

CONSIDERACIONES METODOLÓGICAS SOBRE USO DEL IMPACTO NORMALIZADO EN CONVOCATORIAS SEVERO OCHOA Y MARÍA DE MAEZTU

Methodological considerations on the use of the
normalized impact in the *Severo Ochoa* and *María de
Maeztu* programmes

Daniel Torres-Salinas, Nicolás Robinson-García, Enrique Herrera-Viedma y
Evaristo Jiménez-Contreras



Daniel Torres-Salinas, doctor en documentación científica por la *Universidad de Granada*, trabaja en dicha *Universidad* y en la *Universidad de Navarra*. Es investigador en el grupo de investigación *Evaluación de la Ciencia y la Comunicación Científica (EC3)* sobre temas bibliométricos. Es CEO de la *spin-off* de asesoría científica *EC3metrics* y responsable de la línea de *Ciencia Digital* del *Medialab UGR*. Es co-creador de sistemas de evaluación de la investigación como *Cientificacvn*, *Rankings I-UGR de Universidades*, *Clasificación integrada de revistas científicas (CIRC)*, *Bipublishers* o *UGRinvestiga*. Tiene indexadas más de 55 publicaciones científicas en la base de datos *Web of Science (Clarivate Analytics)*. Imparte cursos sobre comunicación científica, web 2.0 y ciencia y servicios de investigación en bibliotecas.

<https://orcid.org/0000-0001-8790-3314>

Universidad de Granada y Universidad de Navarra, Centro de Transferencia Tecnológica
Gran Vía, 48. 18010 Granada, España
torressalinas@gmail.com



Nicolás Robinson-García investigador postdoctoral en la *School of Public Policy* del *Georgia Institute of Technology*, es doctor en Ciencias Sociales por la *Universidad de Granada*. Anteriormente disfrutó de un contrato postdoctoral *Juan de la Cierva-Formación* en *Ingenio (CSIC-Universidad Politécnica de Valencia)* bajo la supervisión de Ismael Ràfols. Actualmente está interesado en el desarrollo de métodos cuantitativos para estudiar la movilidad científica así como en el análisis de nuevas fuentes de datos y métodos para rastrear el compromiso social de los investigadores.

<http://orcid.org/0000-0002-0585-7359>

Georgia Institute of Technology, School of Public Policy
North Avenue, Atlanta, Estados Unidos
elrobinster@gmail.com



Enrique Herrera-Viedma es profesor de Ciencias de la Computación e Inteligencia Artificial en la *Universidad de Granada* y vicerrector de *Investigación y Transferencia*. Sus líneas de investigación relacionadas con Biblioteconomía y Documentación se centran en recuperación de información, bibliometría, bibliotecas digitales, evaluación de webs, sistemas de recomendación y medios sociales. Ha publicado más de 200 artículos en revistas *JCR* y coordinado más de 20 proyectos de investigación. Es miembro del comité ejecutivo de la *IEEE SMC Society* y editor asociado de las revistas *IEEE Transactions on systems, man, and cybernetics*, *Knowledge based systems*, *Soft computing*, *Fuzzy optimization and decision making*, *Applied soft computing*, *Journal of intelligent systems* y *Fuzzy systems, and information sciences*. Su índice h es de 62 y figura en los listados de investigadores muy citados de *Clarivate Analytics*.

<https://orcid.org/0000-0002-7922-4984>

Universidad de Granada, Departamento de Ciencias de la Computación e Inteligencia Artificial
Centro de Transferencia Tecnológica
Gran Vía, 48. 18010 Granada, España
viedma@decsai.ugr.es



Evaristo Jiménez-Contreras es catedrático de Bibliometría en la *Facultad de Comunicación y Documentación* de la *Universidad de Granada*, y director del grupo de investigación *EC3 (Evaluación de la Ciencia y de la Comunicación Científica)*. Sus líneas de investigación se centran en la evaluación de la actividad científica con metodología bibliométrica. Es uno de los promotores de los índices de impacto de las revistas españolas de ciencias sociales, jurídicas y humanas (*In-Recs, In-Recj, In-Rech*), *CIRC (Clasificación integrada de revistas científicas)*, *Metaranking EC3*, y *Científica*.
<http://orcid.org/0000-0001-5668-7057>

*Universidad de Granada, Departamento de Información y Comunicación
Campus Cartuja, s/n. 18071. Granada, España
evaristo@ugr.es*

Resumen

En 2011 se lanzaron por primera vez las convocatorias *Apoyo y acreditación de Centros de Excelencia Severo Ochoa y a Unidades de Excelencia María de Maeztu*. Desde entonces estas acreditaciones se han convertido en uno de los ejes de la política científica española, distribuyéndose 186 millones de € y acreditando a 26 centros y 16 unidades. A nivel bibliométrico uno de los criterios de evaluación más llamativos es la necesidad de que los investigadores garantes tengan un Impacto Normalizado de, al menos, 1,5. En este trabajo analizamos críticamente el origen, en los años ochenta, de este indicador bibliométrico normalizado, las variantes que se han propuesto y las limitaciones de su uso en esta convocatoria nacional. Finalmente se ofrece una serie de recomendaciones prácticas para un uso más adecuado y preciso del indicador de Impacto Normalizado con fines evaluativos.

