. Retrieved, June 25, 2014, from <http://www.ease.org.uk/sites/default/files/aug13pageslowres.pdf>.
- Pendlebury, D. (2009 November). Discover the power of quantitative analysis: The art and science of identifying future Nobel laureates (slides). Retrieved, October 26, 2014, from http://ip-science.thomsonreuters.com/m/pdfs/Identifying_Nobel_Laureates.pdf
- Rousseau, R., García-Zorita, C. ve Sanz-Casado, E. (2013). The h-bubble. *Journal of Informetrics*, 7, 294-300.
- San Francisco Declaration on Research Assessment: Putting science into the assessment of research. (2012, December 16). Retrieved, June 25, 2014, from <http://am.ascb.org/dora/files/SFDeclarationFINAL.pdf>.
- Schubert, A. (2009). Using the h-index for assessing single publications. *Scientometrics*, 78(3), 559-565.
- Seglen, P.O. (1997, February 15). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314(7079):498-502. Retrieved, June 25, 2014, from <http://www.dcsience.net/seglen97.pdf>.
- Sgroi, D. ve Oswald, A.J. (2013). How should peer-review panels behave? *The Economic Journal*, 123(570), F255–F278.
- Tonta, Y. & Ünal, Y. (2008). Dergi kullanım verilerinin bibliyometrik analizi ve koleksiyon yönetiminde kullanımı. *Türk Kütüphaneciliği*, 22(3): 335-350.
- TÜBİTAK Türkiye Adresli Uluslararası Bilimsel Yayınları Teşvik Programı Uygulama Esasları. (2013). Retrieved, November 2, 2014, from http://www.tubitak.gov.tr/sites/default/files/esaslar_v_2_vers.2_2.pdf.
- UBYT Programı 2012 yılı teşvik miktarları. (2012). Retrieved, June 25, 2014 from <http://www.ulakbim.gov.tr/cabim/ubyt/tesvik.uhtml>.
- UBYT Uygulama Esasları'nda değişiklik. (2012). Retrieved, June 25, 2014 from <http://www.ulakbim.gov.tr/cabim/ubyt/haberler.uhtml>.
- Van der Wall, E.E. (2011). The Hirsch index: more than a number (editorial). *The Netherlands Heart Journal*, 19(5), 209-210. Retrieved, November 4, 2014, from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3087030/>.
- Van Raan, A.F.J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133–43.
- Yaltrak, C. (2014, June 21). TÜBİTAK yayın teşvik sistemini değiştirmeli! *Cumhuriyet Bilim ve Teknoloji*, (1409), 18.