

12-1-2014

Standardizing research metrics and indicators – Perspectives & approaches

Gali Halevi Dr
Elsevier

Lisa Colledge Dr
Elsevier

Follow this and additional works at: <https://www.researchtrends.com/researchtrends>

Recommended Citation

Halevi, Gali Dr and Colledge, Lisa Dr (2014) "Standardizing research metrics and indicators – Perspectives & approaches," *Research Trends*: Vol. 1 : Iss. 39 , Article 6.

Available at: <https://www.researchtrends.com/researchtrends/vol1/iss39/6>

This Article is brought to you for free and open access by Research Trends. It has been accepted for inclusion in Research Trends by an authorized editor of Research Trends. For more information, please contact r.herbert@elsevier.com.

Section 5:
Expert Opinion

Standardizing research metrics and indicators – Perspectives & approaches

Dr. Gali Halevi & Dr. Lisa Colledge, Elsevier

The increase of data availability and computational advances has led to a plethora of metrics and indicators being developed for different levels of research evaluation. Whether at the individual, program, department or institution level, there are numerous methodologies and indicators offered to capture the impact of research output. These advances have also highlighted the fact that metrics must be applied appropriately depending on the goal and subject of the evaluation and should be used alongside qualitative inputs, such as peer review.

However, this has not solved the challenge of finding core quality and validity measures that will guide the current and future development of evaluative metrics and indicators. While innovation in the field of research metrics is ongoing, funders, institutions and departments are already using output metrics to measure specific elements. Such metrics as are being used cannot be scaled up to global indicators, however. This means that the field now faces a divide: although new metrics exist, they are oftentimes not suitable or cannot be scaled up to the global research ecosystem. Therefore, evaluators still use metrics that have already been recognized as unsuitable measures of individuals' performance, such as journal-level indicators. But for lack of agreed upon alternatives, such metrics are being used routinely in inappropriate circumstances despite their shortcomings.

The need for quality and validity measures that will guide the development of research metrics and ensure that they are applied in an appropriate and fair way is at the heart of several discussions carried out via conferences and listservs, especially in the Scientometrics, Science Policy, and Research Funding communities.

One such panel discussion was held at the [Science and Technology Indicators \(STI\) 2014 conference in Leiden](#). The panel focused on the need for standardization in the field of research metrics that will speak to their validity, quality and appropriate use and ways to arrive at a consensus. The panel consisted of Dr. Lisa Colledge (Elsevier, Director of Research Metrics), [Stephen Curry](#) (professor of Structural Biology at Imperial College, London, and member of HEFCE [Higher Education Funding Council for England](#) steering group on the use of research metrics in performance measurement), [Stefanie Hausteine](#) (University of Montreal), [Jonathan Adams](#) (Chief Scientist at Digital Science), and [Diana Hicks](#) (Georgia Institute of Technology).

The [Snowball Metrics initiative](#) (1), presented by Dr. Lisa Colledge, is an example of research universities collaborating internationally to arrive at a commonly agreed upon set of measures of research (outputs as well as other aspects of the research process). Snowball Metrics' aim is for universities to agree on a set of metrics methodologies that give strategic insight into all of a university's activities. These metrics should be understood by everyone in the same way, so that when universities calculate metrics using these "recipes" they all follow the same protocol (2).

Lisa emphasized that Snowball Metrics welcomes feedback from the research community, to improve the existing recipes and to expand the set of recipes available. Elsevier is involved in Snowball Metrics at the invitation of the universities who drive it, to project manage and to provide technical expertise where needed. The Snowball Metrics program has responded to the HEFCE review (3), and this initiative has significantly influenced Elsevier's overall approach to the use of research metrics, expressed in a response to the same HEFCE review (4). The main principles of Elsevier's manifesto are:

1. A set of multiple metrics distributed across the entire research workflow is needed.
2. Metrics must be available to be selected for all relevant peers.
3. The generation and use of metrics should be automated and scalable.
4. Quantitative information provided by metrics must be complemented by qualitative evidence to ensure the most complete and accurate input to answer a question.
5. The combination of multiple metrics gives the most reliable quantitative input.
6. Disciplinary and other characteristics that affect metrics, but that do not indicate different levels of performance, must be taken into account.
7. Metrics should be carefully selected to ensure that they are appropriate to the question being asked.
8. We cannot prevent the inappropriate or irresponsible use of metrics, but we can encourage responsible use by being transparent, and intolerant of "gaming".
9. Those in the research community who apply metrics in their day-to-day work, and who are themselves evaluated through their use, should ideally define the set of metrics to be used. It is highly desirable that this same community, or those empowered by the community on their behalf, maintains the metric definitions.
10. There should be no methodological black boxes.
11. Metric methodologies should be independent of the data sources and tools needed to generate them, and also independent of the business and access models through which the underlying data are made available.
12. Aggregated or composite metrics should be avoided.