Palabras Clave

Indicadores bibliométricos; Métodos bibliométricos; Citación normalizada; Impacto normalizado; Evaluación; Indicador Crown; Política científica; Centros de excelencia.

Abstract

In 2011, the research programmes for *Support and accreditation of Severo Ochoa Centers of Excellence and Units of Excellence María de Maeztu* were launched for the first time. Since then, these accreditations have become one of the axes of the Spanish scientific policy. In these years, € 186 million have been distributed and 26 centers and 16 units have been accredited. At the bibliometric level, one of the most relevant evaluation criteria is the need for guarantor researchers to have a Normalized Impact of, at least, 1.5. In this work, we critically analyze the origin of this bibliometric indicator in the eighties, the variants that have been proposed and the limitations of its use in this national call. Finally we offer a series of practical recommendations for a more accurate use of the Normalized Impact indicator for evaluative purposes.

Keywords

Bibliometric indicators; Bibliometric methods; Normalized citation; Normalized impact; Crown indicator; Research policy; Research evaluation; Assessment; Research performance; Centers of excellence.

Torres-Salinas, Daniel; Robinson-García, Nicolás; Herrera-Viedma, Enrique; Jiménez-Contreras, Evaristo (2018). "Consideraciones metodológicas sobre uso del impacto normalizado en convocatorias Severo Ochoa y María de Maeztu". *El profesional de la información*, v. 27, n. 2, pp. 367-374.

<https://doi.org/10.3145/epi.2018.mar.15>

1. Introducción

Las convocatorias *Apoyo y acreditación de Centros de Excelencia Severo Ochoa y Unidades de Excelencia María de Maeztu* forman parte del *Plan Estatal de Investigación Científica y Técnica y de Innovación*. Su objetivo es financiar instituciones (centros o unidades) que realizan investigación de gran impacto, con liderazgo a nivel internacional e impacto económico/social en su entorno. La primera convocatoria *Severo Ochoa (SO en adelante)* tuvo lugar en 2011 y tres años más tarde, en 2014, se complementó con la *María de Maeztu (MM en adelante)*. La diferencia entre ambas es el tamaño de los centros a acreditar y la financiación que se obtiene: en la *SO* se consiguen cuatro millones repartidos en cuatro años, mientras que en la *MM* la financiación se reduce a dos millones. En las dos convocatorias las ayudas son renovables a la finalización del período. Hasta el momento se ha presentado

un total de 247 solicitudes y el pasado 21 de enero de 2018 concluyó el plazo de la última convocatoria.

En los seis años de funcionamiento se han acreditado 26 centros y 16 unidades (véase dataset disponible en *Zenodo*) (tabla 1).

<https://zenodo.org/record/1161753#.WqT-jZPOUUE>

La financiación global del programa ha sido de 184 M€, incluyendo doce renovaciones de los años 2015 y 2016. La mayor parte recae en las acreditaciones *Severo Ochoa* que acumula el 82% de la financiación.

En la distribución por comunidades, Cataluña cuenta con 19 acreditaciones (45%) y 84 M€ (45%), Madrid 11 acreditaciones y 52 M€, País Vasco y Comunidad Valenciana tienen 4; el resto de comunidades obtienen una porción testimonial, y la mayor parte, 11, no ha obtenido financiación.

Tabla 1. Distribución del número de solicitudes concedidas por comunidad autónoma en las convocatorias de centros de excelencia *Severo Ochoa* (SO) y *María de Maetzu* (MM)

	2011		2012		2013		2014		2015		2016		Total
	SO	MM	SO	MM	SO	MM	SO	MM	SO	MM	SO	MM	
Cataluña	4		2		2		1	4	2	2		2	19
Comunidad de Madrid	3		1		1			2		1	2	1	11
Comunidad Valenciana			1		1		1			1			4
País Vasco					1				1		1	1	4
Andalucía			1									1	2
Galicia												1	1
Islas Canarias	1												1
Resto comunidades (11)													0
Total	8		5		5		2	6	3	4	3	6	42

