



Atrapados en el sistema

Trapped in the system

Se analizan algunas de las fortalezas y debilidades de los procedimientos actuales de evaluación de artículos en revistas científicas y su papel en el marco general de la evaluación profesional.

Palabras clave: difusión de artículos científicos, evaluación por pares, evaluación profesional, evaluación cuantitativa *versus* evaluación cualitativa.

This paper discusses the strengths and weaknesses of the current procedures of evaluating articles in research journals, as well as their role in the overall framework of the professional evaluation.

Keywords: research paper diffusion, peer review, professional evaluation, qualitative versus quantitative evaluation.

Uno de los aspectos que fueron objeto de atención preferente durante la mesa redonda celebrada con motivo del 25 aniversario de la *Revista d'Arqueologia de Ponent* fue la cuestión de la evaluación de las revistas científicas en general y de las de arqueología en particular, y es a ello a lo que dedicaré las líneas que siguen.

Sin embargo, no debemos olvidar que los actuales mecanismos de evaluación de revistas científicas se insieren en un marco más amplio, el de la evaluación científica y profesional, que no es posible obviar a la hora de aproximarse a los primeros. La evaluación científica opera en diversos campos interrelacionados: la incorporación laboral, la promoción y evaluación continua (sexenios) de investigadores y profesores universitarios, la financiación de proyectos y, para terminar por el principio, la formulación de políticas

para la ciencia. En lo que respecta a los profesionales, en realidad se trata de un sistema de valoración piramidal cuya base se sustenta en la evaluación de la producción científica en los medios de difusión especializados y en cuya cúspide se sitúa la consolidación y promoción profesional y la obtención de fondos para la investigación. Así pues, la relevancia de la evaluación de las revistas científicas trasciende la casuística de cada una de ellas, e incluso de su conjunto, para erigirse indirectamente en la columna vertebral del sistema de ciencia de un país.

Los sistemas de evaluación científica se han basado tradicionalmente en la valoración por los propios colegas del campo de conocimiento de que se trata. Sus antecedentes se remontan a los primeros pasos de la llamada institucionalización de la ciencia, que se retrotraen a la fundación de las primeras socie-

dades (Royal Society) y a la edición de las primeras revistas científicas a partir del siglo XVII (*Journal des Savants*, París 1665; *Philosophical Transactions of the Royal Society*, Londres 1665; *Acta Eruditorum*, Leipzig 1682). La Royal Society estableció como norma para la publicación de textos en su revista, *Philosophical Transactions*, un informe previo favorable de un miembro de la sociedad. De manera que el mecanismo actualmente en uso —con sus bondades y flaquezas—, la evaluación por pares, no supone más que la sistematización de una práctica, la autoevaluación cualitativa y colegiada por parte de los científicos, que se remonta en el tiempo a la propia organización de la ciencia como sistema social y a su sanción por consenso.

No sucede lo mismo con los índices de impacto de las publicaciones científicas, basados en indicadores bibliométricos y que constituyen una valoración cuantitativa de la investigación. Su historia es más reciente y quizás el hito principal de la misma es el diseño por parte de Eugene Garfield del llamado Factor de Impacto en los años cincuenta del pasado siglo, la fundación del Institute for Scientific Information (ISI) y el inicio de la publicación del *Science Citation Index* —que desde 1992 pertenece al consorcio editorial Thomson Corporation (Borrego, Urbano 2006). La indexación y categorización de revistas científicas, el otro gran vector de la evaluación de las publicaciones, se basa en buena medida en criterios editoriales y no de calidad de los contenidos, en relación con los cuales el único requisito que se exige es la evaluación por pares.

Así pues, el actual sistema de evaluación de publicaciones científicas en revistas se basa en un doble parámetro, cualitativo (revisión por pares) y cuantitativo. Sin embargo, en la pirámide que, como hemos dicho, constituye el conjunto del contemporáneo sistema de evaluación de la ciencia ha venido primando la evaluación cuantitativa. Un sistema en el que solo el primer escalón, la evaluación por pares, constituye una perspectiva cualitativa. En el caso español, el sesgo cuantitativo ha sido especialmente marcado, aunque no es ni mucho menos el único país occidental en que se ha producido esta tendencia, como ha venido siendo denunciado por investigadores de otros países. A ello se ha sumado en los últimos años lo que algunos llaman sesgo gerencial (Boure 2015), claramente vinculado al destronamiento, aunque no supresión, de la ciencia en favor de la tecnociencia (Echeverría 2003). En la actualidad una presión competitiva enorme por los puestos y por la financiación de proyectos provoca algunos efectos indeseados. Entre los más relevantes el exceso de publicaciones, la re-publicación (previo re-redactado) en diversos medios científicos y la consiguiente falta de reflexión, que inciden en la calidad final de los textos (Lévy-Leblond 2003: 144).