Dr. Ian Viney, Director of [Strategic Evaluation and Impact, Medical Research Council](#), supports this approach, saying that "standards, at least properly described metrics, are important if you want to have reproducibility for your analyses, across different organizations and/or timescales. Evaluation of research, is itself research and development – success and failure should be properly documented." Therefore, "recipes' should be available for discussion; testing and modification and effective approaches should become accepted standards – methods that everyone can apply." Dr. Viney also commented on the gap between research metrics and the research community saying that:

"The link between these outputs and research activity or impact is little understood. What is most interesting is the development of metrics relating to other logically important areas of research activity – e.g. the ways in which researchers influence policy setting processes, or research feeds into policies, the way in which research teams develop new processes and products, the way in which research materials are disseminated and used. We can make a good argument that these activities are intermediate indicators of impact, they logically describe steps along a pathway to impact. They describe activities however not well reported in any standard format, and data is not readily available on these outputs."

Dr. Ian Viney:

"We should be open about our methods, discussion across stakeholders is helpful, and work such as Snowball Metrics will help accelerate the field. I will be convinced that a particular method should become a standard when it has been successfully and reproducibly applied, when it helps us better understand research progress, productivity and/or quality."

The scientometrics community should provide expert advice to stakeholders regarding the development of suitable approaches. This community has a central role in proposing the most promising methods for wider use."

[Dr. Jonathan Adams, Chief Scientist at Digital Science](#), who participated in the panel, cautioned against rigid setting of standards. In his view "It is infeasible to set comprehensive written standards for metrics, indicators or evaluation methodologies when there is a diverse range of contexts, cultures and jurisdictions in which they might be applied and when data access and data diversity are changing very rapidly." Therefore, his opinion is that any attempt to create such standards would create "an artificial vision of security and stability" that might be used inappropriately by research agencies and managers.

[Dr. Paul Wouters](#), Director of [The Centre for Science and Technology Studies \(CWTS\)](#) and professor of Scientometrics, added his concern regarding standardization of metrics stating that "standards may be important for the construction of databases of research products. So at the technical level they can be useful. However, standards can mislead users if they are essentially captured by narrow interests."

Following the conference, CWTS, a part of Leiden University, published the ["The Leiden Manifesto in the Making: proposal of a set of principles on the use of assessment metrics"](#) (5).

In the manifesto Paul Wouters, Sarah de Rijcke and their colleagues summarized some principles around which the debate about standardization and quality should revolve:

1. There should be a connection between assessment procedures and the primary process of knowledge creation. If such a connection doesn't exist then the assessment loses a part of its usefulness for researchers and scholars.
2. Standards developed by universities and data provided should be monitored and benefit from the technical expertise of the Scientometrics community. Although the Scientometrics community does not want to set standard themselves, it should take an active part in documenting them and ensuring their validity and quality.
3. There's a need to strengthen the working relationship with the public nature of the infrastructure of meta-data, including current research information systems, publication databases and citation indexes including those available from for-profit companies.
4. Taking these issues together provides an inspiring collective research agenda for the Scientometrics community.

Dr. Wouters added that the main motivation should be to “prevent misuse or harmful applications by deans, universities or other stakeholders in scientometrics. Although many studies in scientometrics suffer from deficient methods, this problem cannot be solved with standards, but only with better education and software (which may build on some technical standards).”

Dr. Paul Wouters:

“I do not think that global standards are currently possible or even desirable, therefore, Principles of good evaluation practices: YES. Universal technical standards: NO.

The Scientometrics community should analyze, train, educate, clarify, and also take on board the study of how the Scientometric indicators influence the conduct of science and scholarship.”

