Appropriate Use of Bibliometric Indicators for the Assessment of Journals, Research Proposals, and Individuals

(*Adopted by the IEEE Board of Directors 9 September 2013)*

Bibliometric indicators provide numerical scales that are intended to quantitatively determine the value of scientific research and the scholarly publication in which that research is published. Since scientific performance cannot, of course, be directly "measured", citations acquired by each published paper are assumed as a proxy for quality, without prejudging the reasons for the citations.

The application of bibliometrics to quantify the significance of individual journals dates back several decades [1] and the field has now reached a sufficiently high level of maturity to recognize that the scientific impact of journals as evaluated by bibliometrics is a complex, multi-dimensional construct and therefore **more than one indicator is needed for such evaluation** [2]-[4]. Nearly all commonly used bibliometric indices [1],[5]-[7] can be classified fundamentally as measuring either popularity or prestige, two concepts for which citation behaviors are valued in different and complementary ways. These indices also offer differing consideration of self-citations and have various levels of susceptibility to potential manipulation. As such, use of a single bibliometric index to rank, evaluate, and value journals is inappropriate. Rather, the use of multiple metrics with complementary features provides a more comprehensive view of journals and their relative placements in their fields.

Recently, citation counts and proxies thereof have risen in importance as fundamental elements in the determination of the scientific impact of entire departments or universities and research centers [8], funding evaluations of research proposals and the assessment of individual scientists for tenure and promotion [9], salary raises [10], or even to predict future career success [11]. While the first use is technically appropriate, provided it relies on data collected from a sufficiently large set to provide a statistically meaningful analysis, this condition is never satisfied when applied to individual scientists.
Furthermore, while using data appropriate for an individual researcher (such as average citation count or h-index and its variations [12]) can provide information to be adopted *in conjunction with other measures* to form a comprehensive evaluation, the use of the bibliometric index of a journal in which a researcher publishes (typically the Impact Factor (IF)) as a proxy for the quality of his/her specific paper is a common example of a technically incorrect use of bibliometrics [13][29]. There is, in fact, no guarantee that every single article published in a high-impact journal, as determined by any particular metric, will be of high quality and highly cited.
Measuring individual impact by using journal bibliometric indicators is worse when comparing researchers in different areas. In fact, citation practices vary significantly across disciplines and even sub-disciplines, and similarly the number of scientists (and authors) contributing to a specific field can be substantially different. This can result in large numerical differences for some bibliometric indicators (the IF in particular) that have no correlation with the actual scientific quality of the corresponding journals.
When based upon such data as a measurement of "scientific quality," decisions by research funding agencies or by tenure/promotion committees can be misguided and biased.

Such technically incorrect use of bibliometric indices is a problem of severe concern in the scholarly community. Many scientists and science organizations in US, Europe and Australia have raised concerns about or taken strong positions against purely numerical assessment based on bibliometrics (see, e.g., [14]-[18],[29]), highlighting the potential unintended and adverse consequences of these practices. They have proposed clear recommendations on the correct use of such indices [19][29], and described best practices for using peer review in the assessment of scientists and research projects [20]-[23]. A common conclusion is the recognition of the need to use multiple indicators as well of the importance of peer review in the assessment of research quality, which resulted in the **recommendation that 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**.

The IEEE, in its leading position as the world's largest professional association dedicated to advancing technological innovation and in its desire to fulfill its primary mission of fostering technological excellence for the benefit of humanity, recognizes the above concerns about the inappropriate application of bibliometrics to the evaluation of both scientists and research proposals.

More specifically, the IEEE endorses the following tenets in conducting proper assessment in the areas of Engineering, Computer Science and Information Technology:

1. The use of multiple complementary bibliometric indicators [2]-[4] is fundamentally important to offer an appropriate, comprehensive and balanced view of each journal in the space of scholarly publications. The IEEE has recently adopted the *Eigenfactor* and the *Article Influence* [5] in addition to the IF for the internal and competitive assessment of its publications [24] and welcomes the adoption of other appropriate complementary measures at the article level, such as those recently introduced in the framework of the so-called altmetrics [25], once they have been appropriately validated and recognized by the scientific community.
2. Any journal-based metric **is not designed to capture qualities of individual papers and must therefore not be used as a proxy for single-article quality or to evaluate individual scientists** [26]-[28]. All journals' bibliometric indices are obtained by averaging over many papers, and it cannot be assumed that every single article published in a high-impact journal, as determined by any particular journal metric, will be highly cited.
3. While bibliometrics may be employed as a source of additional information for quality assessment **within a specific area of research**, the primary manner for assessment of either the scientific quality of a research project or of an individual scientist **should be peer review**, which will consider the scientific content as the most important aspect, and also the publication expectations in the area, and the size and practice of the research community.

The IEEE also recognizes the increasing importance of bibliometric indicators as independent measures of quality or impact of any scientific publication and therefore **explicitly and firmly condemns** any practice aimed at influencing the number of citations to a specific journal with the sole purpose of artificially influencing the corresponding indices.

