

# 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 communi-

ty. (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/>

## Editorial

# La medición de la producción intelectual: Retos, Mitos y Complejidades

---

De forma general, la medición de productos, investigadores o grupos es un proceso que puede generar tensiones, más aún si se tiene en cuenta que estas mediciones tienen implicaciones económicas e institucionales.

El proceso de medición en el caso de la producción académica pasa en primer lugar por enlistar el tipo de productos que pueden dar cuenta de la actividad académica y que pueden ser objeto de indicadores de medición. Así se encuentran dentro de estos productos libros, capítulos en libro, artículos científicos, artículos de divulgación, materiales de intervención psicosocial, patentes, modelos, normas, software, artefactos de intervención social, acciones de apropiación social del conocimiento, participación en escenarios académicos y sociales de los productos de investigación, las acciones de transformación de política pública, como otros productos de innovación social y los procesos de formación asociados a la producción de investigación.

Luego, en este proceso de medición es necesario evaluar los productos y esto presenta un alto nivel de complejidad. Por ejemplo, en el caso de los artículos publicados en revistas tienen una evaluación más clara por cuanto los mismos deben pasar por procesos de valoración de pares y además, las revistas visibilizan la calidad mediante el reconocimiento que los investigadores dan a su trabajo expresado en los usos de estos contenidos. En este sentido, una revista con una alta citación de sus contenidos puede generar un indicador de la calidad de los artículos de la misma y los sistemas

de información de citación o de descargas pueden darnos indicadores de esta dinámica de evaluación de la calidad de estos productos. Pero esta información no solo es utilizada para valorar los productos sino que además hay estudios que muestran el uso, validez de técnicas, tráfico de videos o impacto de ciertos contenidos sobre una comunidad (Haran & Poliakoff, 2011; Sugimoto et al., 2013; Thelwall, Haustein, Larivière, & Sugimoto, 2013)

Por otro lado y con un grado de complejidad mayor, la evaluación de los libros y capítulos en libro no pasan por los mismos sistemas de valoración y edición previa, lo que añade una distancia en términos de calidad percibida. Sin embargo, las editoriales se han dado cuenta hoy por hoy de la necesidad de contar con procesos transparentes y exigentes) de valoración por pares de los contenidos. Es por esta razón que en gran parte de los sistemas de valoración de productos que los artículos y los libros no tienen la misma asignación de valor.

La valoración de otras formas de producción como la presentación en eventos académicos o la presentación en medios de difusión puede resultar un poco más difusa, pues no todos los eventos académicos hacen evaluación por pares y los medios de comunicación no necesariamente deciden sus contenidos en relación con la calidad de la investigación si no por las dinámicas mediáticas que se escapan de la dinámica de la valoración académica. No obstante, estas actividades podrían tener una valoración en relación con la difusión y las descargas en la red a este tipo

de contenidos y en las citaciones que se hacen de ese contenido no solo en el ámbito académicos, si no social (Thelwall et al., 2013). Además, dado que ha sido un campo creciente, debido a la facilidad en el acceso a la información, se han creado páginas que intentan evaluar los contenidos generados por fuera de sistemas tradicionales de producción textos escritos, como por ejemplo los blogs (Zivkovic, 2011).

Por otro lado, otros productos que pueden ser fácilmente valorables son las patentes; por cuanto pasan por procesos de evaluación y reconocimiento de pares. Por el contrario, es un poco más complejo valorar las intervenciones o las innovaciones sociales como las que se expresan en leyes o en documentos de política pública, ya que deberían tener un reconocimiento importante por el impacto que generan en la sociedad. No obstante, las dinámicas políticas no necesariamente toman en cuenta estos contenidos de incidencia social solo por la fuerza de los hallazgos académicos, sino por las consecuencias políticas o económicas de las mismas. En este sentido, informes de política pública o leyes que se logren con base en hallazgos de investigación deberían tener un peso significativo en la evaluación. Por otro lado están productos un poco más complejos de medir como los que están asociados a incorporar resultados de investigación a cambios en las dinámicas sociales en comunidades. Éstos son más difíciles de medir y valorar ya que no es posible dar cuenta de la calidad de los mismos y de sus impactos. Por lo anterior, es claro que estamos ante el reto de encontrar formas de medir su utilidad, impacto y calidad.

Los procesos de formación pueden ser evaluados midiendo a los estudiantes en los proyectos y grupos de investigación y en el caso de procesos de formación en doctorado seguramente las tesis de maestría y doctorado pueden dar indicadores, tanto a corto como a largo plazo. Por ejemplo, algunos de estos tipos de productos terminan en libros y artículos en revistas generando más indicadores menos inmediatos. Pero, es necesario pensar el tema por cuanto la formación doctoral en contextos como el colombiano es incipiente y pueden existir centros de investigación que no están orientados a la

formación y no podrían dar cuenta de su actividad en esta dimensión.

También, es necesario tener en cuenta que estos sistemas de medida deben compararse contra las dinámicas propias de cada una de las áreas de conocimiento. En este sentido, no es lo mismo evaluar los impactos de la biomedicina, que de la astronomía o de las ciencias sociales que de las humanidades. En el caso colombiano la producción de artículos en ciencias sociales registrados en Scopus es grande y está en crecimiento con relación a otras áreas del conocimiento, por lo tanto, no puede compararse si no contra sí misma. Así, las herramientas de contabilizar y visualizar la producción (en este caso Scopus) permiten ver este crecimiento contrario a lo que algunos académicos sin evidencia pueden llegar a afirmar. La estrategia de Scopus de inclusión de revistas de América Latina en forma exponencial y en especial de abrirle la puerta a las revistas de ciencias sociales y humanidades permite dar cuenta de las dinámicas de citación de estos productos en forma cada vez más fiable.

Por último y como una forma de recoger en muchos casos la discrepancia ante los sistemas de medición, es necesario que se evalúen los sistemas de ponderación de los productos y los pesos de los mismos. Al igual que las ventanas de evaluación temporal y de las limitaciones de los sistemas de registro, pues las fallas en estos sistemas de registro y ponderación pueden deslegitimar el sistema de valoración. En el caso colombiano, estas fallas, por lo general, han generado desconfianza y el sistema debería definir estrategias de ajuste y de mejoramiento en el proceso hasta que el mismo sea suficientemente robusto y confiable. Otro elemento adicional debería ser el comparar estos modelos de medición entre países (mediciones comparables a nivel internacional) para identificar las debilidades y las fortalezas como los impactos que ellos tienen sobre las dinámicas de producción académica.

Con todo lo anterior, lo que sí parece evidente es que no podemos escapar de los procesos de evaluación y que debemos contribuir a mejorar los mismos a estudiar sus impactos y a asegurar la calidad de los mismos, justificando y valorando su uso al interior de la comunidad académica y social.

## Referencias

- 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-research-blogging-org/>