Dr. Peter Dahler-Larsen, a professor at the Department of Political Science at the University of Copenhagen, recently contributed to **The Citation Culture** blog on the topic of development of quality standards for Science & Technology indicators. Dr. Dahler-Larsen commented on his contribution to *Research Trends*, saying that “it is important to follow the discussion of standards, because in some fields standards pave the way for a particular set of practices that embody particular values - for better or for worse.” The main motivation for the development of standards, added Dr. Dahler-Larsen, is “NOT their agreed-upon character” but rather their ability to “inspire ethical and methodological awareness, and this can take place even without much consensus.” Yet, Dr. Dahler-Larsen says that in spite of their importance he “does not have high hopes about the adoption of standards in policy-making.”

Dr. Dahler-Larsen:

“The most important function of standards is to raise awareness and debate. Standards can be helpful in discussion of problematic policy-making initiatives.

The Scientometric community has an important role to play because presumably, they comprise experts who knows about the pros and cons and pitfalls related to particular measurement approaches etc. Their accumulated experience should inform better practice.”

Dr. John T. Green, who chairs the Snowball Metrics Steering Committee, believes that “whilst some argue that it is impossible to define or agree standard metrics because of the diverse range of contexts and geographies, like it or not, funders and governments are using such measures - some almost slavishly and exclusively (as in Taiwan to allocate government funding). Therefore, whilst it is ideologically acceptable for the Scientometric community to take the high ground and claim that because metrics cannot be perfect therefore none should be developed, to do so is ignoring reality – let us at least do our best and develop metrics as best we can (as indeed has happened over time with bibliometrics). I believe it is important for the academic community to engage and ensure that whatever is used to measure them is fit for purpose, or as fit as can be, especially given that they should never be used in isolation – metrics are only a part, albeit an important part, of the evaluation landscape. Thus the approach of Snowball – bottom-up and owned by the academic community.”

Professor Jun Ikeda, Chief Advisor to the President of the **University of Tsukuba**, Japan, supports the development of standards in metrics. In his view they will save researchers time when reporting to funders. Prof. Ikeda pointed to the fact that in many cases there is a real difficulty to compare universities’ performance and says that “If every university defines things in their own way, and calculates metrics in their own way, then seeing a metric that is higher or lower than someone else’s is meaningless because the difference might not be real, but just due to different ways of working with the data. I want to do apples to apples comparisons, to be sure that I can be confident in differences that I see, and confident in taking decisions based on them.”

Research-focused universities need to be active in defining the metrics that they want to use to give insights into their strategies, Prof. Ikeda said. “Ideally the researchers within our universities would also support and use the same metrics to help them to promote their careers and to understand how they are performing relative to their own peers.” Although the debate about whether standardization in research metrics is necessary or even desirable, there is no doubt that the discussion of itself is of importance as it serves as an instrument to raise awareness about the complexity of the topic as a whole. Standards may not be easy to develop or implement, but there is little doubt that consensus regarding their proper use is needed. As more data becomes available and more metrics are developed,

the issue of their usefulness and accuracy in different settings becomes crucial. Data providers, evaluators, funders and the Scientometric community must work together to not only aggregate, calculate and produce metrics, but also test them in different contexts and educate the wider audience as to their proper use.

Prof. Ikeda:

“The biggest gap is for the research community to drive the direction that this whole area is going in. A lot is happening, but we feel a bit like it is all being done to us. There is space for us to take control of our own destiny, and shape things as we would like them to be, and as they make the most sense to us.”

References:

1. Colledge, L. (2014) “Snowball Metrics Recipe Book”, Available at: http://www.snowballmetrics.com/wp-content/uploads/snowball-recipe-book_HR.pdf
2. Snowball Metrics, “Why is this initiative important to the higher education sector?”, Available at: <http://www.snowballmetrics.com/benefits>
3. Snowball Metrics (2014) “Response to the call for evidence to the independent review of the role of metrics in research assessment”, Available at: <http://www.snowballmetrics.com/wp-content/uploads/Snowball-response-to-HEFCE-review-on-metrics-300614F.pdf>
4. Elsevier (2014) “Response to HEFCE’s call for evidence: independent review of the role of metrics in research assessment”, Available at: http://www.elsevier.com/_data/assets/pdf_file/0015/210813/Elsevier-response-HEFCE-review-role-of-metrics.pdf
5. CWTS (2014) “The Leiden Manifesto in the Making: proposal of a set of principles on the use of assessment metrics in the S&T indicators conference”, Available at: <http://citationculture.wordpress.com/2014/09/15/the-leiden-manifesto-in-the-making-proposal-of-a-set-of-principles-on-the-use-of-assessment-metrics-in-the-st-indicators-conference/>