References

[1] E. Garfield, "Citation analysis as a tool in journal evaluation", *Science,* vol. 178, pp. 471-479, 1972

[2] C. Neylon, S. Wu, "Article-level metrics and the evolution of scientific impact'", *Plos Biology*, vol 7, n. 11, e1000242, doi:10.1371/journal.pbio.1000242, 2009

[3] J. Bollen, H. van de Sompel, A. Hagberg, R. Chute, "A principal component analysis of 39 scientific impact measures," *PloS One*, vol. 4, n. 6, e6022, 2009.

[4] L. Leydesdorff, "How are New Citation-Based Journal Indicators Adding to the Bibliometric Toolbox?", *Journal of the American Society for Information Science and Technology*, vol. 60, pp. 1327-1336, 2008

[5] J. D. West, T. C. Bergstrom, C. T. Bergstrom, "The Eigenfactor Metrics: A network approach to assessing scholarly journals," *College of Research Libraries*, vol. 71, pp. 236-244, 2010

[6] B. Gonzalez-Pereira, V. P. Guerrero-Bote, F. Moya-Anegon, "A new approach to the metric of journals scientific prestige: The SJR indicator," *Journal of Informetrics* vol 4, pp. 379-391, 2010

[7] H. F. Moed, "Measuring contextual citation impact of scientific journals", *Journal of Informetrics*, vol. 4 pp. 265-277, 2010

[8] L. Waltman, C. Calero-Medina, J. Kosten, E. C.M. Noyons, R. J.W. Tijssen, N. Jan van Eck, T. N. van Leeuwen, A. F.J. van Raan, M. S. Visser, P. Wouters, "The Leiden Ranking 2011/2012: Data collection, indicators, and Interpretation", arXiv:1202.3941, 17 February 2012

[9] S. Lehmann, A. D. Jackson, B. E. Lautrup, "Measures for Measures", *Nature*, vol. 444, pp. 1003-1004, 21/28 December 2006

[10] J. Shao, H. Shen, "The outflow of academic papers from China: why is it happening and can it be stemmed?", *Learned Publishing*, vol 24, n. 2, pp. 95-97(3), April 2011

[11] D. E. Acuna, S. Allessina, K. P. Kording, "Future Impact: Predicting Scientific Success", *Nature*, vol 489, pp. 201-202

[12] S. Alonso, F. Cabrerizo, E. Herrera-Viedma, F. Herrera, "h-index: a Review Focused in its Variants, Computation and Standardization for Different Scientific Fields", *Journal of Informetrics*, vol 3, pp. 273-289, 2009

[13] G. F. Gaetani, A. M. Ferraris, “Academic Promotion in Italy”, *Nature*, vol. 353, pp. 10, 1991
[14] P. Lawrance, “The Politics of Pubblications”, *Nature*, vol. 422, pp. 259-261, 2003
[15] P. Lawrance, “The Missmeasurement of Science”, Current Biology, vol 17, n.15, p.. R583-R585, 2007
[16] F. Guilak, C. R. Jacobs, “The h-index: Use and Overuse”, Journal of Biomechanics, vol. 44, pp 208-209, 2011
[17] A. Abbott , D Cyranoski , N. Jones , B. Maher, Q. Schiermeier, R. Van Noorden, “Metrics: Do metrics Matter?”, *Nature*, vol. 465, pp. 860-862, 2010
[18] National Health and Medical Research Council, Australia, April 2010, “NHMRC Removes Journal Impact Factor from Peer Review of Individual Research Grant and Fellowship Applications”, http://www.nhmrc.gov.au/_files_nhmrc/file/grants/peer/impact%20factors%20in%20peer%20review.pdf
[19] Institut de France, Académie des Sciences, “On the proper Use of Bibliometrics to Evaluate Individual Researchers”, 17 January 2011, http://www.academie-sciences.fr/activite/rapport/avis170111gb.pdf
[20] European Science Foundation, European Peer Review Guide, Integrating Policies and Practices for Coherent Procedures, March 2011 (http://www.esf.org/activities/mo-fora/peer-review.html)
[21] Peer review in scientific publications, Science and Technology Committee, House of Commons, UK, 18 July 2011 (http://www.publications.parliament.uk/pa/cm201012/cmselect/cmsctech/856/85602.htm)
[22] Swedish Research Council, Quality Assessment in Peer Review, 5 November 2009 (www.cm.se/webbshop_vr/pdfer/2011_01L.pdf)
[23] European Physics Society, “On the use of Bibliometric Indices During Assessment”, http://c.ymcdn.com/sites/www.eps.org/resource/collection/B77D91E8-2370-43C3-9814-250C65E13549/EPS_statement_June2012.pdf
[24] http://www.ieee.org/publications_standards/publications/journmag/journalcitations.html
[25] http://altmetrics.org/manifesto/
[26] P. Campbell, “Escape from the Impact Factor”, *Ethics in Science and Environmental Politics*, vol 8, pp. 5-7, 2008
[27] P. O. Seglen, “Why the Impact Factor of Journals Should not be Used for Evaluating Research”, BMJ, vol 314, 15 February 1997
[28] P. O Seglen, “Causal Relationship Between Article Citedness and Journal Impact”, *Journal of the American Society for Information Science*, vol 45, pp. 1-11, 1994
[29] “San Francisco Declaration on Research Assessment (DORA)”, http://am.ascb.org/dora/, 2013