



LAS CITAS BIBLIOGRÁFICAS EN LA EVALUACIÓN DE LA ACTIVIDAD CIENTÍFICA: SIGNIFICADO, CONSECUENCIAS Y UN MARCO CONCEPTUAL ALTERNATIVO

BIBLIOGRAPHIC CITATIONS IN THE EVALUATION OF THE SCIENTIFIC ACTIVITY: MEANING, CONSEQUENCES AND AN ALTERNATIVE CONCEPTUAL FRAMEWORK

Jorge V. Crisci^{1*}  y Liliana Katinas^{1*} 

“Las medidas eran excelentes en su género, y fueron bien ejecutadas; su defecto residía en que eran inaplicables al caso y al hombre en cuestión”

Edgar Allan Poe, La carta robada (1844)

SUMMARY

1. División Plantas Vasculares, Museo de La Plata, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, Argentina.

*crisci@fcnym.unlp.edu.ar;
katinas@fcnym.unlp.edu.ar

Citar este artículo

CRISCI, J. V. y L. KATINAS. 2020. Las citas bibliográficas en la evaluación de la actividad científica: significado, consecuencias y un marco conceptual alternativo. *Bol. Soc. Argent. Bot.* 55: 327-337.

 DOI: <https://doi.org/10.31055/1851.2372.v55.n3.28723>

Background and aims: In the evaluation of scientific activity, metrics are applied based on the number of bibliographic citations received by articles from a journal or from a scientist. The objectives of this work are to establish the meaning of a bibliographic citation, discuss the consequences of the use of metrics based on bibliographic citations as a synonym of scientific quality including its effect in the field of Botany, and propose an alternative conceptual framework for evaluating the scientific activity.

Results: A series of factors prevent the statistical support of these metrics: size of the potential audience, the variation in publication and citation practices between the different disciplines, the long-tail statistical distribution of citations, the journal and language of publication, and that an observed statistical regularity will tend to collapse once pressure is put on it when used for control purposes. The consequences of using citations to evaluate are: inhibition of creativity, reification of the scientific achievement, assess the product based on the number of its consumers, the journals with the highest number of citations become a power factor, devaluation of topics with local or regional value, science becomes an industrialized activity, metrics replace judgment, and generating bibliographic citations is part of the objectives of a scientific article.

Conclusions: The origin and validity of these metrics are the consequence of a market society, that is, a way of life organized on the basis of market reasoning and morality and where human and social relations are mere consumer relations.

KEY WORDS

Bibliographic citation, conceptual framework, evaluation of science quality, metrics.

RESUMEN

Introducción y objetivos: En la evaluación de la actividad científica se aplican métricas basadas en el número de citas bibliográficas que reciben los artículos de una revista o los de un científico. Los objetivos de este trabajo son establecer el significado de una cita bibliográfica, discutir las consecuencias del uso de las métricas basadas en las citas bibliográficas como sinónimo de calidad científica incluyendo su efecto en el campo de la Botánica, y proponer un marco conceptual alternativo de evaluación de la actividad científica.

Resultados: Ciertos factores impiden sustentar estadísticamente a esas métricas: tamaño de la potencial audiencia, variación de las prácticas de publicación y de citación entre las distintas disciplinas, la distribución estadística “long tail” de las citas, revista e idioma de la publicación y que una regularidad estadística observada tenderá a colapsar una vez que se ponga presión sobre ella al utilizarla con propósitos de control. Las consecuencias del uso de las citas para evaluar son: inhibición de la creatividad, cosificación del logro científico, valorar al producto en función del número de consumidores, las revistas con mayor número de citas se transforman en un factor de poder, desvalorización de los temas con valor local o regional, la ciencia se industrializa, las métricas reemplazan al juicio, y generar citas bibliográficas forma parte de los objetivos de un artículo científico.

Conclusiones: El origen y la vigencia de estas métricas son la consecuencia de una sociedad de mercado es decir, una forma de vida organizada sobre la base del razonamiento y la moral del mercado y donde las relaciones humanas y sociales son meras relaciones de consumo.

PALABRAS CLAVE

Cita bibliográfica, evaluación de la calidad científica, marco conceptual, métricas.

Recibido: 23 Mayo 2020
Aceptado: 2 Agosto 2020
Publicado: 30 Septiembre 2020
Editora: Ana María Gonzalez 

ISSN versión impresa 0373-580X
ISSN versión on-line 1851-2372

INTRODUCCIÓN

La ciencia exhibe una acusada y predominante contemporaneidad. La mayoría de los acontecimientos científicos, correspondientes a todas las épocas, están ocurriendo ahora ante nuestros ojos. Esa contemporaneidad genera una extraordinaria cantidad de productos de la actividad científica. La sociedad que sostiene a la tarea de la ciencia y la ciencia misma necesitan tener, como nunca antes, herramientas de evaluación de esa tarea.

Entre esas herramientas están aquellas que utilizan como su fundamento a las publicaciones y las citas bibliográficas que reciben. Una cita bibliográfica es la mención en un artículo científico de un texto publicado por otro autor al cual se le adjudica una idea, y que se corresponde con una entrada en la literatura citada.

Para esta forma de evaluar, la calidad de una publicación está directamente relacionada con la cantidad de citas bibliográficas que genera. Por ejemplo, se evalúa a las revistas científicas por el número de citas de los artículos que éstas publican, y a los individuos por el número de citas que sus trabajos reciben.

Nuestro tiempo con su extraordinaria acumulación de datos y sus sofisticadas tecnologías de la información, crearon la oportunidad para desarrollar medidas cuantitativas (índices) que midieran calidad a través de las citas bibliográficas. Los índices más utilizados son: el factor de impacto (Garfield, 2006) y sus variaciones (para revistas) y el índice h (Hirsch, 2005) y sus variaciones (para individuos).