Por otra parte, el modelo empresarial y neocapitalista que caracteriza hoy nuestras sociedades tiene una clara repercusión en todo lo relativo a la ciencia y a la enseñanza y, como no podía ser de otro modo, en las publicaciones científicas periódicas, cuyo modelo, tradicionalmente público, ha ido cediendo lugar cada vez más a la empresa privada, cuyos fines son principalmente comerciales y que cuentan con una

batería de poderosas estrategias para posicionarse adecuadamente que no están al alcance de las revistas editadas y patrocinadas por universidades y entes públicos. Muchas de las grandes bases de datos internacionales dependen de estas mismas empresas privadas. Aunque a algunos no les agrada oírlo (no hay más que ver las descalificaciones que ha recibido el Nobel Randy Schekman por su reciente posicionamiento en contra de las que algunos llaman “revistas de lujo”), la situación ha llegado tan lejos que cada vez se levantan más voces de denuncia. Por citar solo algunos ejemplos, el manifiesto llamado *The Cost of Knowledge* llamando al boicot a las grandes corporaciones editoriales privadas (<http://thecostofknowledge.com/>) o la más reciente plataforma de debate *Why are we not boycotting Academia.edu?* (disruptivemedia.org.uk/portfolio/why-are-we-not-boycotting-academia-edu) que denuncia la relación parasitaria de la gran corporación de San Francisco con el sistema público y que no es más que uno de los ejemplos de tales prácticas por parte de empresas privadas, rentabilizar en el sector privado las inversiones realizadas por el sector público. Recientemente Academia.edu, de la que personalmente soy usuaria, además de rentabilizar los datos que aportamos, ha empezado a ofrecer “servicios extra” que ya no son gratuitos. En concreto un *Premium Account* (9,99 € al mes) para tener informes de impacto, analíticas detalladas, búsqueda avanzada... y, lo que es peor, ofreciendo un servicio (pagando, claro) por el cual los artículos del pagador ¡figurarían entre los recomendados, aumentando así su visibilidad! (<http://publishingarchaeology.blogspot.com.es/2016/01/academiaedu-wants-to-commercialize-its.html?m=1>). Sin embargo, a pesar de toda la parafernalia relativa al factor de impacto y a todo lo relacionado con el mismo, los propios datos de ISI muestran que dos tercios de los artículos que se publican nunca son citados y que, en caso de serlo, su visibilidad real es bastante limitada (Lévy-Leblond 2003).

Los actuales sistemas de evaluación de las revistas científicas han tenido efectos beneficiosos para el sistema de ciencia español. Sin embargo, en términos generales tienen también efectos perversos, que han sido puestos de relieve a menudo en la bibliografía especializada: la estandarización de la conducta investigadora, la cohesión y conformidad ideológica entre investigadores, la discriminación por vía del nepotismo y el sexismo son algunas de ellas (Pérez Sedeño 2007; Wennerås, Wold 1997). Igualmente, se ha incidido repetidamente en la debilidad de las ciencias sociales y humanas en el terreno de la evaluación de sus publicaciones periódicas debido a su carácter a menudo delimitado desde el punto de vista territorial, a la frecuente utilización de lenguas vernáculas y a sus generalizaciones débiles frente a las leyes generales de las ciencias naturales, elementos todos ellos que se encuadran con dificultad en un panorama en que los indicadores internacionales actúan miméticamente en relación con las prácticas de las ciencias de la naturaleza. Por otra parte, el enfoque eminentemente cuantitativo de la “pirámide evaluadora” y la enorme presión que soportan los investigadores, especialmente los llamados jóvenes investigadores —por desgracia cada vez mayores

