

Editorial

El futuro de la evaluación de revistas científicas

Yair Alexander Porras Contreras¹

La investigación en las universidades colombianas se posiciona como un eje transversal de los procesos de formación de profesionales, toda vez que rige el rumbo de los planes, los proyectos y las iniciativas involucradas en la producción del conocimiento, su divulgación y apropiación social. Si bien las actividades de docencia significan el mayor porcentaje de los tiempos de dedicación en los planes de trabajo de los profesores, la investigación constituye el aspecto más valorado del trabajo docente (Alperin, Muñoz, Schimanski, Fischman, Niles, y McKiernan, 2019). La importancia de compartir y comunicar los resultados de las investigaciones que se elaboran en las universidades ha propiciado el surgimiento de un sistema de evaluación de las publicaciones, que en principio era solo utilizado por los bibliotecólogos para tomar decisiones sobre el tipo de revistas que se deberían adquirir, convirtiéndose hoy en día en el paradigma de racionalidad institucionalizado sobre el cual se mide la calidad de las revistas.

La revista pionera en la construcción de una nueva visión de mundo centrada en la valoración del conocimiento fue *Nature*, desde 1967, al convertirse en la primera publicación en adoptar la revisión por pares como parte de su política editorial (Norman, 2017). A partir de esa época y hasta hoy, la revisión por pares ha tenido defensores y detractores. Entre las principales críticas a este sistema se encuentra la lentitud del proceso, su ineficiencia, la evaluación sesgada y la posibilidad de abusos. Sobre este asunto, la conferencia SpotOn London 2016 formuló la siguiente pregunta de discusión a un grupo de investigadores, bibliotecarios, editores y científicos: ¿cómo podría ser la revisión por pares en el año 2030? Sin lugar a dudas, una posible respuesta debe trascender el campo de la medición, para considerar otros aspectos propios de los contextos de significación, en los cuales la métrica tradicional aún se considera incipiente.

A nivel mundial, en lo que algunos denominan la geopolítica del conocimiento (Flick y Convoco Foundation, 2011), se resalta la evaluación de las revistas científicas a partir del uso de métricas o marcadores de calidad, que tienen por objeto determinar el número de citas de los artículos que pertenecen a un campo. Dentro de estos indicadores se destaca el factor de impacto (FI), el cual surgió como una alternativa para decantar el número de revistas que, por sus esfuerzos editoriales, merecían un reconocimiento, además de ser un referente para la toma de decisiones para los organismos encargados de asignar los presupuestos. Esta iniciativa fue allanando el camino hacia tres conceptos polémicos,

¹ Profesor del Departamento de Química, Universidad Pedagógica Nacional. Doctor en Innovación e Investigación en Didáctica. Editor de la revista *Tecné, Episteme y Didaxis: TED*.

Correo electrónico: yporras@pedagogica.edu.co
ORCID: <http://orcid.org/0000-0002-7111-0632>

en las discusiones sobre las metodologías de evaluación de las publicaciones: la calidad, la excelencia y el prestigio. Vale la pena recordar que el conteo de referencias para medir el impacto de las publicaciones fue propuesto por Garfield en 1955 (Garfield, 1955), siendo consolidado en 1963 por el denominado *factor de impacto*. Históricamente el factor de impacto (FI), desarrollado por la organización Thomas Reuters, mide la frecuencia promedio con la que se han citado los artículos publicados en una determinada revista en los dos años anteriores. Se calcula con el número de citas en un cierto año, con respecto a los artículos publicados en una revista durante los dos años anteriores, dividido por el número de artículos citables, publicados en la misma revista y período de tiempo (FAZEL & LAMSMA, 2015).

$$FI = \frac{\text{número total de citas que reciben los artículos publicados en una revista en los últimos dos años}}{\text{número total de artículos que se publicaron en esa misma revista en esos dos años}}$$

Vale la pena mencionar que existen otros indicadores de impacto, ideados para complementar la información del FI. Por ejemplo, Google Scholar utiliza el índice H (Hirsh, 2005), que mide la productividad e impacto de la producción científica. Este se define como *el mayor número h, de modo que al menos h artículos en esa publicación fueron citados h veces cada uno en los últimos cinco años calendario completos* (Google, 2019). En otras palabras, un investigador tiene un índice h cuando h de sus documentos recibieron al menos h citas cada uno. Por ejemplo, un investigador o revista con un índice h5 igual a 12, significa que en los cinco años de revisión ha publicado 12 artículos que recibieron 12 citas cada uno.