El factor de impacto, definido de la manera más simple, es el cociente entre el número de veces que los artículos publicados en una revista determinada en los últimos dos años fueron citados y el número total de artículos publicados por esa revista en ese período. Una de las variaciones que intentan mejorar este cálculo es el SJR (*Scimago Journal Rank*; González-Pereira *et al.*, 2010). Expresado de manera muy simple, el indicador SJR se calcula sobre el número de las citas recibidas por las revistas en un período de tres años, otorgando un peso mayor a las citas procedentes de revistas de alto prestigio (aquellas con altas tasas de citación y baja autocitación).

En relación a los individuos, de manera muy

simple también, un científico o investigador tiene índice h si ha publicado n trabajos con al menos n citas cada uno. Indirectamente, este índice exige también cierta productividad del investigador. Se han propuesto numerosas variaciones al índice h (Bornmann *et al.*, 2011). Por más sofisticados que sean (que lo son en muchos casos) los cálculos de estos índices, ya sea de revistas o de individuos, siempre tienen el mismo basamento o materia prima: las citas bibliográficas.

Las críticas al cálculo y/o uso de estos índices por parte de editores, sociólogos, científicos, estadísticos, matemáticos y sociedades científicas son numerosas, variadas y severas (e.g., Lawrence, 2007; Campbell, 2008; Todd & Ladle, 2008; Adler *et al.*, 2009; Alberts, 2013; Hicks *et al.*, 2015; DORA, 2020). Por otro lado, estos índices solo se enfocan en un único aspecto del trabajo de un científico: las publicaciones, ignorando otras dimensiones de la actividad de la ciencia tales como la formación de recursos humanos, la gestión en ciencia, la vitalidad y sostenibilidad del grupo de trabajo, la divulgación y la extensión.

No menos serias son las consecuencias que la aplicación de estos índices está causando en la cultura del trabajo científico donde las métricas reemplazan al juicio, y la búsqueda de generar citas bibliográficas forma parte de los objetivos de un artículo científico.

En este contexto, los objetivos de este trabajo son: (1) intentar establecer el significado de una cita bibliográfica, (2) discutir brevemente las consecuencias que acarrea el uso de citas bibliográficas como sinónimo de calidad científica, incluyendo su efecto en el campo de la Botánica, y (3) proponer un marco conceptual alternativo de evaluación de la actividad científica.

El marco conceptual propuesto no intenta ser una suerte de formulario de aplicación para evaluar, sino una base de discusión sobre una evaluación sin el uso de los índices. Este marco incluye, los principios básicos de la evaluación, los escenarios de evaluación y los elementos a evaluar con especial atención a la forma de evaluar a las publicaciones sin utilizar métricas.

¿Qué significa una cita?

A pesar de la profusa cantidad de trabajos que rodean al significado de una cita (e.g., Brooks, 1985; MacRoberts & MacRoberts, 1987; Cozzens,

1989; Moed, 2005), sigue vigente la controversia acerca de la motivación que mueve al investigador a citar un trabajo, e indirectamente a otorgarle significado. Numerosos sociólogos han analizado el acto de citar desde diferentes enfoques, incluyendo investigaciones empíricas (ver referencias en Bornmann & Daniel, 2008), que revelaron una multitud desconcertante de motivaciones, funciones y causas de referencias en comunicación científica (Wouters, 1999). La conclusión es la falta de una adecuada y satisfactoria teoría de la cita (Cronin, 1981, 1984; Cozzens, 1985; MacRoberts & MacRoberts, 1986, 1989; Luukkonen, 1990, 1997).

Sin embargo, podríamos asumir que una cita significa:

- 1) Valor conceptual (relación del trabajo y sus conceptos con otros artículos): y allí encontramos impacto, influencia, utilidad y relevancia.
- 2) Reputación (relación del autor con sus colegas): y allí encontramos reconocimiento, visibilidad y productividad.

A partir de estos supuestos (no necesariamente ciertos pues son difíciles de probar, como los estudios empíricos lo demuestran) se puede intentar construir un modelo estadístico que interprete a estas representaciones -de valor conceptual y de reputación- como calidad científica. Sin embargo, las citas están bajo la influencia de una serie de factores que dificultan la generación de un modelo estadístico.

Algunos de estos factores son:

- Tamaño de la potencial audiencia.
- Variación de las prácticas de publicación y de citación entre las distintas disciplinas.
- Paradigmas dominantes en las distintas disciplinas.
- El no distinguir citas negativas, donde el artículo está citado dentro de un contexto crítico en el que se rechaza lo planteado en el texto, de las citas positivas.
- La distribución estadística “*long tail*” de las citas (alto número de ocurrencias al comienzo, seguida por una larga serie de bajas ocurrencias).
- Revista donde se publica.
- Idioma en el que se publica.
- Relación del autor con sus inmediatos “competidores” en el tema de trabajo.
- Trabajos con numerosos coautores (a veces centenares e incluso miles, e.g., Aad *et al.*, 2015) que resultan a futuro en numerosas citas, que no serán consideradas autocitas.

- Ley (económica) de Goodhart (1976): “*Cualquier regularidad estadística observada tenderá a colapsar una vez que se ponga presión sobre ella al utilizarla con propósitos de control*”. Ello significa que cuando una medida se convierte en objetivo, deja de ser una buena medida. También se asocia con esta ley el llamado “efecto cobra” (Siebert, 2001), que se generó en India, cuando ese país estaba bajo control británico. Buscando disminuir la cantidad de serpientes venenosas que andaban sueltas por las calles, el gobierno comenzó a ofrecer dinero a cambio de cada serpiente cobra muerta que se entregue. En un comienzo, la medida pareció ser exitosa: la gente comenzó a matar cobras y a entregarlas al gobierno. Pero luego de un tiempo, muchos se dieron cuenta que era más fácil criar cobras que cazarlas. Florecieron criaderos de cobras cuyo negocio consistía en entregar cobras muertas al gobierno, a cambio de dinero. Cuando el gobierno se dio cuenta de esta estrategia, decidieron eliminar el programa de compensaciones. Fue entonces cuando los criadores de cobras dejaron libres a muchas serpientes y la cantidad de serpientes sueltas en las calles aumentó.

