

Reflexiones a propósito de la política nacional para mejorar el impacto de las publicaciones científicas nacionales y el nuevo modelo de clasificación de revistas científicas

Didier Andrés Ávila Roncancio*

En agosto de 2016, el Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias), en cabeza de su Dirección de Fomento a la Investigación, dio a conocer el documento No. 1601 denominado “Política Nacional para Mejorar el Impacto de las Publicaciones Científicas Nacionales”, en el cual se identifica como problemática el bajo nivel de contribución de la comunidad científica nacional a la creación y producción del conocimiento mundial, situación que para el organismo rector del Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI), se encuentra asociado a tres principales razones:

1. El bajo número e impacto de publicaciones científicas de autores nacionales.
2. Los reducidos niveles de impacto de las revistas científicas editadas en el país.
3. Las publicaciones de autores nacionales y las revistas científicas tienen poco reconocimiento y un bajo nivel de citación en la comunidad científica internacional (Colciencias, 2016).

Como respuesta a los problemas detectados, la Dirección de Fomento y un Comité de Expertos en temas editoriales, diseñaron el nuevo *Modelo de Clasificación de Revistas Científicas*, el cual se encuentra constituido en dos etapas: una de ellas de diagnóstico y otra que arrojará los resultados oficiales, los cuales tendrán una vigencia de dos años. Ambas etapas estarán integradas por tres fases: (i) cumplimiento de la gestión editorial, (ii) validación, evaluación y visibilidad, e (iii) impacto de la revista científica.

En la primera de ellas, el cumplimiento de la gestión editorial, es obligación de la revista ubicarse dentro de una de las disciplinas, áreas y grandes áreas del conocimiento de acuerdo con la clasificación de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y según los temas tratados por la revista en sus diferentes fascículos, todo ello con el fin de unificar y homogeneizar las diferentes disciplinas que se incorporarían dentro de una misma gran área del conocimiento para, conforme a ello, fijar y medir los

* Abogado de la Universidad Libre. Estudios en Mercado Bursátil de la Bolsa de Valores de Colombia. Candidato a Magíster en Derecho Económico de la Pontificia Universidad Javeriana. Integrante del Grupo de Investigaciones Socio Jurídicas de la Universidad Libre, reconocido por Colciencias en categoría A. Asistente Editorial de la Revista *Diálogos de Saberes*.

máximos y mínimos en los que se organizará cada cuartil y así ubicar las revistas dentro de las diferentes categorías (A1, A2, B y C). De esta forma, por ejemplo, la Revista *Diálogos de Saberes* al publicar artículos de investigación concernientes sobre todo al Derecho, integra el gran área de las Ciencias Sociales, esto junto a revistas relacionadas con psicología, economía y negocios, ciencias de la educación, sociología, ciencias políticas, geografía social y económica, periodismo y comunicaciones, entre otras.

Quiere decir lo anterior que la dinámica de citación al interior de las comunidades científicas y académicas que componen las diferentes disciplinas del gran área del conocimiento, cobra especial importancia para la categorización de la revista en el Sistema Nacional de Indexación pues, al pertenecer a una misma gran área, los cuartiles se verán afectados por aquellas disciplinas en las cuales los investigadores acostumbran a citarse: a mayor número de citas en un área o disciplina, mayor será el valor del H5, lo que se traducirá a su vez en un mayor número de citas necesario para lograr los márgenes del 25-49% y 50-74, 9% de su gran área de conocimiento y así situarse por lo menos en las categorías B y C.

Es por ello que se considera una falencia en el diseño del Modelo de Medición acoger un enfoque de agrupación por áreas y no por disciplinas, más aún cuando las dinámicas de interacción entre comunidades académicas son muy diferente; así, mientras que en ciencias como la psicología y la economía es usual la citación entre investigadores, por la importancia de los antecedentes experimentales,

existen otras disciplinas en las que el nivel de citación es mucho más reducido, por lo que una agrupación como la propuesta por el Modelo, además de desconocer los hallazgos efectuados por la sociología de la ciencia –y que dan cuenta de las diferentes estructuras de la comunidad científica y sus diferencias en la construcción de la ciencia–, puede llegar a arriesgar los canales de divulgación del conocimiento científico de aquellas disciplinas en las que los niveles de citación son menores y para las cuales los cuartiles para ser reconocidos en categorías B o C resultan inalcanzables dados los altos niveles de citación de otras disciplinas de su misma área.