¿Por qué razón el factor de impacto ha perdido su credibilidad?

Las críticas al FI se centran, en primer lugar, en la sobredimensionada importancia que se les brinda a los datos y no a las reflexiones o a los juicios, en torno a la calidad de un artículo o de una publicación. En este sentido, las evaluaciones centradas en factores de impacto promueven una distorsión en el comportamiento del autor y del editor, particularmente en temas de financiamiento, nombramiento y promoción. De igual manera, los datos se limitan estrictamente al campo científico, por lo que otros campos de conocimiento no son considerados, subestimando la apropiación social de un artículo y de una revista, además de invisibilizar el contexto desde el cual se construyen los conocimientos. Junto a estos factores, resulta preocupante el aumento en el uso de malas prácticas editoriales de muchas revistas, al alentar maniobras de publicación de artículos de alto perfil en los primeros meses del año, con el fin de acumular más citas y con esto mejorar los resultados del FI (Ritter, 2013).

La suma de críticas al FI confluyó en la conformación de algunas organizaciones, cuyo propósito se enfocó en la construcción de propuestas alternativas o complementarias a este tipo de métricas. Tres documentos se constituyen en los referentes de estas propuestas: el primero de ellos es la *Declaración de San Francisco sobre la evaluación de la investigación* (DORA, 2012), con 18 recomendaciones que se basan en tres aspectos: a) eliminar el uso de métricas apoyadas

en revistas, tales como el factor de impacto, para el financiamiento, el nombramiento y/o la promoción; b) la necesidad de evaluar la investigación por sus propios méritos en lugar de basarse en la revista que la publica, y c) la necesidad de aprovechar las oportunidades que ofrece la publicación en línea, para considerar métricas alternativas o *factores de uso* importantes para identificar el impacto de los artículos. Frente a estos temas, DORA aconseja a los editores que reduzcan el énfasis en el factor de impacto como herramienta promocional, alentando el uso de una variedad de métricas basadas en revistas. De igual manera, se recomienda disponer de una variedad de métricas a nivel de artículo para alentar un cambio hacia la evaluación basada en su contenido, fomentando las prácticas de la autoría responsable y reconociendo las contribuciones de cada autor. Especial atención se debe prestar a las citas de fuentes primarias, por lo cual los comités editoriales en sus políticas de publicación deben recomendar a los autores que den los créditos a quienes informaron anticipadamente sobre un hallazgo o el resultado de un estudio.

El segundo documento de referencia, enfocado a mejorar las formas de evaluación de publicaciones, es el *Manifiesto de Leiden* (2015), un decálogo sobre buenas prácticas en la evaluación de revistas, más allá del FI. Entre los puntos relevantes de este documento, asociados a la gestión editorial, se encuentran: el apoyo a la valoración cualitativa; la protección a la excelencia en investigación de relevancia local; la apertura, transparencia y simplicidad en los procesos de recopilación y análisis de datos; la necesidad de considerar las diferencias en las prácticas de publicación y citación entre diversos campos científicos, y el reconocimiento periódico de los efectos sistémicos de la evaluación y los indicadores.

El tercer documento de análisis es el *Manifiesto de Hong Kong para evaluar a los investigadores: promoviendo la integridad de la investigación* (Moher et al., 2019), el cual presenta cinco principios que se enfocan en:

a) prácticas de investigación responsables; b) informes transparentes; c) ciencia abierta (investigación abierta); d) valorar la diversidad de tipos de investigación, y e) reconocer todas las contribuciones a la investigación y a la actividad académica. Queda claro, fruto de estas iniciativas, que el futuro de la evaluación de las publicaciones científicas requiere una transformación estructural, que en el contexto colombiano está siendo liderado por la mesa técnica de Publindex, un espacio de discusión producto de los acuerdos entre el gobierno y los estudiantes, a raíz de las movilizaciones del año 2018. Algunos puntos que son tema de debate y pueden contribuir a las reflexiones frente al futuro de las publicaciones en Colombia van más allá de aceptar el aumento del FI de las publicaciones como el propósito exclusivo de las revistas, considerando que dicho indicador debe ser una consecuencia de los procesos de publicación, divulgación y apropiación social del conocimiento que se produce en las comunidades. De esta misma manera, concentrar la mirada en el uso de indicadores alternativos se constituye en una gran opción para reconocer la visibilidad e impacto de un artículo en las redes sociales (Bar-Ilan et al., 2012).