Estos son algunos de los factores que impiden establecer un modelo estadístico que permita medir cuantitativamente la calidad científica a través de las citas, pero ello no significa que las citas carezcan de valor en la ciencia. Son los índices y su uso regulatorio los que desnaturalizan el posible valor de la cita en un trabajo científico.

Consecuencias del uso de las citas para evaluar la calidad de la ciencia

Dos episodios de la historia reciente de la ciencia son evidencias empíricas de las consecuencias de la sinonimia entre calidad y citas, y al mismo tiempo, son la introducción ideal a un listado específico de esas consecuencias.

En el año 2014 la revista *Nature* (Van Noorden *et al.*, 2014) publicó la lista de los 100 trabajos más citados desde 1900 hasta ese momento, utilizando para ello una base de datos con unos 58 millones de trabajos. El trabajo más citado tenía 305.000 citas y es un trabajo sobre técnicas para medir cantidad de proteínas en una solución. La mayoría de los trabajos de la lista tratan sobre técnicas y software, mientras que Albert Einstein (con 6647 citas),

Watson y Crick y su descubrimiento de la estructura del ADN (con 5307 citas), y el descubrimiento de la capa de ozono (con 1871 citas) no figuran ni cercanamente entre los 100 más citados.

El caso de Francis Peyton Rous, es otro ejemplo de la debilidad del uso de las citas como sinónimo de calidad científica. Rous (1910) publicó a principios del siglo XX un artículo sobre el papel de los virus en los tumores. El hallazgo de Rous fue rechazado por los sostenedores de los paradigmas dominantes de esa época y por lo tanto el trabajo no era citado y Rous se vio obligado por muchos años a abandonar esa línea de trabajo. En 1966, 56 años después de la publicación de su trabajo, Rous recibió el Premio Nobel de Fisiología o Medicina por ese artículo y su mirada pionera.

Más allá de la evidencia empírica que estas anécdotas proveen, la política de utilizar las citas para evaluar la calidad de la ciencia tiene serias consecuencias (Crisci, 2008; Crisci & Apodaca, 2017; Crisci *et al.*, 2019) que conviene listar:

- Inhibición de la creatividad, es decir ajustarse a las modas, no tomar riesgos, quedarse dentro de los paradigmas dominantes eludiendo temas que se salen de ellos. La historia de la ciencia muestra que el mayor progreso ocurre cuando el paradigma dominante es desafiado o revocado (Kuhn, 1970).
- Cosificación (reificación) de la calidad científica, es decir convertir algo abstracto (logro científico) en algo concreto y cuantificable.
- Una “lógica de mercado” de la ciencia, donde el número de consumidores del producto determina el valor del mismo.
- Las revistas con mayor número de citas se transforman en un factor de poder que influye en las decisiones sobre políticas científicas, nacionales e internacionales.
- La desvalorización de temas con valor local o regional, ya que las revistas con alto factor de impacto aceptan solo temas que a su criterio son de interés global.
- Industrialización de la ciencia (Marcuse, 1964), es decir una producción de bienes (en este caso artículos científicos) organizada a gran escala a la manera de una fábrica, con la ayuda de la tecnología. Esta práctica también podría encuadrarse dentro de una forma atenuada del fordismo (Gramsci, 1934): producción industrial y estandarizada en serie.

- El efecto San Mateo: “*Porque, al que tiene, le será dado y tendrá más; y al que no tiene, aun lo que tiene le será quitado*”, del Evangelio según San Mateo 25: 29. La aplicación de estos índices genera revistas e individuos “prestigiosos”. En el caso de las revistas “prestigiosas”, existe el deseo generalizado de publicar en ellas. En el caso de los individuos “prestigiosos”, hay una tendencia a elegir temas y enfoques similares a los que ellos utilizan. Este fenómeno genera una retroalimentación que, con el tiempo, aumenta el valor de los índices para las revistas “prestigiosas” y para los individuos “prestigiosos”.
- Estimula la generación de telarañas burocráticas con un exceso de evaluaciones y un culto desmedido a la cultura del *management* (Lawrence, 2007).
- Cambios en la cultura científica: a) los investigadores, las revistas e incluso las editoriales toman decisiones con el objetivo de aumentar el valor de esos índices (MacDonald & Kam, 2007). Por ejemplo, en la elección de temas y métodos por parte de los investigadores y en la selección del tipo de artículo a publicar por parte de las revistas y las editoriales. Estas decisiones no necesariamente mejoran la calidad de los trabajos, sólo aumentan la posibilidad de ser citados. b) La generación de comunidades no-formales de científicos, denominadas *citation cartels* (Fister Jr. *et al.*, 2016), alrededor de ciertos objetos de estudio, enfoques o métodos populares en ese momento que, al citarse reiteradamente entre sus miembros, producen un aumento de los índices, sin que ello signifique necesariamente aumento de la calidad científica de los artículos. c) La necesidad de los investigadores y las revistas de tener una presencia *online*, pues según estudios empíricos (e.g., Finch *et al.*, 2017; Clayton *et al.*, 2018; Lee, 2019), una mayor presencia *online* aumenta la cantidad de citas que un artículo recibirá. Incluso se ha sugerido una suerte de reemplazo de los indicadores que miden citas, por los indicadores alternativos (Priem, 2014), conocidos como “*altmetrics*”. Estos indicadores alternativos, expresan la cantidad y la diseminación de las menciones que una actividad científica recibe *online* (Nature editorial, 2012) en blogs, Wikipedia, Twitter, Mendeley y en toda otra actividad en redes sociales, en bases de datos, documentos públicos, comentarios de expertos, etc.

La botánica y las citas bibliográficas

La Botánica se ha visto desfavorecida, en relación a otras disciplinas biológicas, por el uso de las citas bibliográficas en la evaluación de la actividad científica. Tres factores han contribuido a ese fenómeno: 1) la hegemonía de los estudios moleculares en la Biología; 2) la desvalorización de los temas botánicos locales o regionales; y 3) la degradación o erosión de la Botánica como disciplina científica.

