

# La difícil tarea de evaluar la actividad científica

Una de las labores cruciales del quehacer científico universal, incluido el de nuestro país, es evaluar la investigación que produce y la que proyecta realizar, así como las personas dedicadas a esas tareas y las organizaciones en que tiene lugar. Como consecuencia, una fracción no despreciable del tiempo de todo científico activo está dedicada a evaluar proyectos y resultados de investigación, informes sobre el desempeño de personas, solicitudes de ingreso en las entidades científicas y de promoción a posiciones académicas superiores, funcionamiento de instituciones, etcétera.

La evaluación es una instancia imprescindible para garantizar la salud del sistema científico. Asegura que lo que se acepta como conocimiento nuevo ha sido revisado con rigor, y que el dinero disponible para investigar se reparte entre los mejores proyectos. Al mismo tiempo, el trabajo de evaluar presenta desafíos no menores, que, a su vez, se someten recurrentemente a debate y revisión entre los propios miembros de la comunidad científica.

Entre sus aspectos más problemáticos, en todos los países, se destaca la enorme cantidad de información que debe ser analizada, en relación con el número de científicos que la pueden revisar. Y en comunidades relativamente pequeñas, como la argentina, la disponibilidad de expertos que puedan evaluar cabalmente y con independencia la labor de sus pares es notoriamente baja en muchas áreas temáticas.

En un escenario ideal, un investigador debería poder ser evaluado por otro lo suficientemente idóneo, y a la vez lo suficientemente alejado, como para garantizar una razonable objetividad. Además, el evaluador debería contar con el tiempo necesario para examinar de manera exhaustiva los resultados de la labor del evaluado, y así emitir un juicio de valor fidedigno y cabal sobre su desempeño.

La realidad muestra que en todos lados se está muy lejos de esa situación ideal. La escasez de evaluadores y del tiempo que estos pueden dedicar a la tarea, en relación con la elevada cantidad de evaluaciones que deben realizar, ha promovido el empleo de una serie de indicadores cuantitativos sistematizados, que facilitan enormemente la actividad. Pero esos indicadores -por ejemplo, el número de artículos publicados en revistas científicas reconocidas de circulación internacional (en muchos casos calificados por un factor que toma en cuenta la importancia de la

revista)- presentan no pocas limitaciones, que comentaremos en este editorial y que en estos momentos se debaten en los medios académicos de todos los países.

En nuestros comentarios simplificaremos una situación que tiene muchas complejidades. Así, no nos detendremos en distinguir las diferencias entre evaluar investigadores, planes de investigación, publicaciones e instituciones, pero advertimos que nuestras afirmaciones pueden aplicarse de forma distinta en cada uno de esos casos. Tampoco nos detendremos en las diferencias entre disciplinas, en especial, las particularidades de aquellas -por lo general las humanidades y las ciencias sociales- en que la producción intelectual tiende a darse a conocer más publicando libros (o capítulos de ellos) que artículos en revistas.

Para cumplir su función, los evaluadores deben considerar distintas facetas de la actividad de sus colegas, las que por lo general incluyen la producción de conocimiento, la formación de investigadores (es decir, la dirección de tesis de doctorado), la gestión de instituciones académicas, la transferencia tecnológica, la docencia universitaria general y, crecientemente, la divulgación de la ciencia al público no especializado. Si bien la forma en que se pondera cada una de esas facetas da lugar a un debate específico y varía entre disciplinas e instituciones, existe bastante acuerdo acerca de que la producción de conocimiento es la dimensión de mayor peso relativo.

Esa creación de conocimiento, en una primera aproximación, puede determinarse en las ciencias exactas y naturales a partir de la variable mencionada: la cantidad de trabajos de investigación que un científico publicó en revistas científicas, las cuales, a su vez, en su amplia mayoría, practican la evaluación por pares del material que reciben para incluir en sus páginas, lo que permite suponer que, para los especialistas que lo examinaron a pedido de la revista, es conocimiento de aceptable calidad. Pero se trata solo de una primera aproximación, porque lo que realmente debe evaluarse en mayor profundidad es, justamente, la calidad de cada trabajo, más allá de que supere el umbral mínimo que requiere la revista en que se publicó: su aporte al stock universal de conocimiento, su relevancia, su originalidad, su rigor metodológico, entre otras facetas.

Sea como fuere, para un evaluador que se enfrenta con una gran cantidad de individuos por evaluar, simplemente contar la cantidad de artículos publicados por año por cada uno resulta un atajo muy tentador, frente al esfuerzo

de leer concienzudamente cada trabajo. En este contexto, el uso de indicadores cuantitativos resulta valioso y, por eso, se han creado algunos bastante más refinados que la mera suma del número de publicaciones, aunque ellos tampoco están exentos de limitaciones.