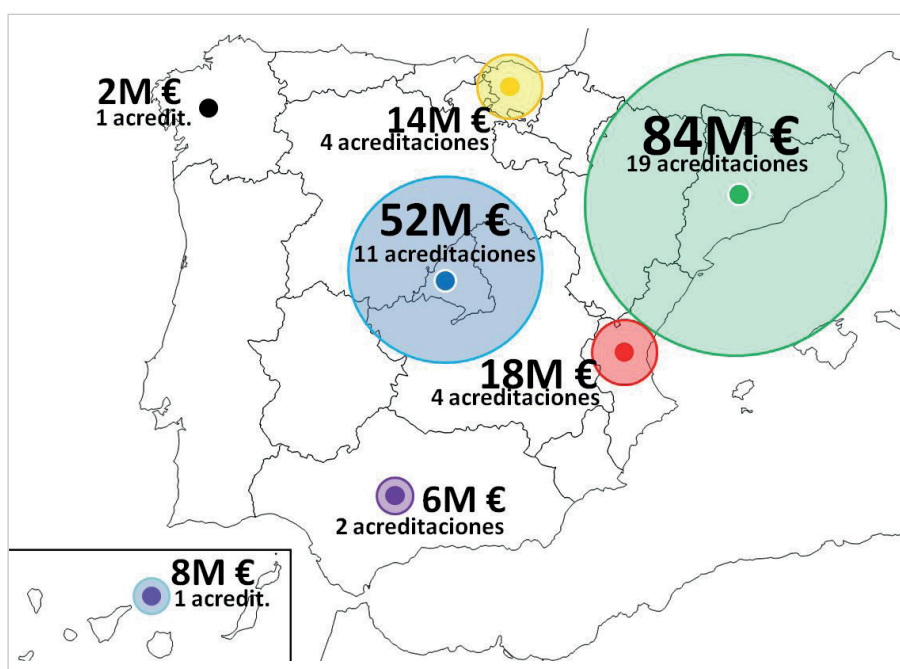


Figura 1. Distribución de la financiación por comunidad autónoma en las convocatorias de centros de excelencia *Severo Ochoa* (SO) y *María de Maetzu* (MM)

Tabla 2. Distribución de la financiación por comunidad autónoma en las convocatorias de centros de excelencia *Severo Ochoa* (SO) y *María de Maetzu* (MM). Con renovaciones. En millones de €

	2011		2012		2013		2014		2015		2016		Total
	SO	MM	SO	MM	SO	MM	SO	MM	SO	MM	SO	MM	
Cataluña	32		16		8		4	8	8	4		4	84
Comunidad Madrid	24		8		4			4		2	8	2	52
Comunidad Valenciana			8		4		4			2			18
País Vasco					4				4		4	2	14
Andalucía			4									2	6
Galicia												2	2
Islas Canarias	8												8
Resto comunidades (11)													0
Total	64		36		20		8	12	12	8	12	12	184

Por áreas temáticas destacan las matemáticas, ciencias experimentales e ingeniería (54%) y dentro de éstas la física. En el área de ciencias de la vida, el 33% de las acreditaciones son de biomedicina. La representación global de las ciencias sociales es mínima con sólo cuatro acreditaciones, todas en economía; y las humanidades no obtienen ningún centro. Es por tanto una convocatoria con enormes sesgos a nivel territorial y temático.

En la documentación que se solicita, el aspecto que mayor interés ha suscitado han sido los criterios de evaluación y su dependencia en indicadores bibliométricos. Entre los requisitos de obligado cumplimiento se exigen 10 investigadores garantes, 6 en el caso de las unidades MM, que tengan un proyecto de investigación activo y superen el 1,5 de impacto normalizado.

Es importante subrayar que por primera vez en una convocatoria nacional se hace uso de una medida de citación normalizada. Sin embargo, el indicador se ha adoptado sin un proceso de reflexión previo como muestra su aplicación a nivel micro, la utilización de una variante de su cálculo con notables limitaciones y sin aval por parte de la comunidad científica.

Considerando la gran carga bibliométrica de las solicitudes y la importancia que estas convocatorias están tomando como uno de los ejes de la política científica española, en este trabajo analizaremos el origen y evolución del indicador de impacto normalizado, las diferentes formas de cálculo y sus limitaciones, y finalmente ofreceremos una serie de recomendaciones para un uso más adecuado del mismo.

