

Ampliar los criterios de evaluación de la investigación: el Manifiesto de Leiden (2015)



Por **Antonio Bolívar**
Catedrático de Didáctica y Organización Escolar de la Univ. de Granada
Editor de la Revista “Profesorado. Revista de Curriculum y Formación del Profesorado”

Publicado el 5 de abril de 2019.

En el Congreso ECER 2014, celebrado en Oporto y dedicado al tema “The past, present and future of educational research in Europe”, Antonio Nóvoa impartió una conferencia (“What is educational research for?”), ahora publicada (Nóvoa, 2015), en que se preguntaba críticamente, a la luz de lo que estaba pasando, por qué esta vorágine de publicaciones, señalando: “Todos sabemos la respuesta: se trata de una lucha por la supervivencia en la selva académica (p. 266). Concluía su intervención: “las formas dominantes de organización del trabajo académico y de evaluación de los profesores están afectando gravemente la idea de universidad y nuestras vidas profesionales y personales. Ha llegado el tiempo de repensar la investigación educativa desde una perspectiva más amplia, con una libertad sin condiciones. Porque la libertad es todo, y todo lo demás es nada” (p. 271).

En efecto, un conjunto de factores confluyentes, comunes a otros países, pero agudizados en éste, han conducido la investigación educativa a una situación que es preciso repensar y, en la medida de lo posible, poner límites. Particularmente porque estamos orientando a los jóvenes investigadores a qué es lo que importa: tener muchas publicaciones y en revistas con mayor factor impacto, al margen de ser fruto de un pensamiento e investigación reposados. Entre otros, está provocando que nuestras revistas estén colapsadas, teniendo algunas que cerrar, periódicamente, la recepción de artículos, para poder gestionar los cientos que se reciben. Momento, pues, para llegar a decir: basta, otros modos de trabajar y difundir la investigación educativa son posibles y deseables. Lo que siempre ha sido es primar la calidad de trabajos bien hechos y reposados sobre la cantidad de papers. La calidad y los planteamientos propios y originales deben primar sobre la cantidad.

España, de los pocos países, que de modo incuestionado ha abrazado las métricas habituales de Thompson & Reuters o Scopus, en lugar de optar por “métricas

responsables” que incrementen la evaluación de equilibrio y equidad. La acreditación a CU o TU, así como los sexenios, dependen del factor de impacto de las publicaciones presentadas, sin cuestionar ni entrar en la calidad o contenido. En lugar de un uso responsable de los indicadores métricos, los aplicamos de forma automática como máquinas evaluativas, para lo que no haría falta un experto, bastaría un administrativo que cuenten el número y el index. Nuestra larga tradición burocrática, al tiempo que aceptar ortodoxias sin cuestionarlas, nos han conducido a una situación que, alejada sobre cómo se está haciendo a nivel internacional, debiera cuestionarse a nivel local.

Algunas de las mejores Universidades norteamericanas, a la hora de seleccionar a su profesorado, solicitan que, de todo su trabajo, seleccione lo que estima más relevante. La Comisión examina y discute con el autor dichos trabajos, no de modo anónimo recontando el número de publicaciones en revistas de factor impacto, sin tener que entrar (ni leer) en lo publicado. Este modo de selección, presente en los modos vigentes de acreditación, es único en el mundo y deberíamos pensar si, por su excepción, es el mejor. Es cierto que, las revistas suelen tener sus revisores o que, después de la acreditación, una Comisión de la propia Universidad revalida la acreditación con el acceso a la plaza. Pero todos sabemos (y conocemos) que son -en su mayor parte- actos protocolarios, una vez acreditado.

Es un clamor la necesidad de métricas alternativas. Hasta el propio Comisario Europeo de Investigación, Ciencia e Innovación (Carlos Moedas) ha abogado por otras formas alternativas de evaluación de la investigación que puedan incentivar mejor la práctica científica, así como la Open Science (European Commission, 2016). La Declaración DORA (Declaración de San Francisco sobre Evaluación de la Investigación), impulsada por la Sociedad Americana de Biología Celular, ya advirtió, con buen criterio, la necesidad de evaluar la investigación por sus propios méritos y no por la revista donde se publica. Va *contra natura* que lo que importe sea dónde se publicó, al margen del contenido, relevancia u originalidad de lo publicado. Los enfoques cuantitativos, concentrados en un solo indicador (factor impacto), son reduccionistas y tienen graves defectos. Se clama por practicar la evaluación cuantitativa de la investigación de manera responsable.

El llamado “Manifiesto de Leiden” (Hicks et. al., 2015) es un de las más relevantes llamadas y criterios en esta dirección. Redactado por cinco expertos en evaluación científica (uno de ellos, Ismael Rafols, es español), presenta un decálogo de principios para el empleo transparente y eficaz de los indicadores bibliométricos, y que ha sido publicado por la prestigiosa revista *Nature*. El Manifiesto agrupa diez principios las mejores prácticas de ejercicios bibliométricos. como guía para el empleo de los indicadores bibliométricos en la evaluación de la actividad investigadora (Coombs y Peters, 2017). En 2016, la Asociación Europea para el Estudio de la Ciencia y la Tecnología le otorgó el Premio Ziman. Los 10 principios se pueden resumir en:

- 1.- Los indicadores, como información cuantitativa que son, no pueden sustituir a los criterios de evaluación valorativos de los expertos, sino que deben de servirles de ayuda. Por tanto, las evaluaciones cuantitativas deben ser evaluadas de forma cualitativa y por expertos. Criterios cuantitativas y cualitativas deben informarse mutuamente.
- 2.- La evaluación de la actividad investigadora debe ajustarse a la misión u objetivos de la institución, investigador o grupo que está siendo evaluado. Las métricas se utilizan para medir los méritos y el rendimiento en relación con objetivos de investigación específicos. Por lo tanto, los objetivos deben hacerse explícitos y

coincidir con los indicadores adecuados. Dado que no puede existir una solución única para todos, las evaluaciones pueden (y deberían) ser individuales.