Algunos de esos indicadores refinados, que resultan de ingeniosos algoritmos, procuran entre otras cosas reflejar la importancia que asigna la propia comunidad científica a una contribución. Entre tales indicadores están los llamados *índices* o *factores de impacto*, de los que existen numerosas variantes. El mecanismo que emplean es identificar la cantidad de veces que el conocimiento generado por un científico fue utilizado por otros científicos, lo que se determina por las veces que un artículo es citado en otros artículos en revistas semejantes. El supuesto central es que los trabajos más valiosos por su contribución al conocimiento son más citados que los que contribuyen menos, lo que normalmente es cierto (con la rara excepción de artículos con numerosas citas que los refutan).

En la actualidad existen bases de datos continuamente actualizadas que consignan el número de veces que determinado trabajo es citado por otros investigadores, lo que, si se acepta este razonamiento, permite determinar de manera casi instantánea el peso de un investigador a partir de esa clase de repercusión de sus trabajos.


Del mismo modo, las revistas en las que se publican los resultados de la labor científica también son evaluadas por las veces que se citan los trabajos que incluyen en sus páginas. Esto se hace calculando el promedio de citas por artículo. Con esto, las revistas resultan ordenadas en un ranking relativamente sencillo y, como consecuencia, los investigadores tienden a enviar sus trabajos a publicar a las revistas cuyos artículos reciben más citas, es decir, tienen mayor factor de impacto. Una consecuencia de este ordenamiento de las revistas es que, en la práctica, se tiende a juzgar el valor de los trabajos que publica un investigador en función del factor de impacto de la revista en que salió, como si todos los artículos de la revista recibiesen la misma cantidad de citas.

Estas herramientas cuantitativas de evaluación se han aplicado masivamente en todo el mundo desde hace por lo menos veinte años. Han proporcionado un buen marco para analizar comparativamente el valor de la producción científica, y una base numérica objetiva para tomar decisiones de asignación de becas y subsidios, y para contratar o promover investigadores. Pero disponer de un buen marco o una buena base para la evaluación no es lo mismo que tener una herramienta infalible que se puede usar en todos los casos y no requiere mucho más análisis. Por desgracia, no pocas veces se ha procedido como si se tratase de esto último, lo que ha creado múltiples trastornos, más allá de arrojar evaluaciones distorsionadas.

Uno de esos trastornos ha sido estimular a los científicos a que se ocupen de los temas de moda o de aquellos en los que trabaja la mayor cantidad de investigadores, lo

mismo que a las revistas a favorecer la publicación de artículos sobre esos temas, pues tenderán a ser citados más veces y lograrán más altos índices de impacto. Esta tendencia no necesariamente conduce a buena investigación ni a innovaciones, y puede estimular a que se publique mucho trabajo de rutina prolijamente citado en nuevos trabajos de rutina.

Otro trastorno ha sido incitar a los investigadores a que orienten su labor a lograr muchas publicaciones con muchas citas, y menos a que se concentren en reflexionar sobre el fondo de los asuntos que investigan y en profundizar sus respuestas a las preguntas que se plantean. Esto lleva a que salgan artículos que hubiesen requerido más maduración, o a que para publicarlos se fraccione un tema en varias partes, que dan lugar a sendos artículos (y su secuencia de citas), en detrimento de la lógica y de la claridad de exposición.

CIENCIA HOY no tiene un remedio infalible para estos trastornos, ni cree que exista. Como en casi todas las cosas, el camino seguramente pasa por el uso simultáneo de varias herramientas, la moderación en la confianza que se deposita en cada una y, sobre todo, en el sentido común. Tampoco escribí este editorial para tomar parte en el intenso debate académico actual sobre el asunto, porque no sería este el órgano para publicarlo (para los interesados en él, se agregan lecturas sugeridas). Lo hizo para ilustrar a lectores ajenos a los medios científicos pero interesados en ellos sobre cuestiones de las que cada tanto suelen oír hablar y les resultan más bien crípticas. Y porque a quienes están sumidos en el debate seguramente les viene bien mirarlo alguna vez con los ojos del lego, ya que es este, en última instancia, el que debe terminar zanjando cuestiones tan esenciales como para qué la sociedad invierte recursos en ciencia (según expusimos ya hace tiempo en un editorial llamado 'Pares e impares', CIENCIA HOY, 6, 34, 1996). 

## LECTURAS SUGERIDAS

**AAVV**, 1996, 'El juicio de los pares', CIENCIA HOY, 6, 33: 7-8.

**AAVV**, 2013, *San Francisco Declaration on Research Assessment*, accesible en la página de la American Society for Cell Biology, [www.ascb.org/SFdeclaration.html](http://www.ascb.org/SFdeclaration.html).

**GARFIELD E**, 2006, 'The history and meaning of the journal impact factor', *Journal of the American Medical Association*, 295, 1: 90-93.

**MISTELIT**, 2013, 'Eliminating the impact of the Impact Factor', *Journal of Cell Biology*, accesible en [www.jcb.org/cgi/doi/10.1083/jcb.201304162](http://www.jcb.org/cgi/doi/10.1083/jcb.201304162).

**OESTERHELD M**, 2013, 'El valor de tener muchas citas', *Ecología austral*, 23: 70-73.

**SEGLEN PO**, 1997, 'Why the impact factor of journals should not be used for evaluating research', *British Medical Journal*, 314, 7079: 498-502.