2. Origen y limitaciones de los indicadores normalizados

El indicador impacto medio esperado que se emplea en la convocatoria se basa en la agregación de promedios de citas por años y disciplina de los trabajos publicados por los investigadores garantes en los últimos cuatro años comparados con la media mundial. El indicador pretende corregir sesgos en la citación causados por la edad de los documentos y las diferencias en los hábitos de citación de cada disciplina científica. Su uso es sin duda más correcto que otros empleados en convocatorias públicas como los cuartiles y el Factor de impacto (**Cabezas-Clavijo; Torres-Salinas, 2015**). Sin entrar todavía en el ámbito de aplicación, se trata de un indicador ampliamente discutido en la bibliografía científica que parte de una larga tradición bibliométrica (**Waltman et al., 2011a**). La idea de base de los indicadores de citación normalizada es obtener un valor que permita realizar

comparativas entre individuos, instituciones o áreas de investigación. Para ello se definen dos tipos de impacto:

- impacto esperado: número de citas promedio que recibieron todos los trabajos del mundo publicados en el mismo año y en el mismo campo temático que el trabajo analizado;
- impacto observado: número de citas que el trabajo que se trata de comparar ha recibido hasta el momento.

De la división entre el impacto observado y el esperado se obtiene el impacto normalizado de un trabajo científico, cuyo valor será menor a 1 cuando se encuentre por debajo de la media mundial, igual a 1 cuando esté en la media mundial, o superior a 1 cuando sea mayor que la media.

Uno de los principales problemas surge a la hora de agregar documentos para obtener las líneas base (*baselines*) sobre las que se establece la comparativa, ya que implica tener acceso completo a bases de datos bibliométricas como *Web of Science* o *Scopus*. Esto explica que sólo centros como el *CWTS (Centre for Science and Technology Studies, Leiden University)* pudieran calcular dichas líneas y el indicador a partir de los años noventa. Sin embargo las primeras tentativas para crear un indicador normalizado datan de los ochenta cuando **Schubert, Glänzel y Braun (1983)** y **Vinkler (1986)** lanzaron sus primeras propuestas. Al carecer de *baselines* mundiales de citación sugerían utilizar el Factor de impacto como medida de impacto esperado. Por ello se considera que fueron los investigadores del *CWTS (De-Bruin et al., 1993; Moed; De-Bruin; Van-Leeuwen, 1995)* los primeros en sugerir un indicador de impacto normalizado denominado CPP/FCSM (citas por publicación entre promedio de citas en el área temática).

El CPP/FCSM o *Indicador Crown*, como fue renombrado por **Van-Raan (1999)**, se define del siguiente modo: dado un conjunto de artículos científicos, se calcula para cada trabajo su impacto observado y se establece por otro lado su

The screenshot shows the website 'Ciencia e Innovación' with a navigation menu and a main content area. The main content area features a title 'Apoyo y acreditación de Centros de Excelencia «Severo Ochoa» y a Unidades de Excelencia «María de Maeztu»' and a detailed description of the program. The page also includes a sidebar with 'Investigación' and 'Enlaces' sections.

<https://goo.gl/6q7THg>

Tabla 3. Ejemplo aplicado del cálculo de la citación normalizada atendiendo a las propuestas del *CWTS* y del *Karolinska Institute*

Un investigador ha publicado tres trabajos:

a) **Determinamos tipo documental, año y categoría. Se calcula el número de citas a las que denominados impacto observado:**

- (Trabajo A) Artículo publicado en 2014 en comunicación con 9 citas
- (Trabajo B) Revisión publicada en 2015 en educación con 5 citas
- (Trabajo C) Artículo publicado en 2016 en comunicación con 2 citas

b) **Establecemos valores de referencia medios mundiales atendiendo a tipo documental, año y área a los que denominados impacto esperado de un trabajo:**

- Los artículos de 2014 en comunicación tienen un promedio 6 citas
- Las revisiones de 2015 en educación tienen un promedio de 5 citas
- Los artículos de 2016 en comunicación tiene un promedio de 4 citas

c) **Con los datos anteriores se puede calcular la citación normalizada:** Según la fórmula del *CWTS* (**Moed; De-Bruin; Van-Leeuwen**, 1995):

- División entre el impacto observado y esperado:

$$(9+5+2) / (6+5+4) = 1,06$$

El valor 1,06 significa que se está un 6% por encima de la media mundial.

Según la fórmula del *Karolinska Institute* (**Lundberg**, 2007):

- Citas del artículo A entre su impacto esperado: $9/6 = 1,5$
- Citas de la revisión B entre su impacto esperado: $5/5 = 1$
- Citas del artículo C entre su impacto esperado: $2/4 = 0,5$
- Cálculo final del indicador: $(1,5+1+0,5)/3 = 1,00$

El valor 1 significa que se está justo en la media mundial.

impacto esperado. El impacto esperado se determina considerando no sólo el año de publicación y el área temática, sino también el tipo documental (artículo, revisión o carta). Posteriormente se divide la suma total del impacto observado entre la suma del impacto esperado de todas las publicaciones (tabla 3).

