



EDITORIAL semFYC

Estar o no estar en el asunto: la evaluación individual del rendimiento científico

To be or not to be in the picture: the individual assessment of scientific performance

A lo largo de su historia, la ciencia ha sufrido una enorme transformación, pasando de una fase caracterizada por aportaciones individuales de autores singulares, como Euclides (300 a.C.), Galileo (1564–1642), Isaac Newton (1643–1727), Charles Darwin (1809–1882) o Albert Einstein (1879–1955), a otra definida por la actividad colectiva a pequeña y gran escala —tómense los proyectos «Genoma humano» o el sincrotrón como ejemplos recientes—. Este espectacular cambio se produjo a caballo entre las guerras mundiales y se conoce como «*from little science to big science*»¹. Si bien los premios Nobel² dan continuidad a la idea de individualidad premiando a investigadores concretos, en la actualidad la ciencia es una actividad predominantemente colectiva³.

Una de las consecuencias del paso a la *big science* ha sido el crecimiento a escala mundial de la comunidad científica, y con éste la aparición de lo que en bibliometría se conoce como homónimos, es decir, distintos autores que publican usando la misma firma bibliográfica. Ya en 1974 E. Garfield⁴, creador del Institute for Scientific Information (ISI), se refería al problema que supone identificar autores que usan la misma firma, sobre todo si trabajan en el mismo centro o, peor aún, si se dedican a la misma disciplina científica. Posteriormente, y desde su perspectiva de compilador de bases de datos, ese autor describe monográficamente las diferencias entre las prácticas de firma de autores de los cinco continentes, puntualizando que éstas pueden además cambiar si el autor publica dentro o fuera de su país de origen⁵. Según Garfield, a esto se suman otros factores como la falta de regularidad en la firma, tanto por parte del propio autor como por la del citante, y la presencia de artículos y preposiciones en las firmas. Con todo, admite que en muchas ocasiones no es posible distinguir el nombre del apellido de un autor, lo cual se ve agravado por la creciente presencia de variantes de las firmas de los autores, lo que se conoce como sinónimos. Como se puede ver, los homónimos y

sinónimos son el principal obstáculo para identificar a un autor a través de sus publicaciones y, lo que es peor, no se trata de un fenómeno específico de un país en particular, sino de un fenómeno mundial.

¿Pero en qué situaciones es necesario evaluar de forma individual el rendimiento de un investigador? En general se podrían describir tres: en convocatorias competitivas donde la solvencia científica del investigador principal tiene un gran peso; en la creación de paneles de expertos, y en la promoción dentro de la institución. Estas situaciones, esenciales en la carrera de todo investigador, dependen del rigor de la evaluación, la cual a su vez depende del grado de complejidad que encierre establecer una correspondencia entre las firmas bibliográficas y el autor en cuestión. Una incorrecta evaluación, ya sea por inclusión u omisión de publicaciones, por lo general va en detrimento del autor objeto de la evaluación.

¿Cómo puede facilitarse la evaluación individual? Teniendo en cuenta las fuentes de información es posible actuar en tres frentes. Si entramos en un proceso de evaluación y se nos requieren las publicaciones, nuestro currículum vitae (CV) es la primera opción. Entre sus ventajas se cuentan su exhaustividad y el hecho de que es posible presentarlo en diferentes formatos: ordenando las publicaciones más recientes primero, y diferenciando los artículos y revisiones indexados del resto de los documentos, por ejemplo. La principal limitación del CV es que no tiene un vínculo con las bases de datos capaces de proveer el número de citas recibidas, y/o algún indicador de la revista de publicación. Establecer dicho vínculo para cada publicación es poco factible. Así pues, la utilidad del CV en evaluaciones es limitada. Con el objetivo de superar esta limitación, la Fundación Española para la Ciencia y la Tecnología (FECYT) está desarrollando el CV Normalizado (CVN)⁶, una aplicación con acceso web que, entre otras características, permitirá enlazar las publicaciones registradas por los autores con las