Una primera definición de las métricas alternativas es la formulada por Priem et al. (2010), en el manifiesto *Altmetrics*, en el que se justifica el uso de indicadores emergentes para resaltar actividades de discusión y debate en las redes sociales o en la web social: "Este grupo diverso de actividades forma un rastro compuesto de impacto mucho más rico que cualquier otro disponible. Llamamos a los elementos de este seguimiento *altmetrics*" (p. 1). En este sentido, las métricas alternativas se constituyen en indicadores, métodos o estrategias para evaluar el impacto de un artículo en ambientes o ecosistemas sociales. Esta propuesta de evaluación involucra redes sociales "tradicionales", *blogs*, servicios de marcadores académicos, gestores de referencias, medios de comunicación, multimedia, canales o *podcasts*, *preprints* y sitios de revisión por pares posteriores a la publicación, entre otros.

De acuerdo con algunos autores (Hodgson, 2014; Williams & Paula, 2015; Alonso-Arévalo, 2016) las ventajas de los indicadores alternativos se centran en:

- Ofrecer mejores maneras de filtrar el impacto e influencia de la investigación, gracias a herramientas centradas en la web, como marcadores, enlaces, *tweets* y *blogs*.
- Evaluar los artículos individuales, en vez de la revista como un todo.
- Contar las descargas en cualquier lugar donde aparezca un artículo en línea, desde repositorios de datos y bibliotecas digitales, hasta el sitio web de un editor o autor.
- Mostrar la atención, la recepción y la respuesta a un trabajo publicado antes de ser citado.
- Se puede aplicar a resultados de investigación no tradicionales como conjuntos de datos y publicaciones de *blog*.
- Mostrar el impacto de la investigación en tiempo real, sin esperar puntajes estandarizados.

Entre las principales desventajas que citan estos autores se encuentran:

- Un alto número de menciones no es necesariamente un buen indicador para reconocer si la investigación es buena o mala.
- Al medir la influencia social, la calidad de una investigación queda en entredicho.
- Falta de consistencia y normalización entre los datos y sitios utilizados para la medición.
- El impacto social no significa un impacto académico, ya que estos pueden variar mucho en una publicación.

Como una posible conclusión, es necesario reconocer las limitaciones del factor de impacto en los procesos de evaluación de revistas; sin embargo, su transición a un sistema de evaluación alternativo aún se aprecia lejano, quizás por la representación social que circula entre autores y editores en torno a la necesidad de mantener un nivel de rigurosidad y seriedad en la evaluación de artículos científicos. En este sentido, vale la pena enfatizar en la crítica a los sistemas de medición que tienen como propósito dar cumplimiento a las políticas de globalización que ponen a las personas al servicio de la economía, como es el caso de las métricas para evaluación de artículos, por solo citar algunos ejemplos. No olvidemos que, de acuerdo con la Ley de Goodhard (1984), "Cuando una medida se convierte en un objetivo, deja de ser una buena medida" y quizás se convierte en la justificación para comportamientos poco éticos. En un interesante artículo de Fire y Guestrin (2019) se ejemplifica la ley de Goodhart en diferentes campos de conocimiento, destacándose su influencia en la evaluación de publicaciones científicas. Para estos autores, al convertir las medidas en los objetivos de la publicación académica, su efectividad puede comprometerse al emerger comportamientos no deseados, tales como las publicaciones salami, las autorías fantasmas, la piratería informática, la manipulación de métricas, la falsificación de datos de investigación, la falsificación de revisiones por pares e

incluso el plagio por parte de revisores (Fire y Guestrin, 2019, p. 2).