La hegemonía molecular, producto de importantes avances y de una gran popularidad, tiene el efecto no deseado de una devaluación de la Botánica. El resultado es que un nivel de organización (moléculas) supera en la percepción de los evaluadores y editores de revistas, la relevancia científica de una disciplina de múltiples niveles (e.g., moléculas, células, tejidos, órganos, organismos, poblaciones, especies, etc.) como es la Botánica (Crisci, 2006). Los trabajos botánicos de taxónomos y morfólogos, por citar solo dos ejemplos, se han visto afectados por el factor de impacto, dado que las revistas que tienen mayor índice de impacto están fuertemente centradas en los trabajos a nivel molecular.

En segundo lugar, existe una desvalorización de los temas botánicos con valor local o regional, ya que las revistas con alto factor de impacto (con el fin de atraer una mayor audiencia), solo aceptan temas que a su criterio son de interés global. Este fenómeno perjudica a los trabajos botánicos con enfoques locales o regionales, por ejemplo sobre morfología, taxonomía y biogeografía de plantas (un gran número de las tesis doctorales botánicas de universidades argentinas caen en este grupo).

En tercer lugar, existe una erosión de la Botánica como disciplina científica frente a evaluadores y editores. Esta erosión es el resultado de la desaparición de los departamentos y cursos de botánica en las universidades de los países centrales. A ello se agrega un uso peyorativo o rechazo del término “botánica” (Crisci *et al.*, 2019). Maddox (1998) critica a la Biología Celular manifestando que por el momento los descubrimientos de nuevos genes, y especialmente todo el proyecto del genoma humano, no son más que taxonomía, un inventario, al que califica peyorativamente de esta manera: “*Much of contemporary cell biology is but high level botanizing*”. Un ejemplo del rechazo al uso del término “botánica” es la Declaración

de Shenzhen sobre Ciencias de las Plantas (The Shenzhen Declaration Drafting Committee, 2017) del Congreso Internacional de Botánica del 2017. En la declaración, no puede hallarse en ninguna parte del texto la palabra “*Botany*”, siendo ésta reemplazada por “*Plant Sciences*”.

Finalmente, una comparación del factor de impacto entre revistas (tomado del sitio web de cada revista en agosto de 2020) con énfasis en temas moleculares y/o globales *versus* revistas que publican trabajos botánicos en niveles no moleculares y/o temas locales o regionales, es ilustrativa de la devaluación que el uso de las citas bibliográficas como método de evaluación le causa a la Botánica. En el primer caso podemos mencionar a: *Journal of Molecular Biology* = 5,067; *Molecular Biology and Evolution* = 11,062; *Molecular Phylogenetics and Evolution* = 3,992. En el segundo caso: *Boletín de la Sociedad Argentina de Botánica* = 0,29; *Phytotaxa* = 1,185; *Systematic Botany* = 1,897.

La demostración de la falta de un modelo estadístico que sustente a las citas como herramientas de la evaluación de la actividad científica y las consecuencias negativas del uso de los índices, y su efecto sobre la disciplina Botánica, nos pone frente a la necesidad de ofrecer al mismo tiempo un método alternativo de evaluación. Por ello, a continuación se presenta un marco conceptual que no utiliza a las citas en ninguna etapa de la evaluación.

Un marco conceptual para la evaluación de la actividad científica

El marco conceptual propuesto contiene tres facetas de la evaluación: los principios básicos, los escenarios posibles y los elementos (ítems) a considerar. Los principios, los escenarios y la evaluación de uno de los elementos, las publicaciones, se aplican a la mayoría de las disciplinas científicas. Para el resto de los elementos a evaluar, es imposible generar un marco conceptual único que abarque a todas las disciplinas, ya que cada una tiene sus peculiaridades. Por ello, listamos y brevemente discutimos un conjunto de elementos que no necesariamente aplican a todas las disciplinas, pero que ofrecen una suerte de mapa de ruta para el diseño de evaluaciones específicas de un estilo de labor científica.

La originalidad del marco propuesto nos interesa menos que su utilidad para evaluar y su potencialidad para generar una discusión constructiva sobre el tema.

Los principios básicos del sistema de evaluación

Para realizar la evaluación:

- 1) Se llevará a cabo por pares, es decir escrutinio de expertos en el mismo campo disciplinar y/o metodológico que el investigador evaluado,
- 2) Se considerarán y revisarán (leerán) trabajos publicados por el investigador solo en revistas con arbitraje (excepcionalmente y en casos debidamente justificados, se aceptará evaluar trabajos publicados en revistas sin arbitraje). Podría argumentarse, que el juicio de los pares (arbitraje) de la revista ya es suficiente para establecer la calidad los trabajos. De aceptar como única evaluación los criterios aplicados por las revistas (con sus consiguientes políticas editoriales, “convicciones” de los editores y revisores sobre temáticas y metodologías y la necesidad de aumentar el factor de impacto de la revista), éstas pasarían a ocupar el lugar de la institución que está evaluando a sus miembros o candidatos a serlo. Indirectamente, las revistas estarían estableciendo las políticas científicas de la institución.
- 3) Además de las publicaciones, se tendrán en cuenta todas las otras dimensiones del quehacer de la ciencia (teniendo en cuenta las peculiaridades de cada disciplina), tales como la vitalidad y sostenibilidad del grupo de trabajo, la formación de recursos humanos, repercusión en la comunidad nacional y en la internacional, la divulgación y la extensión y la gestión en ciencia.
- 4) Se asegurará que todas las formas del quehacer científico y todas las disciplinas sean consideradas con la misma imparcialidad de juicio en la evaluación.
- 5) Se hará teniendo en cuenta a la misión de la institución que alberga al investigador y dentro del contexto socio-económico y cultural que rodea al investigador y a su institución.
- 6) No existe un modelo de evaluación único que se aplique a todas las disciplinas y a todas las circunstancias que rodean a la tarea del investigador.

