

La valoración de la actividad investigadora desarrollada por individuos, grupos y entes es una tarea extremadamente difícil y es por ello que, como en muchos otros ámbitos de la actividad humana, también en este campo se han buscado criterios cuantitativos que proporcionen resultados objetivos al respecto.

El paso final de la investigación científica pública es la presentación de los resultados obtenidos a la sociedad que, por otro lado, es la que suele proporcionar la financiación necesaria. Los métodos para hacer públicos estos resultados son variados, pero la publicación de un artículo en una revista científica, que lo somete al análisis de una comisión de expertos (revisión por pares), se considera como el mejor método para acreditar la validez científica del trabajo realizado. Por ello, la mayoría de los parámetros objetivos usados para valorar la investigación pública se centra sobre los artículos publicados en este tipo de revistas.

El parámetro más sencillo es el número de artículos: quien escribe más artículos es mejor investigador. Producir millares de artículos irrelevantes es mejor que desarrollar una investigación de gran impacto cultural limitada a unas pocas publicaciones. Si este fuera el criterio utilizado, las revistas tendrían interés en publicar la gran mayoría de los artículos que reciben, reduciendo lo más posible el nivel de exigencia de los pares. Pero en este parámetro, la calidad de lo publicado no juega ningún papel.

La búsqueda de un parámetro que tuviera en cuenta la relevancia científica de los artículos llevó a responsabilizar a la propia comunidad científica de la valoración de esta relevancia mediante la cita de la publicación. El número de citas se convirtió en el parámetro objetivo que definía la calidad de una investigación: a mí me citan más que a ti y, por tanto, soy mejor. Sin embargo, es fácil caer en problemas de *popularidad* como los que ocurren con, por ejemplo, los índices de audiencia de los programas de radio y televisión: un artículo puede ser muy citado, independientemente de su calidad. Un ejemplo reciente es el documento de la colaboración OPERA sobre los neutrinos que viajan a una velocidad mayor que la de la luz (<http://arxiv.org/abs/1109.4897>).

Una mejora respecto a este banal conteo de las citas consiste en valorar la calidad de las revistas donde el artículo se ha publicado. Para determinar esa calidad se utiliza el factor de impacto, IF (del inglés, *impact factor*). El cálculo del IF se lleva a cabo considerando el número de las citas anuales de la revista y dividiéndolo por el número total de artículos publicados. Una revista es tanto más prestigiosa cuanto más alto es este número. Publicar pocos artículos muy citados es mejor que publicar muchos pero menos citados.

Se pueden hacer muchas críticas al IF. Quiero mencionar aquí las dos que en mi opinión son más relevantes. En primer lugar, se trata de un índice dependiente de la disciplina científica. Es evidente que el número de científicos que componen una comunidad dada, que se autocita por definición, influye sobre el valor del IF. Por ejemplo, una comunidad de tres personas puede producir como máximo seis citas, mientras que una de cuatro investigadores, veinticuatro. Además, los tiempos de publicación y de desarrollo de la investigación son típicos de cada disciplina y, por

tanto, también lo son las respuestas de confirmación o falsación de un nuevo artículo. Por ejemplo, la física teórica tiene tiempos de respuesta mucho más rápidos que los de la física experimental de altas energías; ésta se organiza en grupos de centenares de personas, mientras que los físicos teóricos trabajan solos o en grupos reducidos.

En segundo lugar, el IF es un valor medio de una distribución que no es normal: un único artículo muy citado desplaza significativamente el valor medio. Esto permite *trampear* sin muchas dificultades con el IF: en 2007 una revista científica que polemizaba sobre el (ab)uso del mismo escribió un editorial en el que autocitaba todos los artículos publicados por la propia revista del 2005 al 2006 provocando un aumento de su propio IF de 0.66 a 1.44.

Por lo que conozco, la última invención en términos de parámetros objetivos es el índice h , cuyo nombre se debe al apellido de su inventor, Jorge E. Hirsch, un físico de la universidad de California. Más allá de la complejidad de la formulación original, el método consiste en ordenar las publicaciones de una persona en sentido decreciente del número de citas. El índice h viene dado por la posición del artículo que coincide con su número de citas.

Este índice resuelve el problema de sobrevaloración de un currículum con un gran número de publicaciones irrelevantes, pero adolece de algunos aspectos negativos. Por ejemplo, h no tiene en cuenta el número de autores de un artículo. El comentario de más arriba acerca de físicos teóricos y experimentales pone en evidencia la dificultad de utilizar este parámetro en contextos amplios. Además, y en cualquier caso, este índice está ligado al número total de publicaciones. Un científico tan relevante por el impacto de sus obras como Galois tendría un $h=2$ y en 1906, después de haber publicado sus tres famosos artículos sobre el movimiento browniano, el efecto fotoeléctrico y la relatividad restringida, Einstein habría tenido un h de 4 o 5, insuficiente para ser admitido a un concurso de investigador del nivel más bajo en Italia.

Entre las distintas tareas que mi estado académico me impone, la que más detesto es formar parte de comisiones de concursos para seleccionar nuevos profesores o investigadores para la universidad. Junto a otros colegas debo confrontar personas que trabajan en contextos diferentes y juzgar cuál es la más adecuada para las tareas objeto del concurso. Utilizar parámetros objetivos sería un alivio y evitaría que la comisión asumiera la responsabilidad de la elección. Pero si tales criterios objetivos existieran, no sería necesaria una comisión: un simple ordenador bastaría para seleccionar a la persona más digna.

En Italia, el Ministerio de la Universidad y de la Investigación indica a las comisiones de concurso criterios objetivos a los que atenerse para expresar sus valoraciones. Un colega ha sintetizado este proceso diciendo que seremos forzados a juzgar currícula y no personas. Pero como suele ocurrir, las acciones preparadas por la burocracia centralizada no resuelven los problemas y, en compensación, crean otros.

El análisis cuantitativo es muy útil en algunos ámbitos del conocimiento, pero no puede ser generalizado. Un análisis cualitativo es esencial para poder estudiar estructuras complejas y no homogéneas. El análisis cualitativo es subjetivo y los sujetos en los que se delega el ejercicio de este análisis deben asumir las correspondientes responsabilidades. ¿Qué pensáis acerca de la hipótesis de seleccionar a los jugadores del equipo nacional de fútbol sólo mediante datos estadísticos (goles marcados, pases realizados, balones recibidos, paradas realizadas respecto al número de goles recibidos, etc.)? ¿Objetivo! ¿no?