Alentamos a nuestros lectores el uso de las redes sociales de la revista, en las cuales sometemos a discusión los artículos construidos por docentes en el campo de la educación en ciencias, las matemáticas y las tecnologías. De igual manera, invitamos a la lectura del presente número, el cual cuenta con catorce artículos, de los cuales diez corresponden a *investigación*, un artículo de *caso educativo*, dos de *reflexión* y un artículo de *revisión de tema*.

Los 10 artículos de investigación inician con el documento de los profesores María Antonia Manassero Mas y Ángel Vázquez Alonso, titulado "Evaluación de destrezas de pensamiento crítico: validación de instrumentos libres de cultura", cuyo propósito es presentar la validación de dos pruebas de evaluación de destrezas de pensamiento crítico con dos grupos de estudiantes de primaria y secundaria, que participan en un programa de aprendizaje basado en pensamiento. Entre los resultados se destacan la relación entre las puntuaciones de las pruebas y los criterios de validación independientes y externos, al igual que la importancia de estos estudios en el diseño de instrumentos libres de cultura.

Un segundo artículo, elaborado por los profesores Rodrigo Rodríguez Cepeda, Jaime Augusto Casas Mateus y Diana Esperanza Martínez Cárdenas, "Laboratorio de química bajo contexto: insumo para el desarrollo de habilidades de pensamiento crítico", presenta los resultados de un estudio en el que se aplican actividades y estrategias en torno al desarrollo de habilidades de pensamiento crítico, cuyo tema de estudio gira en torno a la química de los alimentos. Además de la propuesta didáctica se destaca el desarrollo de miniproyectos y el uso de instrumentos de evaluación propios de una metodología investigativa.

El tercer artículo de investigación formulado por la profesora Marisol Uribe Pérez tiene como título "Concepciones de profesores de ciencias en formación inicial sobre interculturalidad y su relación con la enseñanza. Reflexio-

nes en el contexto colombiano". Este presenta las concepciones acerca de interculturalidad en profesores de ciencias en formación inicial vinculados a la práctica docente, de nueve programas de licenciatura en ciencias de universidades públicas de Colombia. Entre los aportes de este trabajo, se destacan la interculturalidad para todos, la interculturalidad como encuentro, interrelación e intercambio entre diferentes culturas.

Un cuarto artículo titulado "La caja de polinomios y el método tradicional: dos alternativas didácticas para la enseñanza de la multiplicación y la división de polinomios", elaborado por los profesores José Martín Villarroel Solís y Natalia María Mazo Barrera, pretende comparar el nivel de aprendizaje que alcanzaron los estudiantes de grado octavo al trabajar multiplicación y división de polinomios desde dos estrategias, la caja de polinomios, fundamentada en la teoría de aprendizaje significativo y el método tradicional, que se fundamenta en los procesos de Explicar-Ejemplificar-Ejercitar. Los resultados del estudio no revelan cambios significativos en los promedios de los estudiantes; sin embargo, permiten abrir el espacio de discusión frente a nuevas maneras de trabajar el tema de los polinomios en clase de matemáticas.

El quinto documento de investigación de autoría de la profesora Yamile Pedraza Jiménez, titulado "La investigación-acción-participación para problematizar la ambientalización curricular universitaria", tiene por objeto describir el proceso de conformación de un grupo de discusión sobre la ambientalización curricular, en una universidad pública colombiana, mostrando el análisis del lenguaje, los discursos y la manera como se problematiza la ambientalización curricular.

El sexto artículo denominado "Desarrollo del pensamiento crítico y reflexivo a partir de la enseñanza del sonido", cuya autoría corresponde a los profesores Olga Lucía Castiblanco Abril y Yenny Lucila Fonseca, se propone desarrollar habilidades de pensamiento crítico y reflexivo en clase de física. La estrategia metodológica se centró en el abordaje del

sonido en fases, promoviendo cinco habilidades de pensamiento crítico: análisis, interpretación, inferencia, evaluación y autorregulación. Entre los hallazgos encontrados se resaltan las mejoras de las habilidades en el aprendizaje de los estudiantes en conceptos, tales como el tono, el timbre y la intensidad del sonido, además de mejorar la disposición para el debate, la expresión de diversas ideas y la argumentación.