La situación descrita puede tener como efecto la reducción de los medios de divulgación científica, monopolizando las revistas científicas –en el caso de algunas disciplinas–, a través de las cuales los investigadores pueden dar a conocer sus resultados, lo cual desincentivará la producción intelectual y el trabajo científico al reducir las probabilidades de reconocimiento al interior de la respectiva comunidad académica y científica; en otros casos, puede llegar a dejar sin indexación algunas revistas dedicadas a la divulgación de temas muy específicos, lo cual le quitaría pluralismo a la ciencia, concretamente a aquellas que cuentan con un número reducido de científicos o expertos.

Frente a la fase (ii), la cual tiene por objetivo la revisión del proceso de evaluación y de visibilidad, es de resaltar los requisitos propuestos por el Modelo para que las revistas garanticen la exogamia, exigiendo, por ejemplo, que un porcentaje significativo de los artículos publicados sea producto de investigaciones

avaladas por instituciones diferentes a la editora, que los editores no publiquen más de un artículo al año en la revista por ellos editada, que al menos el 80% de los miembros del Comité editorial/científico tenga filiación externa a la institución editora y, finalmente, que por lo menos el 60% de los evaluadores de los artículos publicados en los diferentes fascículos de la revista no pertenezcan al Comité Editorial/Científico ni a la institución editora. Todos estos requisitos consideramos fortalecen la rigurosidad y científicidad de los procesos editoriales, contribuyendo a la calidad de los trabajos publicados en las revistas editadas por instituciones nacionales.

Pese a la importancia reconocida en los anteriores requisitos, es necesario anotar que en esta fase también se evalúa la visibilidad de la revista, para lo cual el Modelo exigió el cumplimiento de por lo menos una de las siguientes condiciones: a) Estar incluida en si quiera un Índice Bibliográfico Citacional (IBC); b) en al menos un índice Bibliográfico (IB); o c) en al menos una Base Bibliográfica con Comité Científico de Selección (BCCS). No obstante, y si bien se entiende el rol que cumplen dichos canales de comunicación científica en la divulgación y difusión de las revistas, no se comprende la justificación de que los IBC, IB y BCCS establecidos para cumplir o satisfacer dicha condición, fueran exclusivamente los dictados por Colciencias en el listado de Sistemas de Indexación y Resumen (SIR), más aún cuando no se ofreció una metodología para identificar los SIREs que integrarían la lista, lo cual de alguna manera deja entrever arbitrariedad en el proceso, sobre todo cuando sus efectos pueden ser el de eliminar el reconocimiento

como revista científica al medio de publicación en cuestión, pues la falta de cumplimiento de dicha condición inmediatamente excluye a la revista de ser indexada.

Por otra parte, y frente a la fase (iii) en la que se evalúa el impacto de los artículos publicados en la revista en términos de citas recibidas, el modelo de medición desconoce cómo influyen las diferentes publicaciones en el estado de la ciencia en el que se encuentre adscrita, así como de su disciplina, pues ubicándose en el campo exclusivo del Derecho, el impacto de un artículo científico será mayor en la medida en que tenga efectos transformadores en el estado del arte de la mencionada ciencia, pues no tendrá el mismo impacto aquella publicación, por más que haya sido citada en otros artículos, frente a una que ha sido considerada como fuente de Derecho de carácter doctrinal para la solución o respuesta a un problema jurídico, en una sentencia o como referencia para la construcción de una política pública para la solución de un problema social.

Finalmente, es importante poner de presente que el Modelo, así como la Política para Mejorar el Impacto de las Publicaciones, desconoció los **diez principios del Manifiesto de Leiden sobre Indicadores de Investigación**, por lo que llamamos la atención a las autoridades encargadas para que estos sean analizados y repensados y, a partir de ellos, sea replantado el Modelo, teniendo presente la necesidad de proteger e incentivar la creación de ciencia que se adecue a las necesidades del contexto en el que se produce y que sea evaluada con parámetros que vayan más allá de indicadores cuantitativos que deshumanizan la ciencia y la restringen a una serie de factores que

no necesariamente responden a la calidad y validez de sus resultados, así como a su riqueza, entendida como su aporte a la solución de necesidades que contribuyan a una mejor calidad de vida individual y en comunidad.

Los mencionados principios indican:

1. La evaluación cuantitativa tiene que apoyar la valoración cualitativa por expertos.

Los indicadores pueden corregir la tendencia a perspectivas sesgadas que se dan en la revisión por parte de los Pares y facilitar la deliberación. En este sentido, los indicadores pueden fortalecer la evaluación por Pares puesto que tomar decisiones sobre colegas es difícil sin varias fuentes de información. Sin embargo, los evaluadores no deben ceder a la tentación de supeditar las decisiones a los números. Los indicadores no pueden sustituir a los razonamientos informados. Los decisores tienen plena responsabilidad sobre sus evaluaciones.