en edad— ha tenido también como consecuencia el seguidismo de un lema, en parte cierto y en parte convertido en una especie de leyenda urbana, según el cual es mucho mejor para el currículum profesional publicar dos artículos indexados que una monografía (con el valor añadido de que los primeros dan mucho menos trabajo). Por ello, a pesar de contar aún con manifiestas debilidades, creo muy relevante la relativamente reciente implementación de la plataforma SPI (Scholarly Publishers Indicators, <<http://ilia.cchs.csic.es/SPI/>>). La consecuencia ha sido un descenso de la edición de este tipo de publicaciones, un vehículo de comunicación científica importante en las ciencias sociales y humanas. Igualmente, los editores de revistas deben procurar no solo por el nivel de las mismas sino por su visibilidad, lo cual les obliga a tomar decisiones dictadas por esta necesidad. Buela-Casal (2002) propone, en tono irónico, un decálogo —que denomina los Diez Mandamientos— para incrementar las citas: incrementar la difusión de la revista, incluirla en el mayor número posible de bases de datos, publicar artículos polémicos, publicar revisiones, publicar en inglés, publicar sobre temas de actualidad, publicar a autores muy citados, establecer acuerdos con medios de comunicación y recomendar que se citen trabajos de la misma revista. Sin abandonar el terreno de la ironía, permítaseme afejar al autor de tales mandamientos el que no haya incluido algo así como: mantener contra viento y marea una línea editorial, se adapte o no a los estándares; un planteamiento que, a mi modo de ver, es el último recurso del editor frente a la sobrepresencia de dichos estándares (responsables, entre otras cosas, de que mantener una sección de *Debat* como esta sea para una revista más una losa que un mérito cuantificable en el marco de los requerimientos al uso). A mi modo de ver, la calidad de una revista reside también en la existencia de una línea editorial que cumpla unos objetivos que constituyan una aportación personalizada en su campo y en el interés y calidad de los textos que publica. Esto último depende en buena medida de la evaluación por pares, que los investigadores realizamos de forma gratuita y a la que no siempre dedicamos una atención suficiente, no tanto por esa falta de remuneración como por falta de tiempo. Sin embargo, no todo depende de la política editorial, las políticas personales de publicación de los investigadores tienen también una incidencia importante en el panorama de las publicaciones científicas periódicas, en nuestro caso las de arqueología. En general los sistemas de valoración al uso impelen a todos los investigadores, aunque muy especialmente a los no consolidados, a publicar o intentar publicar en las revistas que reciben calificaciones más altas presentando textos que se adapten a sus líneas editoriales que, en términos generales, se caracterizan por requerir brevedad, novedad y capacidad de atracción de citas. Todo ello constituye una justificación y retroalimentación del sistema, no siempre con resultados relevantes.

Efectivamente, el proceso de implantación de este tipo de evaluaciones ha tenido también como consecuencia la reducción del número de páginas aceptadas por las grandes revistas, la necesidad de titular los artículos de forma atractiva, venga o no a cuento

(hay sumarios de alguna prestigiosísima revista de arqueología que más parecen un índice de *National Geographic*), de plantear siempre temas “innovadores” y, ante todo, publicar solo interpretaciones y los mínimos datos posibles, pues los datos son antiguos y, como algunos repiten a menudo, “no son ciencia”, olvidando que si bien los datos, efectivamente, no son por sí mismos ciencia, sin ellos solo hay pura especulación. Un elemento a tener en cuenta en relación con esta última cuestión es la irrupción de los llamados macrodatos (*Big Data*) que, en parte con sólidas razones, han despertado el entusiasmo de muchos investigadores (Kristiansen 2014), pero respecto de los cuales hay que recordar, por una parte, que “las nuevas tecnologías son inútiles a menos que científicos brillantes planteen preguntas relevantes e interpreten los datos en el contexto que corresponda” (Rossell 2014) y que está por ver la concreción en nuestro campo y las condiciones de acceso futuras.

Algunos han reflexionado sobre las estrategias y el lema “escribir es comunicar para ser leído” (Ruiz Zapatero 2014). Desde luego, pero igualmente no nos van a leer. Como ya he comentado, todos los estudios bibliométricos indican que muy poco de lo que se escribe (y a pesar de las tácticas mercadotécnicas para llamar la atención) es realmente leído y citado y, lo que es más, como los electrodomésticos, cae en el plazo de cuatro o cinco años en la obsolescencia. El problema no es solo cómo escribimos, la cuestión afecta al entero sistema de la ciencia occidental. El resultado de todo ello es que investigadores y editores públicos están atrapados por el sistema, con el que se puede ser crítico, pero que no se puede ignorar so pena de ser lanzado a las profundidades del anonimato y a la penuria laboral, según sea cada caso.

Las reiteradas críticas al exagerado sesgo cuantitativo de la pirámide evaluadora han propiciado que se trate de corregirlo mediante evaluaciones más cualitativas, que están ya en marcha en algunos países. En España se ha publicado recientemente un decreto relativo a la evaluación para la acreditación nacional para el acceso a los cuerpos docentes universitarios (BOE 144, 17/06/2015) cuyo objetivo es orientar la actividad de ANECA en esta dirección. Sin embargo, si no se dota de los medios necesarios a la agencia, y más bien todo indica lo contrario, puede que nos pase como a aquel que estando bien quiso estar mejor; pues la evaluación cualitativa, si no va acompañada de unos estrictos procedimientos y de los medios económicos para desarrollarla en plenitud puede entrañar vicios importantes, y el nepotismo solo es uno de ellos.