Un séptimo artículo denominado “La cocina como estrategia para mejorar la enseñanza/aprendizaje de los conceptos de bioquímica”, de autoría de la profesora Liliana Pérez Cardona, pretende mejorar el proceso de enseñanza y aprendizaje de los conceptos de bioquímica, mediante la aplicación de guías de interaprendizaje acordes al modelo escuela nueva, utilizando la cocina como estrategia didáctica. Los resultados permiten inferir que el empleo de alternativas novedosas en el aprendizaje de la bioquímica fortalece y contribuye a la aplicación de procesos investigativos en la vida cotidiana.

El octavo artículo, de los profesores Nora Bahamonde, María Cecilia Cremer, Paula Natalia Mut, Eduardo Enrique Lozano, “El desarrollo de una línea disciplinar para la enseñanza del modelo de presión arterial en la formación del profesorado en biología”, expone una interesante propuesta de diseño, implementación y evaluación de una Unidad Didáctica para la modelización del fenómeno de presión arterial (PA) y su regulación, desde modelos teóricos del campo de la didáctica. En esta investigación se destaca el abordaje de cuestiones socio-científicas y la progresión en el aprendizaje desde modelos inicialmente simples propios de la organización celular, hasta modelos sistémicos del fenómeno de la presión arterial y su regulación.

El título del noveno artículo es “El papel de la reflexión y la mediación didáctica en la enseñanza de las ciencias naturales: un estudio de caso de profesores en formación”, de autoría del profesor Felipe Alejandro Quiroz Londoño. En esta investigación se propuso determinar la manera como la mediación didáctica y la reflexión en la acción contribuyen al desarrollo de actividades experimentales en docentes en formación en ciencias naturales. Para cumplir este propósito se desarrolló un proceso de reflexión desde el estudio de caso, con dos docentes del programa de Licenciatura en Educación Básica con énfasis en la Enseñanza de las Ciencias Naturales y Educación Ambiental. Los resultados apuntan hacia la importancia de incluir la mediación didáctica y la reflexión como estrategias que contribuyen a mejorar los procesos de formación docente.

El décimo artículo de investigación se titula “Atividades experimentais mediadas por tecnologias digitais de informação e comunicação em licenciaturas em ciências da natureza”, elaborado por los profesores Renata Isabelle Guaita y Fábio Peres Gonçalves. El propósito de esta investigación consistió en analizar la manera como los formadores de profesores del área de ciencias de la naturaleza entienden y promueven actividades experimentales mediadas por Tecnologías Digitales de Información y Comunicación (TDIC). De igual manera, se desarrollaron entrevistas desde la perspectiva del análisis textual discursivo (ATD), del que surgen las categorías “situaciones límite” e “inéditos viables” respaldadas por el educador Paulo Freire, demostrando la relación directa entre la experimentación, la motivación y la apropiación de técnicas de laboratorio, minimizando el potencial de las actividades experimentales mediadas por TDIC.

El artículo de *caso educativo*, de autoría de los profesores Rafael Enrique Gutiérrez Araujo y Luis Andrés Castillo Bracho, titulado “Simuladores con el software GeoGebra como objetos de aprendizaje para la enseñanza de la Física”, tiene por objeto presentar dos simuladores computacionales elaborados con el software GeoGebra para abordar los movimientos parabólico y armónico simple, caracterizándolos como objetos de aprendizaje, desarrollando un tratamiento didáctico a partir de dos planeaciones de clase.

Los artículos de reflexión presentados en este número 47 se detallan a continuación:

El primero de ellos se titula “Ciencia consentida. Resignificando los sentidos en la enseñanza de la ciencia”, cuya autoría corresponde al profesor José Joaquín García García. En este documento se aborda la noción de racionalidad moderna como abstracción, centro de debate en torno a la invisibilización del cuerpo. A partir de este planteamiento el autor propone recuperar el cuerpo para devolverle vida a las aulas usando los sentidos como fuentes de significado en la educación científica. De igual manera, se expone la necesaria explicación sobre la naturaleza de cada uno de los sentidos como receptores de ondas mecánicas y electromagnéticas, y como lectores de señales químicas.