2. El desempeño debe ser medido de acuerdo con las misiones de investigación de la institución, grupo o investigador.

Los objetivos de un programa de investigación deben ser especificados al principio y los indicadores usados para medir el desempeño deben estar claramente relacionados con estos objetivos. La elección y usos de los indicadores han de tener en cuenta los contextos socio-económicos y culturales. Los científicos tienen diversas misiones de investigación. La investigación para avanzar las fronteras del conocimiento académico es diferente de la

investigación focalizada en proveer soluciones a problemas sociales. La evaluación puede estar entonces basada en méritos relevantes para la industria, el desarrollo de políticas o para los ciudadanos en general, en vez de méritos basados en nociones académicas de excelencia. No hay un modelo de evaluación aplicable en todos los contextos.

3. La excelencia en investigación de relevancia local debe ser protegida.

En muchas partes del mundo, excelencia en investigación se asocia únicamente con publicaciones en inglés. La ley española, por ejemplo, explicita el deseo y la conveniencia de que los académicos españoles publiquen en revistas de alto impacto. El factor de impacto se calcula para revistas indexadas por Web of Science, que es una base de datos basada en el sistema estadounidense y que contiene una gran mayoría de revistas en inglés. Estos sesgos son especialmente problemáticos en las ciencias sociales y las humanidades, áreas en las que la investigación está más orientada a temas regionales y nacionales. Muchos otros campos científicos tienen una dimensión nacional o regional, por ejemplo, epidemiología del VIH en el África subsahariana.

Este pluralismo y la relevancia social tienden a ser suprimidos cuando se crean artículos que resulten de interés a los guardianes del alto impacto: las revistas en inglés. Los sociólogos españoles muy citados en *Web of Science* han trabajado en modelos abstractos o estudiando datos de los Estados Unidos. En ese proceso se pierde la especificidad de los sociólogos con alto impacto en las revistas en castellano: temas como la ley laboral local,

atención médica para ancianos o empleo de inmigrantes. Indicadores basados en literatura de alta calidad no inglesa que servirían para identificar y recompensar la excelencia en investigación localmente relevante.

4. Los procesos de recopilación y análisis de datos deben ser abiertos, transparentes y simples.

La construcción de las bases de datos necesarias para evaluar debe seguir procesos establecidos antes de que la investigación sea completada. Ésta ha sido la práctica común entre los grupos académicos y comerciales que han desarrollado metodologías de evaluación durante varias décadas. Estos grupos publicaron los protocolos de referencia en la literatura revisada por pares. Esta transparencia permite el escrutinio y control de los métodos. Por ejemplo, en 2010, un debate público sobre las propiedades técnicas de un importante indicador utilizado por uno de nuestros grupos (el Centro de Estudios de Ciencia y Tecnología –CWTS, de la Universidad de Leiden, en los Países Bajos), se saldó con una revisión en el cálculo de este indicador. Las nuevas empresas comerciales en el campo deben responder a los mismos estándares. Nadie tiene porque aceptar evaluaciones automáticas salidas de caja negras o procesos impenetrables. La simplicidad es una virtud en un indicador porque favorece la transparencia. Pero indicadores simplistas pueden distorsionar la evaluación (véase el Principio 7). Los evaluadores deben esforzarse en encontrar un equilibrio: indicadores simples que sean respetuosos con la complejidad de los procesos de investigación descritos.

5. Los datos y análisis deben estar abiertos a verificación por los evaluados.

Con el fin de asegurar la calidad de los datos, los investigadores incluidos en estudios bibliométricos tienen que estar en la capacidad de comprobar que sus contribuciones han sido correctamente identificadas. Los responsables y gestores de los procesos de evaluación deben garantizar la exactitud de los datos usados mediante métodos de auto-verificación o auditoría por terceras partes. Las universidades podrían implementar este principio en sus sistemas de información. Este debería ser un principio rector en la selección de proveedores de dichos sistemas. La compilación y proceso de datos de alta calidad, precisos y rigurosos, lleva tiempo y cuesta dinero. Los responsables deben asignar presupuestos a la altura de estas necesidades de calidad.