Merece destacarse que el sistema de evaluación por pares es el principal mecanismo de evaluación de los artículos por las revistas y de la labor de los individuos por las instituciones. Estudios empíricos y/o teóricos (e.g., Cole *et al.*, 1981; Horrobin, 1990; Moxham & Anderson, 1992; Weingart, 2005), han demostrado algunas de las debilidades de este sistema, como por ejemplo:

- cierto grado de inevitable subjetividad por parte de los evaluadores. La subjetividad en este caso no incluye a las arbitrariedades, que son un tema ético y que deben estar fuera de cualquier tipo o forma de evaluar,
- frecuentes y profundas divergencias de opinión entre los evaluadores,
- falta de consistencia en el corto plazo de un mismo evaluador en sus criterios de evaluación.

Sin embargo, y a pesar de sus debilidades, la evaluación por pares sigue siendo el mejor sistema posible.

Escenarios de evaluación

Los escenarios más comunes de evaluación son:

- 1) Informe de las tareas realizadas en un periodo determinado.
- 2) Ascenso en un cargo. Se evalúan las tareas realizadas en la categoría de investigación vigente y, secundariamente, el historial científico completo del investigador.
- 3) Ingreso a un cargo o beca. Esto exige evaluar al investigador en un contexto de competencia, o sea generar por los evaluadores un ranking de aquellos que cumplen con las exigencias del cargo.
- 4) Solicitud de subsidio. En este caso aparecen otras dimensiones a evaluar como el plan de trabajo, el equipamiento y el soporte institucional, que no tienen que ver directamente con la calidad científica del solicitante. O sea que la evaluación de la calidad científica es solo una parte de la evaluación total. También incluye, en la mayoría de los casos, al final de la evaluación la necesidad de un ranking.
- 5) Otorgamiento de reconocimientos o premios. Esto exige evaluar el historial científico del investigador, generalmente en un contexto de competencia, o sea la necesidad de generar por los evaluadores de un ranking de aquellos que cumplen con las exigencias del reconocimiento o premio.

Elementos a evaluar

En este marco se utilizan los mismos elementos que son la base de la mayoría de los sistemas de evaluación vigentes. Las diferencias están en: (1) el rechazo al uso de los índices basados en citas bibliográficas. (2) El reemplazo de los índices por la lectura y análisis (originalidad, importancia, rigor) de las publicaciones. (3) Una sugerencia hacia una mirada más totalizadora de la labor científica, que incluye elementos hasta ahora desvalorizados en la mayoría de las evaluaciones tales como la divulgación y la extensión.

Queda en manos de las instituciones evaluadoras la valoración diferencial de cada uno de los elementos y en qué medida se exigirá cada uno de ellos en función de la edad, el nivel y la categoría en la que se encuentra (o a la que aspira) el investigador y la disciplina que cultiva. Esa valoración diferencial estará expresada en el porcentaje de importancia que la institución le da a cada elemento sobre el total de todos los elementos y, en caso de ser necesaria, establecer una gradación de su cumplimiento.

En este marco conceptual y a manera de sugerencia se da una valoración en porcentaje a cada uno de los elementos presentados. Los elementos pueden tener sub-ítems que en cada caso se listan (la lista no es necesariamente exhaustiva). Para cubrir el total del porcentaje adjudicado a ese elemento, no es necesario que el investigador cumpla con todos los sub-ítems, sino haberlo hecho en forma satisfactoria en más de uno de ellos.

1) Publicaciones (valoración del 45 %)

No se utilizará en la evaluación de los artículos tablas de jerarquía de revistas y ningún trabajo será privilegiado o desfavorecido sobre la base del medio donde fue publicado. Solo se exigirá que la revista donde fue publicado el artículo tenga un sistema de evaluación por pares. Para la evaluación lo importante es el contenido del artículo, por ello nada sustituye a su lectura por el evaluador. En el caso de evaluar el historial científico completo de un investigador con numerosos trabajos publicados, se puede acotar el análisis y lectura, por ejemplo, a 10 trabajos seleccionados por el evaluador.

En trabajos con más de dos autores se considerará favorablemente la primera posición o la última (de acuerdo con la disciplina) en la lista de autores. Pero, de ninguna manera se desvalorizará o ignorará la labor de los autores que no ocupen esos lugares,

pues ello iría en contra del trabajo en equipo. Los procesos de evaluación, deben promover ambientes de colaboración y no de competencia por el lugar en la lista de autores.

Los contenidos de los trabajos serán evaluados utilizando tres criterios: originalidad, importancia y rigor.

Por originalidad se entenderá la medida en que el artículo realiza una contribución innovadora para la comprensión y el conocimiento de la disciplina que lo contiene. Los artículos que demuestran originalidad pueden hacerlo de las siguientes maneras: producir e interpretar nuevos hallazgos empíricos o nuevo material; comprometerse con problemas nuevos y / o complejos; desarrollar métodos de investigación innovadores, metodologías y técnicas analíticas; mostrar un alcance imaginativo y creativo; proporcionar nuevos argumentos y / o nuevas formas de expresión, innovaciones formales, interpretaciones y / o ideas; y recopilar y participar con nuevos tipos de datos.

La importancia está expresada en la capacidad del trabajo en producir un avance de la disciplina conocimiento, habilidades, comprensión y erudición en la teoría, práctica, educación, gestión y / o política de la disciplina. También la importancia está en la aplicabilidad de los resultados del artículo a políticas en temas tales como: biodiversidad, enfermedades emergentes, pobreza y exclusión social, cambio climático, economía mundial, alimentos, energía, agua, globalización, avances tecnológicos y sociedades del conocimiento.

Por rigor se entenderá la medida en que el artículo demuestra coherencia e integridad intelectual, y adopta conceptos, análisis, fuentes, teorías y / o metodologías robustas y apropiadas. Se analizará además, el rigor científico y excelencia, con respecto al diseño, método, ejecución y análisis, coherencia lógica del argumento y a la escala, el desafío y la dificultad logística que plantea la investigación.

Respecto a la productividad, se tendrá en cuenta el número de trabajos publicados por el investigador, siempre y cuando los trabajos reflejen coherencia interna y no sean rehechos, recompuestos o refundidos de otros trabajos, o subdivisiones artificiales de un trabajo mayor.