Su principal limitación radica en que al ser un promedio global de citas observadas y esperadas, los campos con mayores tasas de citación pueden acabar determinando el valor final. Esta limitación motivó la principal crítica a cargo de **Lundberg** (2007) del *Karolinska Institute*, quien indicaba la necesidad de establecer los promedios entre citas observadas y esperadas para cada artículo y después sumar los totales. La tabla 3 muestra mediante un ejemplo, las diferencias en los valores obtenidos según la formulación del indicador normalizado empleada.

La crítica de **Lundberg** (2007), refrendada por **Ophtof y Leydesdorff** (2010), dio lugar a un enconado debate (**Van-Raan**

et al., 2010) a raíz del cual los investigadores del *CWTS* decidieron redefinir la propuesta de Moed y desarrollar el indicador normalizado “definitivo”: el MNCS o *mean normalized citation score* (puntuación del promedio de citas normalizada). Sin embargo, lo que hicieron fue copiar la formulación de **Lundberg** (2007). La única diferencia que introducen tiene que ver con el tratamiento que hacen de las áreas temáticas ya que el MNCS aplica un factor corrector para las revistas que están en más de una categoría temática. Anteriormente una de las limitaciones del impacto esperado radicaba en los trabajos publicados en revistas indexadas en más de una categoría temática *Web of Science* ya que se realizaba un promedio de citación con todas las categorías, lo que introducía un sesgo evidente.

Asimismo **Waltman et al.** (2011b) indica dos limitaciones adicionales del MNCS:

- la primera se refiere a la asimetría que caracteriza a las distribuciones de citas, que hacen que sea problemático el uso de medias aritméticas;
- en segundo lugar, al normalizar por años los artículos más recientes pueden influir sustancialmente en el valor del MNCS.

La idea de base de los indicadores de citación normalizada es poder realizar comparativas entre individuos, instituciones o áreas de investigación

Posteriormente han surgido nuevas propuestas, sobre todo enfocadas a limitar los problemas y variaciones debidas al empleo de una clasificación temática u otra. Sin embargo la formulación de impacto normalizado de **Waltman** se ha mantenido y es la aceptada por la comunidad bibliométrica. Las variaciones en los valores obtenidos según la procedencia del indicador (*CWTS*, *InCites*, *Scival*, etc.) tienen más que ver con el trato que se hace a la multicategoría y no con los fundamentos teóricos del indicador. Por ejemplo *InCites* calcula el promedio de impactos esperados en las distintas categorías en las que un mismo trabajo está indexado mientras que el *CWTS* apuesta, como se ha reseñado, por un cálculo fraccionado.

Más allá de las consideraciones técnicas del indicador y las

Tabla 4. Ejemplo de cálculo del indicador de impacto normalizado según la convocatoria *Severo Ochoa / María de Maetzu* 2017. Período 2012-15. Sólo se incluyen artículos y revisiones indexados en *Web of Science*

	A Nº trabajos	B Nº citas	C B/A Promedio citas	D1 Categoría	Baseline empleado	D Valor medio	E C/D Comparación valores	F % Trabajos año	G E*F final
2012	3	17	5,67	Medical Informatics	WoS	10,61	0,53	0,18	0,10
2013	5	58	11,60	Artificial Intelligence	WoS	10,09	1,15	0,31	0,36
2014	5	29	5,80	Neurosciences	WoS	10,23	0,57	0,35	0,18
2015	3	12	4,00	Neurosciences	WoS	5,91	0,68	0,18	0,13
Impacto normalizado →									0,76

diferencias entre las propuestas, hay una serie de características que hacen que este indicador no sea recomendable al analizar grupos reducidos de publicaciones. La principal viene dada por ser un indicador relativo que es independiente del tamaño del conjunto de datos (se puede tener el mismo impacto normalizado con 1 trabajo que con 10).

Esto hace que se pueda premiar a investigadores muy poco productivos frente a otros más productivos que teniendo artículos de alto impacto, no son capaces de mantener los mismos niveles de citación en todos sus trabajos. De hecho es práctica habitual emplear este indicador junto con otros indicadores absolutos (número total de artículos publicados, número de artículos altamente citados, etc.) para facilitar su interpretación. A este tipo de limitaciones tendríamos que añadirle su poca funcionalidad en áreas de ciencias sociales y humanidades donde cualquier variación en el número de citas puede provocar una subida sustancial en el indicador, al ser áreas con promedios muy bajos.