* * *

A pesar de las críticas —algunas de ellas cuando menos dignas de tener en consideración— vertidas en los últimos años en relación con el proceso conocido como “Transición”, no hay duda de que el cambio de régimen político en el Estado español supuso una modernización notable en muchos terrenos que tornó España de un país cerrado en uno de una mayor apertura. El campo científico no fue ajeno a estos cambios: la Ley de la Ciencia de 1986 (sustituida

en 2011 por la Ley de Ciencia, Tecnología e Innovación), la entrada en la UE al mismo tiempo que se proclamaba el Acta Única Europea y poco después del Primer Programa Marco, la implementación de Planes Nacionales, la puesta en marcha progresiva de agencias evaluadoras (ANEP, CNEAI, ANECA), entre otras iniciativas, han ido construyendo el Sistema Español de Ciencia y Tecnología. La evaluación científica constituye en este marco una herramienta esencial para la valoración y la toma de decisiones y, como ya hemos comentado, la evaluación de las revistas científicas constituye uno de sus pilares esenciales. Si bien el balance, pues, es globalmente positivo —especialmente si se compara con la situación de la ciencia durante la dictadura— ha adolecido de una cierta improvisación, inestabilidad, insuficiencia de fondos y poca capacidad crítico-creativa, hechos que han redundado en un seguidismo acrítico y a veces mal entendido de iniciativas surgidas en otros países o en otros foros.

Un sistema de ciencia que se precie debe, no solo evaluar, planificar y tomar decisiones sobre la distribución de los fondos, cada vez más magros, destinados a la investigación, sino que debe también realizar una tarea importante de promoción. Este es, creo, uno de los puntos débiles del Sistema Español de Ciencia y Tecnología que, en lo tocante a la evaluación de las revistas científicas, se manifiesta, entre otros aspectos, en la discontinuidad y desactualización de sus plataformas propias (de hecho, vienen siendo utilizadas últimamente CARHUS, CIRC y, más recientemente, MIAR), reflejo de la extrema falta de sensibilidad hacia los campos de conocimiento relacionados con lo humano y lo social. Ello coincide con una merma en la autoafirmación de las disciplinas con ellos relacionados por parte de los propios profesionales, atrapados en el sistema, con poca capacidad de reacción y focalizados en muchos casos en la lamentación. Si hablamos de ciencia, y no de cientifismo, quizás debamos reivindicarnos recordando las ya lejanas propuestas de Otto Neurath en relación con la prioridad epistemológica de las ciencias sociales y con la necesidad de una unificación de la ciencia inspirada prioritariamente en el tipo de regularidades propias de estas, las generalizaciones débiles (Fernández Buey 1991: 238 y ss.). Muy al contrario, en los últimos años se da la paradoja de que la reivindicación del papel de las ciencias sociales y humanas proviene en muchos casos de profesionales de las ciencias naturales (Gould 2004; Lévy-Leblond 2003).

Núria Rafel Fontanals
Universitat de Lleida
nrafel@historia.udl.cat

Bibliografía

- BORREGO, A. URBANO, C. (2006). La evaluación de revistas científicas en Ciencias Sociales y Humanidades. *Información, cultura y Sociedad*, 14: 11-27.
- BOURE, R. (2015). Les modes d'évaluation de l'enseignement et de la recherche. Épisode 1. *Mondes Sociaux*. Mayo <<http://sms.hypotheses.org/4202>>.
- BUELA-CASAL, G. (2002). La evaluación de la investigación científica: el criterio de la opinión de la mayoría, el factor de impacto, el factor de prestigio y “Los Diez Mandamientos” para incrementar las citas. *Análisis y Modificación de Conducta*, 28: 455-476.
- ECHVERRÍA, J. (2003). *La revolución tecnocientífica*. Fondo de Cultura Económica de España. Madrid.
- FERNÁNDEZ BUEY, F. (1991). *La ilusión del método. Ideas para un racionalismo bien temperado*. Ed. Crítica. Barcelona.
- GOULD, S. J. (2004). *Érase una vez el zorro y el erizo. Las humanidades y la ciencia en el tercer milenio*. Col. Drakontos. Ed. Crítica. Barcelona.
- KRISTIANSEN, K. (2014). Towards a New Paradigm? The Third Science Revolution and its Possible Consequences in Archaeology. *Current Swedish Archaeology*, 22: 11-71.
- LÉVY-LEBLOND, J.-M. (2003). Una cultura sin cultura. Reflexiones críticas sobre la “cultura científica”. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad*, 1, 1: 139-151.
- PÉREZ SEDEÑO, E. (2007). Evaluación, transparencia y democracia. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad*, 3, 8: 173-181.
- ROSSELL, D. (2014). Macrodatos y estadística. La perspectiva de un estadístico. *Mètode. Science Studies Journal*, 83 <<http://metode.cat/es/revista/69>>.
- RUIZ ZAPATERO, G. (2014). Escribir como arqueología, arqueología como escritura. *Anales de la Universidad de Murcia*, 30: 11-28.
- WENNERÅS, C., WOLD, A. (1997). Nepotism and sexism in peer review. *Nature*, 347: 341-343.