En ciertas disciplinas, es fundamental la transferencia de los desarrollos científico-tecnológicos a la sociedad o a los sectores socio-productivos,

como por ejemplo una patente. Estas actividades de transferencia se equiparan con publicaciones.

2) Formación de recursos humanos (valoración del 15 %)

Se evaluará la capacidad del investigador para transmitir conocimientos y experiencias directamente relacionadas con su actividad, así como para dirigir y formar personal de investigación. Esto se comprobará a través de: dirección de tesis de grado y posgrado, becarios de grado y de posgrado, y pasantes; destino científico y laboral de los discípulos generados por las actividad de la dirección de tesis de posgrado; docencia universitaria de grado y posgrado; publicación de libros de texto, técnicos o metodológicos.

3) Repercusión nacional e internacional de la labor del investigador (valoración del 15 %)

Las palabras “nacional” e “internacional” se refieren a estándares de calidad, no a un alcance geográfico de temas particulares, ni al lugar donde se realizan las investigaciones.

Este elemento se comprobará mediante las siguientes actividades divididas en nacionales e internacionales: (a) cursos, seminarios y conferencias dictados, participación en simposios, mesas redondas. En todos los casos tener en cuenta la institución organizadora de las actividades y, en el caso de conferencias en congresos, el tipo de conferencia (inaugural, plenaria, simultánea); (b) profesor invitado, universidad que invita; (c) becas obtenidas; (d) subsidios obtenidos; (e) participación en proyectos de investigación inter-institucionales; (f) participación en comisiones, instituciones que las cobijan; (g) editor o miembro comité editorial de revistas científicas, para ello se considerará la revista, la entidad o empresa a la que pertenece y el tipo de tarea editorial realizada; (h) premios y reconocimientos, se tendrá en cuenta el tipo de premio o reconocimiento y la entidad otorgante.

4) Grupo de trabajo al que pertenece el investigador (valoración del 10 %)

Se esperará del grupo vitalidad y sustentabilidad. La vitalidad se entenderá como la medida en que el grupo se apoye en una cultura de investigación próspera e inclusiva para todo el personal, incluyendo los doctorandos y en la capacidad de atraer investigadores de posgrado y postdoctorado.

La sustentabilidad se entenderá como la medida en que el entorno de investigación garantiza la salud futura, la diversidad, el bienestar, la infraestructura y la contribución científica del grupo. En la sustentabilidad, es fundamental la capacidad del grupo para obtener subsidios a la investigación.

5) Divulgación y/o extensión (valoración del 10 %)

Como parte de las responsabilidades sociales, los científicos deben desarrollar medios de comunicar la ciencia al público (publicaciones, libros, videos, material *online*, etc.) y apoyar la apertura de sus instituciones para que esto suceda. Los científicos además tienen la responsabilidad social de cooperar con los medios de comunicación en el fortalecimiento de la alfabetización en ciencia de los ciudadanos. Por ello, la divulgación, debe formar parte de la calidad de la actividad científica. Muchas veces se ha expresado este pensamiento, pero muy pocas instituciones lo han aplicado en sus evaluaciones.

Los programas y proyectos de extensión constituyen prácticas institucionales y comunitarias de diverso grado de complejidad que vinculan a las instituciones científicas con el medio social, basándose en el conocimiento científico acumulado en la institución y en su capacidad de volcar en la comunidad aquellos aspectos del quehacer científico que mejoren la calidad de vida de la gente. Su misión es la de consolidar el vínculo entre la institución y la comunidad (niños, jóvenes, adultos, adultos mayores, instituciones, organismos gubernamentales, entidades privadas, organizaciones sociales, etc.).

6) Gestión en la ciencia (valoración del 5 %)

Se tendrá en cuenta la actividad llevada a cabo en la creación, organización y/o desarrollo de centros de investigación o su actuación en organismos de planeamiento, promoción o ejecución científica así como la participación en comisiones o en actuación como pares evaluadores en revistas o en instituciones, participación activa en sociedades científicas como miembro de comisiones directivas.

CONCLUSIONES

La importancia y la actualidad que el tema de la evaluación de las actividades de la ciencia tiene, están reflejadas en la exorbitante cantidad de

publicaciones que en los últimos 20 años se han ocupado del tema. Cuatro factores, aparentemente no relacionados entre sí, son los causantes de este fenómeno: la gran cantidad de productos que la ciencia actual genera, la necesidad que la sociedad y la ciencia misma tienen de evaluar esos productos, la gran acumulación de datos y las modernas tecnologías de la información para analizar esos datos.

En este contexto los datos han pasado a gobernar la evaluación de las actividades científicas y florecen las métricas que los analizan. Esas métricas están basadas en el postulado que las citas bibliográficas son el reflejo de la calidad de una publicación, y se generan índices para evaluar revistas e índices para evaluar individuos.

A pesar que el significado de una cita no se puede establecer con certeza y que la aplicación de los índices genera serias consecuencias, estas métricas se siguen utilizando como método de evaluación. La Botánica, más allá del nivel molecular, está desfavorecida con respecto a otras disciplinas por estas métricas.

A menudo, se utilizan estos índices reconociendo sus deficiencias, pero con la contradictoria justificación de que son fáciles y rápidos de aplicar, frente a la gran cantidad de evaluaciones que las instituciones tienen que realizar.

El origen y la vigencia de estas evaluaciones métricas son la consecuencia de una sociedad de mercado. Una economía de mercado es simplemente un mecanismo para la asignación eficiente de recursos, pero una sociedad de mercado es una forma de vida organizada sobre la base del razonamiento y la moral del mercado. La reificación, la transformación de las relaciones humanas y sociales en meras relaciones de consumo, se ha convertido en una segunda naturaleza en la actual sociedad de mercado. En sus formas extremas, es utilizar a las personas de manera instrumental (como medir la inteligencia a través de pruebas de coeficiente intelectual) al convertir lo abstracto en concreto (Lukács, 1971; Gould, 1981).