3.- Se han de desarrollar indicadores que reflejen el impacto de actividades de investigación de ámbito local o regional y que se desarrollan en otras lenguas distintas al inglés. Por tanto, lo prioritario es proteger la excelencia en la investigación relevante a nivel local

4.- Los indicadores deben ser simples, respetando la complejidad, para facilitar la transparencia de las evaluaciones. La transparencia y la apertura con respecto a la recopilación de datos y los métodos para el desarrollo de indicadores y los informes de evaluación son fundamentales, ya que a la vez proporciona autoridad para el análisis de los mismos y permite la revisión y el ajuste.

5.- Los evaluados deben poder verificar el análisis de los indicadores por los que se les ha evaluado, y si no están de acuerdo pedir una reevaluación. El cumplimiento de este principio garantiza la integridad de la evaluación de la investigación y de las personas que la ejecutan.

6.- Se deben tener en cuenta las diferencias en el impacto que existe entre distintos campos de investigación a la hora de elaborar los indicadores. No todos los campos de investigación son similares, ni se pueden valorar con los mismos indicadores. Este principio es clave para valorar responsablemente la investigación.

7.- No se pueden usar los indicadores sin tener en cuenta todo el contexto del investigador. La evaluación de base de investigadores individuales tiene que hacerse dentro de la comunidad profesional en que trabaja, en un juicio cualitativo de su *currículum vitae*.

8.- Hay que evitar las falsas concreciones y precisiones. Comprender las métricas, los datos y los métodos para evitar la falsa precisión al mostrar resultados; los indicadores deben presentarse de una manera que deje clara su precisión limitada para el usuario (por ejemplo, indicadores múltiples de identificación con referencias a otra fuente de datos esperada); evitar extrapolaciones engañosas o conclusiones generales.

9.- Hay que tener en cuenta el efecto que provocan algunos indicadores en tanto que incentivos que pueden favorecer unas actividades y desfavorecer otras. Por eso, se requiere reconocer los efectos del sistema de evaluación y los indicadores. El uso de un solo indicador invita juicios no deseados donde la medición se convierte en el objetivo.

10.- Los indicadores han de ser revisados y actualizados periódicamente. Es preciso examinar los indicadores de manera regular para su progresiva adecuación y mejora.

En fin, el Manifiesto de Leiden cuestiona los cánones incuestionados y plantea la necesidad de transparencia de los datos y métodos subyacentes; abogando por una combinación por defecto de análisis cualitativos y cuantitativos; reconocimiento las limitaciones de los datos y métodos utilizados, y abogando por una coherencia o alineación entre el propósito de la evaluación, elección de indicadores y disciplina. Como finaliza el Manifiesto: “no se debe permitir que la información cuantitativa se convierta en un objetivo en sí misma. Las mejores decisiones se toman combinando estadísticas robustas sensibles a los objetivos y la naturaleza de la investigación evaluada”. las métricas no pueden *sustituir* las decisiones informadas.

Como señala Nóvoa (2015) en el trabajo citado, este tipo de prácticas y orientaciones están destruyendo nuestra vida académica y nuestra libertad intelectual. Estas ideologías están empobreciendo el trabajo científico, en particular en el campo de la investigación en educación (p. 269). La valoración de la calidad de una investigación o de un artículo debe basarse en la investigación misma y no, por ejemplo, en el factor de impacto de la revista donde se haya publicado. No hacerlo así significa que los evaluadores abandonan su responsabilidad¹. Entre las nuevas formas de conspiración contra la ciencia (Kaufman y Kaufman, 2018), están proliferando las revistas depredadoras, que denunciamos en otra contribución anterior.

Referencias bibliográficas:

- Coombs, S.K., Peters, I. (2017). The Leiden Manifesto under review: what libraries can learn from it. *Digital Library Perspectives*, 33 (4), 324-338.
- European Commission (2016). *Open Innovation, Open Science, Open to the World - A Vision for Europe*. Luxembourg: DG Research and Innovation, Publications Office, doi: 10.2777/061652
- Hicks et al. (2015). The Leiden Manifesto for research metrics. *Nature*, 520, 429-431. <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>. Trad. cast.: www.ingenio.upv.es/manifiesto
- Nóvoa, A. (2015). Em busca da liberdade nas universidades: para que serve a pesquisa em educação? *Educação e Pesquisa*, São Paulo, 41 (1), 263-272. <http://www.scielo.br/pdf/ep/v41n1/1517-9702-ep-41-1-0263.pdf>
- Kaufman, A. y Kaufman, J. (Eds.) (2018). *Pseudoscience : the conspiracy against science*. Cambridge, Massachusetts: The MIT Press.

Cómo citar esta entrada:

- Bolívar, A. (2019). Ampliar los criterios de evaluación de la investigación: el Manifiesto de Leiden (2015). *Aula Magna 2.0*. [Blog]. Recuperado de: <https://cuedespyd.hypotheses.org/6166>

¹Sobre los manifiestos DORA y Leiden sobre indicadores de investigación, ver el blog de Lluís Codina <https://www.lluiscodina.com/leiden-manifiesto-dora-metricas/>