Uno de los aspectos más destacados de las convocatorias *SO* y *MM* es que la fórmula de citación normalizada empleada (tabla 4) no coincide con ninguna de las propuestas descritas. Para las acreditaciones *SO/MM* se permite el uso de *Scopus* y *Web of Science* y el propio Ministerio pone a disposición los *baselines* de cada base de datos. Se ofrecen dos *baselines* por base de datos: una incluye conjuntamente artículos, revisiones y actas de congresos. Aquí encontramos el primer error técnico ya que no se distingue entre los tipos documentales y, por tanto, no se respetan las diferencias de citación entre ellos. Por ejemplo un artículo *Web of Science* publicado en 2013 tiene un promedio de citas de 9,54 y las revisiones tienen un promedio de 24,05. En la tabla 5 se ofrecen algunos datos para diferentes categorías y tipos documentales de *Web of Science* que evidencian la necesidad de presentar *baselines* desagregados por tipos documentales.

En la asignación de las categorías también se obvian las propuestas habituales que asignan a cada trabajo su categoría científica correspondiente y se normaliza en este sentido ítem a ítem. En la fórmula *SO/MM* el impacto esperado se selecciona para cada año y no para cada trabajo. La categoría donde más se publique cada año es la que se toma como

Tabla 5. Ejemplo de diferencias de citación en diferentes tipos documentales y categorías *Web of Science* en 2013.

Categoría <i>Web of Science</i>	Promedio citas artículo	Promedio citas revisión	Promedio citas actas de congreso
Global	9,54	24,05	0,88
Neurosciences	13,95	27,32	1,09
Sociology	3,49	4,14	0,39
Ecology	12,46	40,06	0,67
Artificial intelligence	10,39	31,19	1,84

referencia, y en caso de empate entre categorías en un año determinado se escoge la que considere más representativa el investigador, una situación que puede dar lugar a diferencias en el cálculo. Imaginemos que un investigador ha publicado 3 trabajos en *Scientometrics* en 2013 (tabla 5). Como la revista está en dos categorías (*Information Science & Library Science* y *CS-Interdisciplinary Applications*) el investigador podría escoger cualquiera de los dos aun existiendo diferencias significativas entre las mismas. Las diferencias entre escoger una categoría u otra pueden ser sustanciales, como se muestra en la tabla 6.

Otro aspecto problemático es que se dé la posibilidad de calcular el impacto normalizado a través de *WoS* o *Scopus*, con *baselines* totalmente diferentes en cuanto a la categorización y promedio de citas. Esta situación genera que ante un mismo investigador se calculen impactos normalizados diferentes, siendo el indicador casi siempre superior cuando usamos *Scopus*. Así, se puede dar la circunstancia contradictoria de que un investigador garante no cumpla el criterio en *WoS* pero sí en *Scopus*.

Sin embargo la casuística es mucho más amplia. La tabla 7 muestra el impacto normalizado calculado en ambas bases de datos para una serie de investigadores de distintas áreas. Mientras que Lucas Alado presenta prácticamente el mismo valor independientemente de la fuente, José Luis Verdegay cuenta con un impacto normalizado en *WoS* de 0,91 y en *Scopus* de 1,51. En este caso el investigador tiene el mismo número de trabajos en *WoS* y *Scopus* en la categoría Artificial Intelligence donde además los promedios de citas de los *baselines* son similares en ambas bases de datos, sin embargo el mayor número de citas recopiladas por *Scopus* hace que el indicador aumente significativamente. Casos más extremos son los de Leandro di Stasi (más de un punto de diferencia) o Luis Fermín Capitán (casi un punto de diferencia).

Tabla 6. Categorías de indexación de *Scientometrics* y ejemplo de su influencia en el cálculo del impacto medio esperado de las convocatorias *Severo Ochoa* y *María de Maetzu*

Categorías <i>Web of Science</i> de la revista <i>Scientometrics</i> en 2013	Artículos publicados por un investigador	Número citas	Impacto esperado de la categoría	Cálculo final del impacto normalizado
Information Science & Library Science	3	38	5,13	2,47
CS-Interdisciplinary Applications			9,09	1,39

Tabla 7. Cálculo del impacto normalizado para diferentes investigadores de la Universidad de Granada con las bases de datos Web of Science y Scopus para el período 2012-2015

Investigador UGR	Área conocimiento	Impacto normalizado baseline WoS	Impacto normalizado baseline Scopus
José Luis Verdegay Galdeano	Informática	0,91	1,5
Jonatan Ruiz Ruiz	Ciencias de los deportes	1,77	1,88
Luis Fermín Capitán Vallvey	Química	1,38	2,19
Lucas Alados Arboledas	Ciencias de la Tierra	1,66	1,67
Leandro di Stasi	Psicología	1,36	2,58

