

## Editorial

# The measurement of scientific production: Myths and Complexities

---

Generally speaking, the measurement of products, researchers, or groups, is a complex process that may give rise to tensions, especially when considering that these measurements have economic and institutional implications.

The first step in the measurement process, in the case of scientific output, is to list the type of products that can account for academic activity and that may be subject to measurement indicators. These include books, book chapters, scientific papers, divulgation papers, psychosocial intervention materials, patents, models, guidelines, software, social intervention artifacts, social knowledge appropriation actions, participation in academic and social events with research products, actions of public policy transformation, other social innovation products and the processes linked to the production of research.

The next step is to assess those products, which is a very complex matter. For example, clear evaluation criteria are available for journal papers, since they need to undergo peer review and journals transform quality into visibility by taking advantage of the recognition that researchers give to them in terms of citations. A journal with a high citation count can provide an indicator of its quality and information systems can also provide indications of these quality assessment dynamics. This information is not only used to assess products but has also been used in studies about usage, techniques, video traffic or impact of certain contents in a community. (Haran & Poliakoff, 2011; Sugimoto et al., 2013; Thelwall, Haustein, Larivière, & Sugimoto, 2013).

On the other hand, and with a greater degree of complexity, assessing books and book chapters do not have the same systems of evaluation and edition, which decreases their perceived quality. Nevertheless, publishing houses have realised that transparent and demanding assessment processes are needed. This is why both assessment systems are not assigned the same value.

Giving value to other forms of production, such as presenting in academic events or in general mass media, may be a little more difficult, because not all academic events have peer reviewing processes and mass media do not necessarily choose what contents to publish based on the quality of the research but on their own dynamics, which are completely different from those found in academia. However, these activities could be assessed by using download, diffusion and citation indicators both in academic and social settings (Thelwall et al., 2013). With the associated growth in this field, new websites that attempt to assess the quality of the contents created outside traditional written production settings, such as blogs, have been created (Zivkovic, 2011).

Another type of product that can be easily evaluated is patents, since their assessment has a peer reviewing component. Social interventions or innovations expressed in laws or public policy documents, however, are more difficult to assess, despite the important potential recognition they

should have due to their impact on society. Political dynamics do not necessarily take into account the research value of these contents despite their social incidence, but also their political or economical consequences. In terms of assessment, reports on public policy or laws based on research findings should have a more significant weight. Even more difficult would be to measure the impact of research findings on social dynamics within communities, since it is not easy to assess their quality and impact. The challenge clearly is to find those ways of measuring usefulness, impact, and quality.

Educational processes can be assessed by measuring the performance of students in projects and research groups, and associated documents (master's and doctoral theses), both in the short and long term. Some of these products end up being part of books and journal papers, but the difficulty lies in the weight and maturity of these processes in certain contexts: for example, in the Colombian context, doctoral training is still incipient, and a number of research centres are not geared towards training and could not, therefore, account for their own activity in this dimension.

These measurement systems should compare their results within each field of knowledge and taking their own dynamics into account. It is not the same to assess the impact of biomedicine or astronomy or social science, or the humanities. In Colombia, the production in the field of social science recorded by Scopus is strong and growing more than other knowledge areas, and thus cannot be compared against any other field than itself. The tools used to tally and visualise production (Scopus in this case) allow us to observe this growth and counter the arguments of some academics that state otherwise without evidence. Scopus' strategy of covering Latin American journals, and especially of opening its doors to the Social Sciences and the Humanities, enables us to monitor their citation dynamics in an increasingly reliable way.

A final step would be to revise the processes of weighting products, the assessment windows and the limitations of recording systems, because failures in these systems and in the weighting process may contribute to reduce the perceived quality of the assessment system. In Colombia, these problems have created a lack of trust in the system, and adjustment and improvement strategies should be implemented until the whole process is robust and reliable. Another element is to compare these models of measuring the performance of countries to identify strengths and weaknesses, along with the impacts on academic output.

What is ultimately self-evident is that we cannot seem to escape assessment processes and that we need to contribute to improve them, to enhance their impact and their quality, and to justify and value their use within academic and social communities.

**WILSON LÓPEZ-LÓPEZ**  
EDITOR

## References

- Haran, B., & Poliakoff, M. (2011). SPORE series winner. The periodic table of videos. *Science* (New York, N.Y.), 332(6033), 1046–7. doi:10.1126/science.1196980
- Sugimoto, C. R., Thelwall, M., Larivière, V., Tsou, A., Mongeon, P., & Macaluso, B. (2013). Scientists popularizing science: characteristics and impact of TED talk presenters. *PloS One*, 8(4), e62403. doi:10.1371/journal.pone.0062403
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS One*, 8(5), e64841. doi:10.1371/journal.pone.0064841
- Zivkovic, B. (2011). What is: ResearchBlogging.org | The Network Central, Scientific American Blog Network. Scientific American blogs. Retrieved April 16, 2014, from <http://blogs.scientificamerican.com/network-central/2011/10/19/what-is-researchblogging-org/>