bases de datos bibliográficas, de forma que se puedan elaborar indicadores bibliométricos que complementen su evaluación. Si bien el CVN supone un gran avance, la información bibliográfica continuará proviniendo de las grandes bases de datos generalistas, lo que nos lleva al siguiente frente de acción. Supongamos que un organismo decide crear un panel de expertos y, por lo tanto, a priori no sabe qué personas lo compondrán. Una de las opciones será realizar una búsqueda temática en una o varias bases de datos para recopilar las publicaciones pertinentes y, a partir de éstas, identificar a los autores más activos y/o más visibles. En este caso los potenciales integrantes del panel desconocen el proceso de evaluación, por lo que no podrán intervenir aportando sus publicaciones, y su inclusión o no dependerá únicamente del rigor en el proceso de selección y la experiencia de los seleccionadores. En vista de que al final siempre se recurre a las bases de datos generalistas, la opción más razonable es evitar la utilización de homónimos y minimizar la probabilidad de que se generen sinónimos de nuestra firma, de forma que sea más fácil identificarnos a través de las publicaciones.

Para evitar utilizar firmas homónimas tan sólo basta con buscar la firma en cuestión en PubMed⁷, Web of Science (WOS)⁸, Google Scholar⁹ o SCOPUS¹⁰, y podremos evaluar su singularidad. La probabilidad de ser objeto de una evaluación incorrecta aumenta drásticamente cuanto mayor es el número de autores que comparten una firma. Este número se puede aproximar por diferencias en el centro de filiación,

la disciplina (revistas de publicación) o porque publican con distintos subgrupos de coautores¹¹. El número de homónimos depende de la distribución de apellidos en el país de que se trate¹². Según datos del INE¹³, una cuarta parte de la población española, unos 11,5 millones de personas en números redondos, comparten tan sólo 20 apellidos, de los cuales García y González son los más frecuentes (tabla 1). Si reproducimos este análisis con los 20 nombres de pila más frecuentes, el resultado es que una tercera parte de la población, unos 15 millones de habitantes, comparten tan sólo 20 nombres distintos, que además se pueden resumir en 13 iniciales, lo que indica que a cada uno de estos nombres corresponde más de 1 millón de personas, el doble que la proporción de habitantes por apellido (tabla 2). Si extrapolamos estos resultados a la comunidad científica española, los nombres de cuyos integrantes están inevitablemente afectados por lo expuesto anteriormente, se puede ver que la estrategia de elegir una firma bibliográfica construida únicamente a partir del primer apellido y la inicial del nombre de pila resulta, a todas vistas, inapropiada para identificar, ni siquiera con las mínimas garantías, a un determinado autor. Los artículos y resúmenes publicados desde España durante 2006–2008 disponibles en la WOS así lo confirman. Por un lado, las firmas de los 20 autores más productivos tienen el formato de primer apellido e inicial del nombre, pudiéndose resumir el listado en tan sólo 9 apellidos y 5 iniciales distintas; por el otro, el elevado número de documentos, 306 para la primera firma de la

Tabla 1 Distribución de la población según primer apellido, 2007^a

Ordinal	Primer apellido	% ^b	Acum ^c	Población ^d
1	García	3,32		
2	González	2,08	5,4	2.441.911
3	Fernández	2,08	7,48	3.381.787
4	Rodríguez	2,07	9,55	4.316.863
5	López	1,96	11,51	5.203.224
6	Martínez	1,87	13,39	6.050.249
7	Sánchez	1,83	15,21	6.876.440
8	Pérez	1,75	16,97	7.668.579
9	Martín	1,11	18,07	8.168.451
10	Gómez	1,1	19,17	8.666.005
11	Jiménez	0,86	20,03	9.054.217
12	Ruiz	0,82	20,85	9.426.048
13	Hernández	0,79	21,64	9.783.579
14	Díaz	0,75	22,4	10.123.503
15	Moreno	0,7	23,1	10.441.870
16	Álvarez	0,64	23,74	10.731.947
17	Muñoz	0,62	24,37	11.013.972
18	Romero	0,48	24,85	11.232.339
19	Alonso	0,45	25,3	11.437.216
20	Gutiérrez	0,43	25,74	11.632.527

^aSolamente se muestran los primeros 20 apellidos.