El segundo artículo de reflexión, de autoría de los profesores Gastón Mariano Pérez y Leonardo Martín González Galli, se titula “Actividades para fomentar la metacognición en las clases de Biología”. En este texto los autores se proponen abrir un espacio de discusión frente a la adopción de criterios para clasificar actividades que fomenten la metacognición en el aula. Se destaca la ejemplificación de dichos criterios con actividades propias del campo de la Biología, enfatizando en un ejemplo específico sobre la enseñanza de la Biología evolutiva, con el fin de adelantar un tratamiento y abordaje de los obstáculos epistemológicos.

El artículo de revisión de tema titulado “El conocimiento tecnológico, pedagógico de contenido (TPCK): un análisis a partir de la relación e integración entre el componente tec-

nológico y conocimiento pedagógico de contenido”, elaborado por el profesor Jorge Mario Ortega, tiene como propósito la revisión de la categoría *conocimiento tecnológico/pedagógico del contenido* como una alternativa a la integración de la dimensión tecnológica en el currículo.

Agradecemos a nuestros lectores la divulgación de estos textos y la reflexión que puedan generar en las comunidades académicas que conforman el horizonte de sentido de la didáctica de las ciencias, las matemáticas y las tecnologías.

Referencias

- Alonso-Arévalo, J. (2016). Altmetrics: la integración del impacto científico y el impacto social de la investigación. En *Epistemología, acceso abierto e impacto de la investigación científica* (pp. 139-163). Cuenca: Universidad Católica de Cuenca.
- Alperin, J., Muñoz, C., Schimanski, L., Fishman, G., Niles, M. y McKiernan, E. (2019). Meta-research: How significant are the public dimensions of faculty work in review, promotion, and tenure documents? *eLife*. Recuperado de <https://elifesciences.org/articles/42254>
- Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H. y Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social web. *Proceedings of the 17th International Conference on Science and Technology Indicators*, Repro-UQAM, Montreal, Canada, 98-109.
- Declaración de San Francisco sobre la evaluación de la investigación*. Recuperado de <https://sfdora.org/read/es/>
- Fazel, S. y Lamsma, J. (2015). Beyond the impact factor? *Evidence Based Mental Health*, 18, 33-35.
- Fire, M. y Guestrin, C. (2019). Over-Optimization of academic publishing metrics: observing goodhart's law in action. *GigaScience*, 8(6), 1-20.

- Flick, C. M., y Convoco Foundation. (2011). *Geographies of the World's Knowledge*. Oxford: Oxford Internet Institute, Oxford University.
- Garfield, E. (1955). Citation indexes for science: a new dimension in documentation through association of ideas. *Science*, 122, 108-111.
- Goodhart, C. A. E. (1984). *Monetary theory and practice: The uk experience*. Londres: Macmillan.
- Google Scholar Metrics. (s. f). Recuperado de scholar.google.com/intl/en/scholar/metrics.html
- Hicks, D. y Wouters, P. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. Recuperado de <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569-16572.
- Moher, D., Bouter, L., Kleinert, S., Glasziou P. y Sham M. (2019). *The Hong Kong Manifesto for assessing researchers: Fostering research integrity*. Recuperado de https://www.wcri2019.org/uploads/files/2019/Hong_Kong_Manifesto_v9.pdf. Visitado octubre/2019
- Norman, F. (2017). *The history of peer review, and looking forward to preprints in biomedicine*. In *What might peer review look like in 2030?* Londres: SpotOn, BioMed Central, Digital Science.
- Priem, J., Taraborelli, D., Groth, P. y Neylon, C. (2010). *Altmetrics: A manifesto*. Recuperado de <http://altmetrics.org/manifesto>
- Ritter, D. (2013). Beyond the impact factor: New measures of journal impact. *Science Editor*, 36(3), 94.
- Williams, C. y Padula, D. (2015). *The evolution of impact indicators: From bibliometrics to altmetrics* [e-Book]. Recuperado de <https://scholasticahq.com/altmetrics-the-evolution-of-impact-indicators>