6. Las diferencias en las prácticas de publicación y citación entre campos científicos deben tenerse en cuenta.

La mejor práctica en evaluación es proponer una batería de indicadores y dejar que los distintos campos científicos escojan los indicadores que mejor les representan. Hace unos años, un grupo de historiadores recibió una puntuación relativamente baja en una evaluación nacional de pares porque escribían libros en vez de artículos en revistas indexadas por Web of Science. Estos historiadores tuvieron la mala suerte de formar parte del departamento de psicología. La evaluación de historiadores y científicos sociales requiere la inclusión de libros y literatura en la lengua

local; la evaluación de investigadores en informática necesita considerar las contribuciones a conferencias.

La frecuencia de citación varía según los campos: las revistas más citadas en *rankings* de matemáticas tienen un factor de impacto de alrededor de 3; las revistas más citadas en *rankings* de biología celular cuentan con factores de impacto de alrededor de 30.

Por lo tanto, se necesitan indicadores normalizados por campo. El método más robusto de normalización está basado en percentiles: cada publicación es ponderada según el percentil al que pertenece en la distribución de citaciones de su campo (por ejemplo, el percentil 1%, 10%, 20% más alto). Una única publicación altamente citada mejora un poco la posición de una universidad en un *ranking* basado en percentiles, pero puede propulsar la universidad de un lugar medio a las primeras posiciones en un *ranking* basado en promedios de citas.

7. La evaluación individual de investigadores debe basarse en la valoración cualitativa de su portafolio de investigación.

El índice-*h* aumenta con la edad del investigador, aunque este ya no publique. El índice-*h* varía por campos: los científicos en las ciencias de la vida pueden llegar a 200; los físicos a 100 y los científicos sociales a 20 o 30. Se trata entonces de un índice que depende de la base de datos: hay informáticos que tienen un índice-*h* de 10 en Web of Science, pero de 20 o 30 en Google Scholar. Leer y valorar el trabajo de un investigador es mucho más apropiado que confiar en un único número.

Incluso cuando se comparan un gran número de científicos, es mejor adoptar un enfoque que considere información diversa sobre cada individuo, incluyendo sus conocimientos, experiencia, actividades e influencia.

8. Debe evitarse la concreción impropia y la falsa precisión.

Los indicadores de ciencia y tecnología tienden a la ambigüedad conceptual y a la incertidumbre. Se fundamentan en hipótesis que no están universalmente aceptadas. Por esta razón, las buenas prácticas usan múltiples indicadores con el fin de construir un retrato robusto y plural. En la medida que sea posible cuantificarla, la información sobre incertidumbre y error debería estar acompañada por los valores de los indicadores publicados, por ejemplo, usando barras de error. Si esto no fuera posible, los productores de indicadores deberían al menos evitar el ofrecer un falso nivel de precisión. Por ejemplo, el factor de impacto de revistas se publica con tres decimales para evitar empates. Sin embargo, dada la ambigüedad conceptual y la variabilidad aleatoria de las citas, no tiene sentido distinguir entre revistas por pequeñas diferencias en el factor de impacto. Se debe evitar la falsa precisión: solo un decimal está justificado.

9. Deben reconocerse los efectos sistémicos de la evaluación y los indicadores.

Los indicadores cambian el sistema científico a través de los incentivos que establecen. Estos efectos deberían ser anticipados. Esto significa que una batería de indicadores es siempre

preferible puesto que un solo indicador es susceptible de generar comportamientos estratégicos y sustitución de objetivos (y allí la medida se convierte en un fin en sí mismo). Por ejemplo, en los años 90, Australia financió investigación en universidades de acuerdo con una fórmula basada, sobretodo, en el número de publicaciones de un instituto. Las universidades podían calcular el “valor” de una publicación en una revista arbitrada; en el año 2000, el valor se estimó en Aus\$800 (US\$480) destinados a recursos de investigación. Como era de esperarse, el número de artículos publicados por autores australianos subió, pero en revistas menos citadas, lo que sugiere que la calidad de los artículos disminuyó.

10. Los indicadores deben ser examinados y actualizados periódicamente.

Las funciones de la investigación y los objetivos de la evaluación cambian o se desplazan, y el sistema de investigación co-evoluciona con ellos. Medidas que fueron útiles en su momento pasan a ser inadecuadas por lo que nuevos indicadores aparecen. Por lo tanto, los sistemas de indicadores tienen que ser revisados y tal vez modificados. Al darse cuenta de los efectos de su fórmula simplista de evaluación, en 2010 Australia adoptó la iniciativa Excellence in Research for Australia, que es más compleja y hace especial énfasis en el factor *calidad*.