^bTanto por ciento de la población que comparte este primer apellido.

^cTanto por ciento acumulado de población.

^dPoblación en valor absoluto calculado sobre 45.200.737 habitantes, según datos del Padrón Municipal 2007. Fuente: Instituto Nacional de Estadística.

Tabla 2 Distribución de la población según nombre de pila, 2007^a

Ordinal	Nombre de pila	% ^b	Acum ^c	Población ^d
1	Antonio	3,8		
2	Jose	3,6	7,4	3.361.491
3	Manuel	3,2	10,7	4.826.572
4	Francisco	2,9	13,6	6.127.317
5	Juan	2	15,5	7.022.939
6	David	1,5	17	7.683.410
7	José Antonio	1,5	18,5	8.342.754
8	José Luis	1,4	19,9	8.989.701
9	Jesús	1,4	21,3	9.618.966
10	Javier	1,3	22,5	10.191.457
11	Carlos	1,2	23,8	10.751.280
12	Miguel	1,2	25	11.307.264
13	Pedro	1,2	26,2	11.859.549
14	Rafael	1,2	27,5	12.408.000
15	José Manuel	1,1	28,6	12.915.633
16	Ángel	1,1	29,7	13.418.989
17	Daniel	1,1	30,8	13.900.816
18	Francisco Javier	1,1	31,8	14.380.068
19	Luis	1	32,9	14.849.755
20	Fernando	1	33,9	15.310.531

^aSolamente se muestran los primeros 20 apellidos.

^bTanto por ciento de la población que comparte este primer nombre de pila.

^cTanto por ciento acumulado de población.

^dPoblación en valor absoluto calculado sobre 45.200.737 habitantes, según datos del Padrón Municipal 2007. Fuente: Instituto Nacional de Estadística.

Tabla 3 Autores más productivos en el periodo 2006–2008^a

Firma bibliográfica	nDocs ^b
Rodríguez, A	306
Martínez, A	268
Sánchez, A	256
Fernández, A	216
González, A	215
García, A	207
García, J	204
Fernández, E	189
Martin, J	178
González, J	177
González, M	174
Muñoz, A	174
García, C	164
Rodríguez, J	163
Martínez, C	161
Martin, A	158
Fernández, J	150
Martínez, E	149
Moreno, A	149
Fernández, M	148

^aPrimeros 20 resultados de una búsqueda según país en el campo «address» en la Web of Science el 4 de agosto de 2008.

^bNúmero de documentos citables (artículos, revisiones y *proceedings*).

lista, es un claro indicio de que estamos frente a un homónimo; en caso contrario estaríamos hablando de un autor que es capaz de publicar en promedio 100 documentos al año o un documento cada 3 días (tabla 3). En resumen, para evitar utilizar una firma homónima, tan sólo es necesario hacer una búsqueda bibliográfica en una base de datos de cobertura mundial y comprobar que no exista. En caso de que por cualquier situación la creatividad esté bajo mínimos, invito al lector a visitar la página de IRALIS¹⁴, en la que se pueden obtener algunas sugerencias de firmas bibliográficas.

El tratamiento de las variantes o firmas sinónimas es más complejo, por lo que empezaremos por revisar qué y dónde se generan. El primer y principal agente generador de sinónimos es el propio autor. La solución en este caso pasa por elegir una firma, lo más singular posible, y mantenerla a lo largo de toda la carrera científica. También es cierto que en algunos casos, como por ejemplo en las estancias cortas, nuestros colegas nos incluyen entre los autores de publicaciones posteriores a nuestra estancia, a veces sin ni siquiera informarnos, de forma que no podemos controlar que nuestra firma aparezca correctamente. Otra fuente nada despreciable de variantes son los compiladores de bases de datos bibliográficas. Por un lado, las hoy anecdóticas «mutaciones» de letras causadas por errores en el reconocimiento automático de texto, gracias a la transmisión electrónica de datos, y por otro, y más importante, la aplicación por igual del patrón de firma anglosajón —apellido e inicial del nombre de pila— que