Por otro lado, la reificación es una forma de disciplinar en el sentido de Foucault (1976) y una forma de hegemonía cultural en el sentido de Gramsci (1952). Este proceso disciplinario influye en la percepción que los científicos tienen de sí

mismos y en la percepción que la sociedad tiene de los científicos, y permite calificar, clasificar y recompensar o sancionar. La gratificación o la sanción vienen a través de, por ejemplo, la aceptación de los manuscritos por parte de las revistas, la asignación de empleos y promociones, y la adjudicación de subsidios y premios.

Este proceso disciplinario genera polos de poder, como los que ejercen las editoriales, los editores, las instituciones científicas y los líderes de los paradigmas dominantes. Polos de poder que, por más bienintencionados que sean, tienen sus propios objetivos y estrategias, los cuales no necesariamente están en línea con los objetivos del investigador, de su institución o de su país.

En última instancia, el trabajo científico debe ser visto como una labor creativa, y no en términos de producción y consumo de bienes y servicios, cuyo objetivo final es la búsqueda de citas. Es importante mencionar que las citas no carecen de valor en la ciencia, son los índices y su uso regulatorio los que desnaturalizan el posible valor de la cita de un trabajo científico.

Tal vez para finalizar nada mejor que regresar en el tiempo 2000 años, y escuchar lo que tiene para decirnos el filósofo romano Lucio Séneca (4 a.C. - 65 d.C.): *“La sabiduría reside en saber separar las cosas del ruido que ellas producen”*.

CONTRIBUCIÓN DE LOS AUTORES

JVC diseñó el trabajo y ambos autores discutieron ideas y redactaron el manuscrito.

AGRADECIMIENTOS

Agradecemos la lectura crítica del manuscrito a María José Apodaca, Gabriel Bernardello, Victoria Crisci, Edgardo Ortiz-Jaureguizar y Facundo Palacio, sin que ello signifique responsabilidad alguna con el contenido de este trabajo, y a la directora de la revista y a los revisores. Ambos autores agradecen el apoyo del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT 2017-0965), y Facultad de Ciencias Naturales y Museo, UNLP (11/N 814).

BIBLIOGRAFÍA

- AAD, G., B. ABBOTT, J. ABDALLAH (y 5151 autores más). 2015. Combined measurement of the Higgs Boson mass in *pp* collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments. *Phys. Rev. Lett.* 114: 191803-1–191803-33. <http://doi.org/10.1103/PhysRevLett.114.191803>
- ADLER, R., J. EWING & P. TAYLOR. 2009. Citation statistics. A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). *Stat. Sci.* 24: 1-14. <http://doi.org/10.2307/20697661>
- ALBERTS, B. 2013. Impact factor distortions. *Science* 340: 787. <http://doi.org/10.1126/science.124031>
- BORNEMANN, L. & H.-D. DANIEL. 2008. What do citation counts measure? A review of studies on citing behavior. *J. Doc.* 64: 45-80. <http://doi.org/10.1108/00220410810844150>
- BORNEMANN, L., R. MUTZ, S. E. HUG & H.-D. DANIEL. 2011. A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *J. Informetr.* 5: 346-359. <https://doi.org/10.1016/j.joi.2011.01.006>
- BROOKS, T. A. 1985. Private acts and public objects: An investigation of citer motivations. *J. Am. Soc. Inf. Sci. Tec.* 36: 223-229. <https://doi.org/10.1002/asi.4630360402>
- CAMPBELL, P. 2008. Escape from the impact factor. *ESEP* 8: 5-7. <https://doi.org/10.3354/esepp00078>
- CLAYTON, T., C. T. LAMB, S. L. GILBERT & A. T. FORD. 2018. Tweet success? Scientific communication correlates with increased citations in Ecology and Conservation. *Peer J* 6: e4564. <https://doi.org/10.7717/peerj.4564>
- COLE, S., J. R. COLE & G. A. SIMON. 1981. Chance and consensus in peer review. *Science* 214: 881-886. <https://doi.org/10.1126/science.7302566>
- COZZENS, S. E. 1985. Comparing the sciences - citation context analysis of papers from neuropharmacology and the sociology of science. *Soc. Stud. Sci.* 15: 127-153. <https://doi.org/10.1177/030631285015001005>
- COZZENS, S. E. 1989. What do citations count? The rhetoric-first model. *Scientometrics* 15: 437-447. <https://doi.org/10.1007/BF02017064>
- CRISCI, J. V. 2006. One-dimensional systematist: Perils in a time of steady progress. *Syst. Bot.* 31: 217-221. <https://doi.org/10.1600/036364406775971859>
- CRISCI, J. V. 2008. La barbarie del “especialismo” en un tiempo de extinciones. *Anales Acad. Nac. Agron. Veterin. Buenos Aires* 62: 97-107.
- CRISCI, J. V. & M. J. APODACA. 2017. Los rankings globales de universidades y su función disciplinaria. *Revista Mus. La Plata* 2: 12-18. <https://doi.org/10.24215/25456377e039>
- CRISCI, J. V., M. J. APOCADA & L. KATINAS. 2019. El fin de la Botánica. *Revista Mus. La Plata* 4: 41-50. <https://doi.org/10.24215/25456377e067>
- CRONIN, B. 1981. Agreement and divergence on referencing practice. *J. Inf. Sci.* 3: 27-33. <https://doi.org/10.1177/016555158100300104>
- CRONIN, B. 1984. *The citation process. The role and significance of citations in scientific communication.* Taylor Graham, Oxford, UK.
- DORA (San Francisco Declaration on Research Assessment). 2020. [consultado el 9 de mayo de 2020]. Accesible en: <http://www.ascb.org/dora>
- FINCH, T., N. O'HANLON & S. P. DUDLEY. 2017. Tweeting birds: Online mentions predict future citations in ornithology. *R. Soc. Open Sci.* 4:171371. <https://doi.org/10.1098/rsos.171371>
- FISTER JR., I., I. FISTER & M. PERC. 2016. Toward the discovery of citation cartels in citation networks. *Front. Phys.* 4(49): 1-5. <https://doi.org/10.3389/fphy.2016.00049>
- FOUCAULT, M. 1976. *Vigilar y castigar. Nacimiento de la prisión.* Siglo Veintiuno Editores, Buenos Aires, Argentina.
- GARFIELD, E. 2006. The history and meaning of the journal impact factor. *JAMA* 295: 90-93. <https://doi.org/10.1001/jama.295.1.90>
- GONZÁLEZ-PEREIRA, B., V. P. GUERRERO-BOTE & F. MOYA-ANEGÓN. 2010. A new approach to the metric of journals' scientific prestige: The SJR indicator. *J. Informetr.* 4: 379-391. <https://doi.org/10.1016/j.joi.2010.03.002>
- GOODHART, C. A. E. 1976. Monetary relationships: A view from threadneedle street. In: RESERVE BANK OF AUSTRALIA (ed.), *Papers in Monetary Economics*, vol. 1, Revised version of six papers presented at the Conference in Monetary Economics, Sydney, July 1975, in various pagings. Sidney, Australia.
- GOULD, S. J. 1981. *The mismeasure of man.* W. W. Norton & Co., New York.
- GRAMSCI, A. 1934. *Cuadernos de la cárcel, tomo 6, Cuaderno 22 (V), Americanismo y fordismo.* Traducción al español de la edición crítica del