3. Recomendaciones

En esta nota se ha expuesto el origen y cálculo de un indicador cuyo uso se está extendiendo notablemente a nivel nacional y que, en el caso de las convocatorias *SO* y *MM*, es un requisito de imprescindible cumplimiento para los investigadores garantizando limitando sustancialmente el margen de acción de los centros. Por ello se han querido subrayar las limitaciones de la fórmula empleada a fin de que las evaluaciones sean más precisas y rigurosas. Teniendo en cuenta lo expuesto queremos ofrecer una serie de recomendaciones para un uso más acertado del indicador de citación normalizada y una evaluación más pertinente de los programas de excelencia analizados:

- Emplear indicadores bibliométricos para asesorar y evaluar solicitudes en lugar de utilizarlo como requerimiento administrativo para poder presentar una solicitud. Actualmente, tal y como está definida la convocatoria, investigadores de la altura de Kahneman, Premio Nobel en Economía, no tienen un impacto normalizado suficiente para poder ni tan siquiera presentar una solicitud.
- Apoyar el impacto normalizado con otros indicadores; por ejemplo, aquellos que contemplan la carrera científica completa como el índice *h* o reflejan mejor la excelencia científica como el número de trabajos altamente citados. De esta forma se diferenciaría entre investigadores con algunos trabajos excelentes y otros con una distribución de citas más equitativa por trabajo (Waltman *et al.*, 2011b). Debería incluirse un número mínimo de trabajos que garantice la validez estadística, ya que ahora mismo con tres trabajos ya se podría ser investigador garante.

Se debe apoyar el impacto normalizado con otros indicadores, como los que contemplan la carrera científica completa (índice *h*) o reflejan mejor la excelencia científica (número de trabajos altamente citados)

- Utilizar un indicador de citación avalado por la comunidad científica, como ocurre con el *MNCS* del *CWTS* o el *Normalized citation impact* implementado en *InCites*, que siguen el estándar planteado por Lundberg (2007). De esta forma se eliminarían deficiencias manifiestas en el cálculo

lo actual y se haría uso de un estándar internacional que goza del aval de la comunidad bibliométrica.

- Puesto que no se destinan recursos de cara a la evaluación, al contrario de lo que sucede en otros países (Robinson-García, 2017), al menos se debería permitir que se pueda emplear el indicador ofrecido por *InCites* o *Scival*. De esta forma los gestores podrían realizar análisis diacrónicos de sus investigadores de una forma más efectiva sin depender de los *baselines* anuales del Ministerio. Facilitaría la verificación por parte de los evaluadores ya que, por ejemplo, desde *InCites* se pueden descargar en diferentes formatos los trabajos con su impacto esperado y observado asociado. De esta forma el indicador es más fácil de calcular para el solicitante y de verificar por el evaluador. Asimismo, en muchos centros reducirían los costes, ya que el acceso a los datos, el cálculo de indicadores personalizados y su uso conlleva un coste económico importante y un coste de conocimiento todavía mayor que parece que recae invariablemente en gestores de política científica y en usuarios finales, más que en los expertos.

Se debería considerar la singularidad de las áreas a la hora de establecer el umbral de excelencia

- Considerar la singularidad de las áreas a la hora de establecer el umbral de excelencia ya que actualmente se solicita a todas superar el 1,5. No en todas las áreas es igual de fácil alcanzar el mínimo exigido. La UE en su conjunto en el período 2012-2015 en diferentes categorías alcanza el siguiente impacto: Physics (1,66), Medicine, General & Internal (1,48), Engineering - Industrial (1,19), Philosophy (1,14). Esto significa que en física con 1,50 se supera la media mundial pero no se alcanza el estándar europeo mientras que en filosofía con un 1,5 no sólo implicaría estar por encima del mundo si no también sustancialmente por encima del estándar europeo.
- Dadas las diferencias que se producen cuando se calcula el impacto normalizado en función de si se emplea *Web of Science* o *Scopus* sería pertinente indicar en la convocatoria para qué disciplinas científicas se puede hacer uso de una fuente u otra y no dejar en manos del solicitante la decisión. No parece tener sentido que en las ciencias exactas, experimentales o de la salud se utilice una fuente como *Scopus*, cuando por ejemplo todas las evaluaciones

de *Aneca/Cneai* en dichas áreas se articulan en torno a *Web of Science*, produciéndose cierta contradicción y falta de alineación de las políticas de evaluación científica.