crea artefactos difíciles de reconocer, incluso para el propio autor. Son especialmente sensibles a esta manipulación las firmas que incluyen nombres de pila compuestos, artículos y/o preposiciones o simplemente los dos apellidos. A este respecto, en 2006 un grupo de bibliómetras de España consensuó un documento de recomendaciones dirigidas a minimizar y contrarrestar el efecto de la aplicación del mencionado patrón de firma anglosajón¹⁵. En general se recomienda utilizar guiones para unir las diferentes partículas que forman la firma. Así, por ejemplo, las firmas que incluyan artículos y preposiciones adoptarían un aspecto similar a éstas: de-los-Cobos, A; del-Río, J; aquellas que incluyeran dos o más apellidos quedarían como: González-García, J; Morales-Suárez-Varela, A; etc. Asimismo, en mi experiencia^{16–18}, he podido observar que los nombres compuestos pueden ser otra fuente de variantes. A diferencia de lo que ocurre habitualmente con nombres como María del Carmen, que coloquialmente pueden ser utilizados indistintamente como Maricarmen o simplemente Carmen, es altamente recomendable no reproducir estas variantes en la firma bibliográfica. Los cambios de idioma pueden igualmente dar pie a la creación de variantes muy difíciles de identificar, piénsese en la metamorfosis que supone el paso de la letra J a X en nombres como Javier o José en su paso a catalán o gallego respectivamente, o la de la letra C a K en el caso de Carlos a euskera. En definitiva, los autores y las manipulaciones de los compiladores de bases de datos son las principales fuentes de variantes de firmas o sinónimos.

Bien, ya conocemos la enfermedad y además tenemos elementos para conjurar un remedio, ¿pero qué pasa con las publicaciones que ya están recopiladas en las bases de datos o registro histórico? Ni Thomson-ISI ni SCOPUS, únicos proveedores generalistas de bases de datos bibliográficas, normalizan de forma sistemática la información referente a centros o firmas contenidas en las publicaciones, si bien han lanzado herramientas como Research ID¹⁹ o Author Identifier²⁰ con el objetivo de paliar el déficit de información normalizada en cuanto a la producción de autores individuales. Precisamente, el gran problema de la evaluación bibliométrica individual tiene su raíz en la ausencia de correspondencia entre firmas bibliográficas y autores. En la actualidad dicha correspondencia se obtiene con garantías suficientes únicamente a través de procesos manuales y, en el mejor de los casos, semiautomáticos²¹ aunque, por lo general y debido su complejidad, estos procesos suelen afectar a reducidos volúmenes de datos. Determinar las correspondencias entre firmas y autores del registro histórico pasaría por disponer de un sistema diseñado con objetivos evaluativos, como por ejemplo el CVN⁶, que estando en conexión con bases de datos bibliográficas nos permitiera registrar nuestras publicaciones bajo un único e inequívoco identificador y, además, permitiera elaborar diferentes indicadores. Sería igualmente deseable que este sistema recopilara por mecanismos automáticos las nuevas publicaciones y, previa validación por nuestra parte, actualizara nuestro registro personal. Y así sucesivamente... En la actualidad las tecnologías de la información hacen posibles herramientas similares, por lo que ésta es la parte fácil. El verdadero reto es conseguir la aceptación y el uso generalizado y asiduo de una herramienta de estas características por la comunidad científica.

A falta de un sistema de registro y evaluación de amplia aceptación, invito al lector a que incluya en su CV, bajo un epígrafe que podría ser «Ayuda o guía para la evaluación», su particular colección de sinónimos y coautores más duraderos buscando sus publicaciones en PubMed, WOS o SCOPUS. Además de constatar lo difícil que resulta identificar incluso las propias publicaciones en estas bases de datos y descubrir las variantes de firmas creadas por sus compiladores, contribuirá a que sus evaluaciones sean más justas.

