

Let's make science metrics more scientific

Martin Fenner, Gobbledygook

March 29, 2010

In the March 25 edition of *Nature*, Julia Lane, Program Director of the Science of Science and Innovation Policy Program at the National Science Foundation, wrote an interesting opinion piece about the assessment of scientific performance. She argues that the current systems of measurement are inadequate, as they have several inherent problems and do not capture the full spectrum of scientific activities. Good scientific metrics are difficult, but without them we risk making the wrong decisions about funding and academic positions.

Julia Lane suggests that we develop and use standard identifiers both for researchers and their scientific output (examples given include the DOI for publications and ORCID as unique author identifier), that we develop standards for reporting scientific achievements (e.g. using the Research Performance Progress Report format), and that we open up and connect the various tools and databases that collect scientific output. She cites the Lattes database for Brazilian researchers as a successful example for systematically collecting scientific output. Another example given is the ongoing STAR METRICS project which measures the impact of federally funded research on economic, scientific and social outcomes.

The article emphasizes that is not enough to think about how to best collect and report scientific output, but that it is equally important to understand what these data mean and how to use them, and this may differ from field to field. Knowledge creation is complex and measuring this can not be reduced to counting scientific papers and the number of times they are cited. Social scientists and economists should be involved in this step. Julia Lane suggests an international platform supported by funding agencies in which ideas and potential solutions for science metrics can be discussed.

Flickr photo by jepoirrier.

The article contains a lot of food for thought and has already collected some insightful comments. In perfect timing, *Nature* this week not only made *Nature News* available without a subscription, but also added commenting to all their articles. I would like to add some thoughts on topics that were not covered because of space constraints and different perspective.

What are the standard identifiers for research output?

Using standard identifiers for research output is an essential first step, and the standard identifier for scientific papers is the DOI. So why is it that PubMed (the most important database for biomedical articles, published by the U.S. National Institutes of Health) still uses their own PMID and doesn't display the DOI in their abstract and summary views? And where is the DOI in abstracts, full-text HTML or PDF of articles published in the New England Journal of Medicine, to take just one popular medical journal as an example? Both PubMed and the NEJM obviously use the DOI, but why do they make it so difficult for others?

The unique author identifier ORCID was mentioned in the article (disclaimer: I am a member of the ORCID technical working group). There are many other initiatives for uniquely identifying researchers, most of them older than **ORCID** which was started in November 2009. But is very important that we can agree on a single author identifier that is supported by researchers, institutions, journals and funding organizations. **ORCID** already has support from a growing list of ORCID members and is our best chance for a widely supported and open unique author identifier. But this list of **ORCID** members is very short on funding organizations (with notable exceptions such as the **Wellcome Trust** and **EMBO**). What is holding them back, and that includes the **National Science Foundation** (where Julia Lane works) and the U.S. **National Institutes of Health** (NIH)?

Persistent identifiers are essential to attribute, cite and share primary research data sets. We have a long tradition for this with sequence data, and there is growing demand in other research areas, especially when huge amounts of data are collected (one example is PANGAEA for earth system research). DataCite is a new initiative that aims to improve the scholarly infrastructure around datasets, and to increase acceptance of research data as legitimate, citable contributions to the scientific record.

With the focus on research papers, we forget that we do not have standard identifiers for many aspects of scientific activity, including

- research grants
- principal investigator in clinical trials
- scientific prizes and awards
- invited lectures
- curation of scientific databases
- mentoring of students

How do we measure scientific output?

Citations are the traditional way to measure the impact of a scientific paper. Some of the problems with this approach are well-known and were for example highlighted in a 2007 editorial in the Journal of Cell Biology (Show me the data). We need a metric that is open and not proprietary, and that measures

the citations of an individual paper and not the journal as a whole. We should also not forget that the number of citations can't be compared between different fields.

A 2009 analysis by the MESUR project indicates that scientific impact of a paper can not be measured by any single indicator (A Principal Component Analysis of 39 Scientific Impact Measures). Alternatives to citations are usage statistics such as HTML page views and PDF downloads, popularity in social bookmarking sites, coverage in blog posts, and comments to articles. The PLoS article level metrics introduced in September 2009 combine these different metrics, and make the data openly available.

How best to measure the other aspects of scientific output is largely unknown. It is possible to count the number of research grants or the total amount of money awarded, but should we simply count the number of submitted research datasets, invited lectures, science blog posts, etc., or do we need some quality indicator similar to citations?

Why do we need all this?

Julia Lane emphasizes that we need science metrics to make the right decisions about funding and academic positions. And I fully agree with her that we need more research by social scientists and economists to better understand what these data mean and how best to use them. There is a lot of anecdotal evidence that suggests that science metrics alone may be poor indicators of future scientific achievements, simply because there are too many confounding factors. Maybe we also need to find a better term, as metric implies that scientific output can be reduced to one or more numbers.

Another important motivation for improving science metrics, and not mentioned in the article, is to reduce the burden on researchers and administrators in evaluating research. The proportion of time spent doing research vs. time spent applying for funding, submitting manuscripts, filling out evaluation forms, doing peer review, etc. has become ridiculous for many active scientists. Initiatives such as the standardized **Research Performance Progress Report** format mentioned in the paper or automated tools to create a publication list or CV can reduce this burden. Funding organizations are also trying to reduce the burden of evaluating research, e.g. by increasing the time of funding from 3 to 5 years, reducing the number of papers that can be listed in grant applications (German Research Foundation says that numbers aren't everything), or funding investigators and not projects.

Science metrics are not only important for evaluating scientific output, they are also great discovery tools, and this may indeed be their more important use. Traditional ways of discovering science (e.g. keyword searches in bibliographic databases) are increasingly superseded by non-traditional approaches that use social networking tools for awareness, evaluations and popularity measurements of research findings.

References

Lane, J. (2010). Let's make science metrics more scientific *Nature*, *464* (7288), 488-489 <https://doi.org/10.1038/464488a>