- Instituto Gramsci a cargo de Valentino Garretana, 2000, Ediciones Era, Puebla, México.
- GRAMSCI, A. 1952. *Il materialismo storico e la filosofia di Benedetto Croce*. Primera edición. Einaudi Editor, Torino, Italia.
- HICKS, D., P. WOUTERS, L. WALTMAN, S. DE RIJCKE & I. RAFOLS. 2015. The Leiden manifesto for research metrics. *Nature* 520: 429-431. <https://doi.org/10.1038/520429a>
- HIRSCH, J. E. 2005. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U. S. A.* 102: 16569-16572. <https://doi.org/10.1073/pnas.0507655102>
- HORROBIN, D. F. 1990. The philosophical basis of peer review and the suppression of innovation. *JAMA* 263: 1438-1441. <https://doi.org/10.1001/jama.1990.03440100162024>
- KUHN, T. S. 1970. *The structure of scientific revolutions*. Enlarged 2nd ed. University of Chicago Press, Chicago.
- LAWRENCE, P. A. 2007. The mismeasurement of science. *Curr. Biol.* 17: R583-R585. <https://doi.org/10.1016/j.cub.2007.06.014>
- LEE, J.-S. M. 2019. How to use Twitter to further your research career. Nature blogs, <https://www.nature.com/articles/d41586-019-00535-w>
- LUKÁCS, G. 1971. *History and class consciousness*. The MIT Press, Cambridge, Massachusetts.
- LUUKKONEN, T. 1990. *Citations in the rhetorical, reward and communications systems of science*. PhD thesis, Acta Universitatis Tamperensis, ser A, vol. 285, University of Tampere, Tampere, Finland.
- LUUKKONEN, T. 1997. Why has Latour's theory of citations been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis. *Scientometrics* 38: 27-37. <https://doi.org/10.1007/BF02461121>
- MACDONALD, S. & J. KAM. 2007. Aardvark *et al.*: Quality journals and gamesmanship in management studies. *J. Inform. Sci.* 33: 702-717. <https://doi.org/10.1177/0165551507077419>
- MACROBERTS, M. H. & B. R. MACROBERTS. 1986. Quantitative measures of communication in science: A study of the formal level. *Soc. Stud. Sci.* 16: 151-172. <https://doi.org/10.1177/030631286016001008>
- MACROBERTS, M. H. & B. R. MACROBERTS. 1987. Measurement in the face of universal uncertainty: A reply to Stigler. *Soc. Stud. Sci.* 17: 335. <https://doi.org/10.1177/030631287017002008>
- MADDOX, J. 1998. *What remains to be discovered*. Free Press, New York.
- MARCUSE, H. 1964. *One-dimensional man: Studies in the ideology of advanced industrial society*. 1st edition. Beacon Press, Boston.
- MOED, H. F. 2005. *Citation analysis in research evaluation*. Springer, Dordrecht.
- MOXHAM, H. & J. ANDERSON. 1992. Peer review. A view from the inside. *Sci. Technol. Policy* 5: 7-15.
- NATURE EDITORIAL. 2012. Alternative metrics. *Nature Materials* 11: 907. <https://doi.org/10.1038/nmat3485>
- PRIEM, J. 2014. Altmetrics. In: CRONIN, B. & C. R. SUGIMOTO (eds.), *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact*, pp. 263-288. 1st edition. The MIT Press, Cambridge, Massachusetts.
- ROUS, P. 1910. A transmissible avian neoplasm (sarcoma of the common fowl). *J. Exp. Med.* 12: 696-705. <https://doi.org/10.1084/jem.12.5.696>
- SIEBERT, H. 2001. *Der Kobra-Effekt. Wie man Irrwege der Wirtschaftspolitik vermeidet*. Deutsche Verlags-Anstalt, Munich.
- THE SHENZHEN DECLARATION DRAFTING COMMITTEE. 2017. The Shenzhen Declaration on Plant Sciences—Uniting plant sciences and society to build a green, sustainable Earth. *PhytoKeys* 86: 3-7. <https://doi.org/10.3897/phytokeys.86.20859>
- TODD, P. A. & R. J. LADLE. 2008. Hidden dangers of a 'citation culture'. *ESEP* 8: 13-16. <https://doi.org/10.3354/ese00091>
- VAN NOORDEN, R., B. MAHER & R. NUZZO. 2014. The top 100 papers. *Nature* 514: 550-553. <https://doi.org/10.1038/514550a>
- WEINGART, P. 2005. Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics* 62: 117-131. <https://doi.org/10.1007/s11192-005-0007-7>
- WOUTERS, P. 1999. *The citation culture*. Doctoral thesis, University of Amsterdam, The Netherlands.