En definitiva, las evaluaciones individuales con frecuencia son determinantes en la carrera de todo investigador, aunque la existencia de homónimos y sinónimos puede reducir su validez, ya que falsean la correspondencia entre firmas y autores. Frente a esta situación, y mientras no dispongamos de un sistema de registro y evaluación de amplia aceptación, es imperativo elegir una firma lo más singular posible y mantenerla a lo largo de nuestra carrera profesional.

Bibliografía

- Price D de Solla. Little Science, Big Science. New York: Columbia University Press; 1963.
- Premios Nobel [citado Ago 2008]. Disponible en: <http://nobelprize.org>.
- Wuchty S, Jones FJ, Uzzi B. The increasing dominance of teams in production of knowledge. Science. 2007;316:1036–9.
- Garfield E. Errors-theirs, ours and yours. Current Contents. 1974;25:5–6.
- Garfield E. What's in a surname?. Current Contents. 1981;7:5–9.
- FECYT. Curriculum Vitae Normalizado (CVN) [citado Ago 2008]. Disponible en: <http://cv.normalizado.org>.
- PubMed [citado Ago 2008]. Disponible en: www.pubmed.gov.
- Web of Knowledge (WOK) [citado Ago 2008]. Disponible en: <http://www.accesowok.fecyt.es>.
- Google Scholar o académico. [citado Ago 2008]. Disponible en: <http://scholar.google.es>.
- SCOPUS. Disponible en: www.scopus.com.
- Soler JM. Separating the articles of authors with the same name. Scientometrics. 2007;72:281–90.
- Aksnes DW. When different persons have an identical author name. How frequent are homonyms?. Journal of the American Society for Information Science and Technology (JASIST). 2008;59:838–41.
- Instituto Nacional de Estadística (INE) [citado Ago 2008]. Disponible en: <http://www.ine.es>.
- IRALIS [citado Ago 2008]. Disponible en: <http://www.iralis.org>.
- Recomendaciones sobre firmas bibliográficas [citado Ago 2008]. Disponible en: http://www.accesowok.fecyt.es/recomendaciones_publicaciones.html.
- Camí J, Suñén E, Méndez-Vásquez RI. Caracterización bibliométrica de grupos de investigación biomédica en España (1994–2000) [citado Ago 2008]. Barcelona; 2003. Disponible en: <http://84.88.71.251/webs/mapabiomedicogrupos/index.htm>.
- Méndez-Vásquez RI, Suñén-Pinyol E, Sanz G, Camí J. Caracterización bibliométrica de los grupos de investigación en temática cardio-cerebrovascular. España 1996–2004. Barcelona; 2007 [citado Ago 2008]. Disponible en: <http://bibliometria.prbb.org/GruposCardiocerebrovascular>.
- Méndez-Vásquez RI, Suñén-Pinyol E, Torrens M, Castro-Fornieles J, Camí J. Caracterización bibliométrica de los grupos de investigación en psiquiatría, psicología clínica y drogodependencias. España 1996–2004. Barcelona; 2007 [citado Ago 2008]. Disponible en: <http://bibliometria.prbb.org/psiquiatria>.
- Researcher ID es un recurso de la empresa Thomson-ISI [citado Ago 2008]. Disponible en: <http://www.researcherid.com/Home.action>.
- Author Identifier es un recurso accesible a través de suscripción en SCOPUS [nota de prensa]. Disponible en: http://info.scopus.com/news/press/pr_060613.asp.
- Torvik VI, Weeber M. A probabilistic similarity metric for medline records: A model for author name disambiguation. Journal of the American Society for Information Science and Technology (JASIST). 2005;56:140–58.

Raúl Isaac Méndez Vásquez
 Grup de Recerca en Bibliometria i Avaluació en
 Ciència (BAC), Agència d'Avaluació de
 Tecnologia i Recerca Mèdiques (AATRM),
 Barcelona, España
 Correo electrónico: rmendez@prbb.org