- Buscar necesariamente otro indicador u otro criterio de selección mejor adaptado a los perfiles de publicación de las ciencias sociales y las humanidades, así como a sus características bibliométricas. Actualmente difícilmente se podrá aprobar una unidad en estas áreas (a excepción de economía y psicología experimental), no sólo porque no se puede aplicar a los libros sino porque al publicarse menos trabajos y existir menos citas, la variabilidad del indicador es mayor y pierde validez.

4. Referencias

Cabezas-Clavijo, Álvaro; Torres-Salinas, Daniel (2015). *Los sexenios de investigación*. Barcelona: UOC, colección El profesional de la información, n. 29. ISBN: 978 84 9064 530 7

De-Bruin, Renger E.; Kint, André; Luwel, Mark; Moed, Henk F. (1993). "A study of research evaluation and planning: The University of Ghent". *Research evaluation*, v. 3, n. 1, pp. 25-41. <https://goo.gl/k3PFGX>
<https://doi.org/10.1093/rev/3.1.25>

Lundberg, Jonas (2007). "Lifting the crown—citation z-score". *Journal of informetrics*, v. 1, n. 2, pp. 145-154. https://www.researchgate.net/publication/220065906_Lifting_the_crown_-_citation_z-score
<https://doi.org/10.1016/j.joi.2006.09.007>

Moed, Henk F.; De-Bruin, Renger E.; Van-Leeuwen, Thed N. (1995). "New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications". *Scientometrics*, v. 33, n. 3, pp. 381-422
<https://goo.gl/QjBiWd>
<https://doi.org/10.1007/BF02017338>

Ophof, Tobias; Leydesdorff, Loet (2010). "Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance". *Journal of informetrics*, v. 4, n. 3 pp. 423-430.
<https://arxiv.org/abs/1002.2769>

<https://doi.org/10.1016/j.joi.2010.02.003>

Robinson-García, Nicolás (2017). "Los resultados del REF2014 del Reino Unido marcan el camino a seguir en la evaluación científica". *Blok de BiD*, 12 diciembre. <http://www.ub.edu/blokdebid/es/node/831>

Schubert, Andras; Glänzel, Wolfgang; Braun, Tibor (1983). "Relative citation rate: A new indicator for measuring the impact of publications". En: Tomov, D.; Dimitrova, L. (eds.). *Proceedings of the First national conf with intl participation on scientometrics and linguistics of scientific text*, Varna, pp. 80-81. <https://goo.gl/P1X3UU>

Van-Raan, Anthony (1999). "Advanced bibliometric methods for the evaluation of universities". *Scientometrics*, v. 45, n. 3, pp. 417-423. <https://goo.gl/8dcnoK>
<https://doi.org/10.1007/BF02457601>

Van-Raan, Anthony; Van-Leeuwen, Thed N.; Visser, Martijn S.; Van-Eck, Nees J.; Waltman, Ludo (2010). "Rivals for the crown: Reply to Ophof and Leydesdorff". *Journal of informetrics*, v. 4, n. 3, pp. 431-435. <https://arxiv.org/pdf/1003.2113.pdf>
<https://doi.org/10.1016/j.joi.2010.03.008>

Vinkler, Peter (1986). "Evaluation of some methods for relative assessment of scientific publications". *Scientometrics*, v. 10, n. 3-4, pp. 159-177. <https://goo.gl/QuV9LB>
<https://doi.org/10.1007/BF02026039>

Waltman, Ludo; Van-Eck, Nees J.; Van-Leeuwen, Thed N.; Visser, Martijn S.; Van-Eck, Nees J. (2011a). "Towards a new crown indicator: Some theoretical considerations". *Journal of informetrics*, v. 5, n. 1, pp. 37-47. <https://arxiv.org/pdf/1003.2167.pdf>
<https://doi.org/10.1016/j.joi.2010.08.001>

Waltman, Ludo; Van-Eck, Nees J.; Van-Leeuwen, Thed N.; Visser, Martijn S.; Van-Eck, Nees J. (2011b). "Towards a new crown indicator: An empirical analysis". *Scientometrics*, v. 87, n. 3, pp. 467-481. <https://doi.org/10.1007/s11192-011-0354-5>

La **Fundación Biblioteca Social** es una institución sin ánimo de lucro que se constituyó en el año 2014. Tiene como objetivo contribuir a compensar los desequilibrios sociales apoyando proyectos que llevan a cabo las bibliotecas públicas, dirigidos a los sectores más vulnerables de la sociedad.

¿Colaboras?

fundacionbibliotecasocial.org
info@fundacionbibliotecasocial.org
@Biblio_Social
FundacionBibliotecasSocial



Mapa de proyectos de bibliotecas públicas para la inclusión social. 2016.

