
ESTUDIOS / RESEARCH STUDIES

Productivity and impact of Spanish researchers: reference thresholds within scientific areas

Evaristo Jiménez-Contreras*, Nicolás Robinson-García*,
Álvaro Cabezas-Clavijo*

Abstract: Reference thresholds for the scientific production and impact of internationally visible Spanish research within the areas defined by the Spanish National Agency for Evaluation and Prospective (ANEP) are presented. These percentile reference tables are constructed from the population of researchers who applied for a project within Spain's National R&D Plan 2007 ($n = 3.356$) and are to serve as benchmarks, permitting comparisons between researchers' bibliometric behavior and mean performance in their respective scientific disciplines. Data relating to mean production, impact and visibility for each ANEP area are also presented. The internationalization of these areas between 2000 and 2006 is discussed, with special emphasis on the Social Sciences. Finally, we suggest funding agencies and research institutions use these reference thresholds as assessment tools in their selection processes.

Keywords: Bibliometric indicators, reference thresholds, benchmarking, Spanish National Agency for Evaluation and Prospective (ANEP), science, scientific research, Spain.

Productividad e impacto de los investigadores españoles: umbrales de referencia por áreas científicas

Resumen: Se presentan umbrales de referencia de producción e impacto científico de la investigación española con visibilidad internacional para las áreas definidas por la Agencia Nacional de Evaluación y Prospectiva (ANEP) en sus convocatorias. Tomando como población los solicitantes de proyectos del Plan Nacional de I + D 2007 ($n = 3.356$) se construyen tablas de referencia por percentiles que funcionan a modo de benchmarks, permitiendo efectuar comparaciones entre el comportamiento bibliométrico de un investigador y los registros de referencia en su área científica. Igualmente se ofrecen los datos de producción, impacto y visibilidad promedios para las áreas ANEP, y se discute el proceso de internacionalización de dichas áreas en el período 2000-2006,

* EC3: Evaluación de la Ciencia y la Comunicación Científica, Departamento de Biblioteconomía y Documentación, Universidad de Granada. Correo-e: evaristo@ugr.es; elrobin@ugr.es; acabezasclavijo@gmail.com.

Received: 05-01-2011; revision received: 03-02-2011; accepted: 04-02-2011.

Spanish version on the web.

con una especial atención a las Ciencias Sociales. Finalmente, se sugiere el uso de umbrales de referencia como método de evaluación tanto para agencias financiadoras como para instituciones de investigación en sus procesos selectivos.

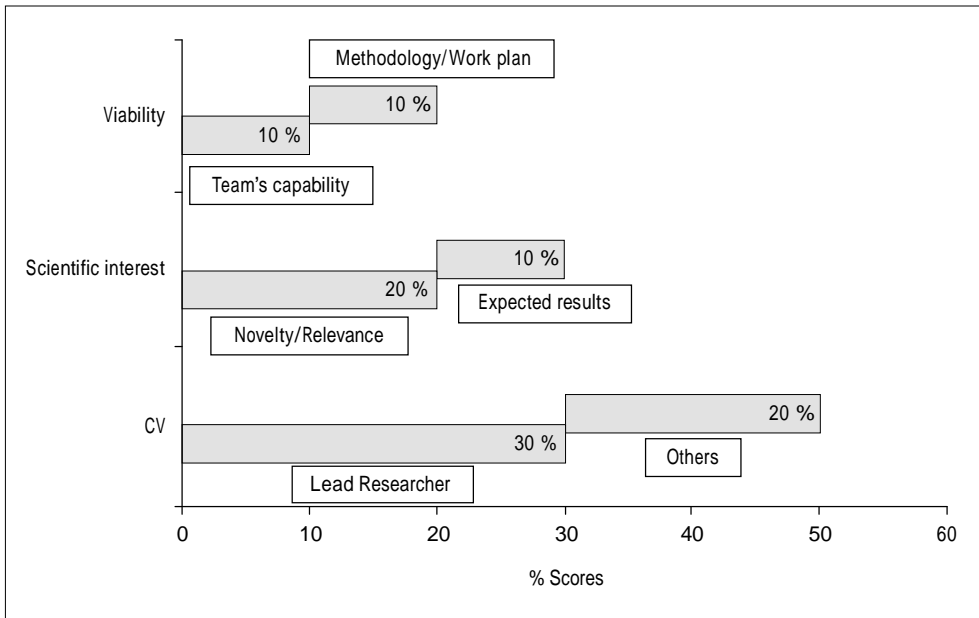
Palabras clave: *Indicadores bibliométricos, umbrales de referencia, benchmarking, ANEP, ciencia, investigación científica, España.*

1. Introduction

In Spain, the growing importance of bibliometric parameters in University selection and promotion procedures or in awarding grants or funding for research projects has meant the public sector agencies involved need sound indicators, adapted to the different aspects of assessment, that guarantee the fair assessment of competing applications. Over the last 25 years, the Spanish scientific community has become increasingly professional (Sanz-Menéndez, 1997; Cruz-Castro and Sanz-Menéndez, 2007) and a number of national and regional public sector agencies charged with managing scientific activity have appeared. These include the assessment agencies among which one of the most important is the National Agency for Evaluation and Prospective (*Agencia Nacional de Evaluación y Prospectiva*, hereafter ANEP), which has been developing an assessment system that delimits and assesses aspects of content, budget and curriculum vitae (CV) as separate entities. Since its creation in 1986, peer-review has been the preferred method of assessing proposals. Each application is assessed by two expert reviewers chosen by the scientific discipline coordinator. They make independent written reports that include the corresponding score for each section and a final total (Gordillo et al., 2004). Although ANEP assesses competitive funding proposals for the Ministry of Science and Innovation and other public bodies, its most important activity concerns applications to the National R&D Plan (PN). The single most important source of public finance for most Spanish researchers, the PN constitutes an opportunity to obtain financial support to conduct research for a period of three years. Essentially, the criteria used to award funding are based on an assessment of the lead researcher's CV (30 %) and of those of the other research team members (20 %) (figure 1). In this process, sound experience, together with a good team or a technically acceptable proposal ensure success. In the exact and experimental sciences and, increasingly, the social sciences too, reviewers frequently assess CV quality—defined as the value and repercussion of publications—through the number of international publications (i.e., those found in the ISI Web of Science [WoS]), and the impact factors (IFs) of the journals in which they appear. In other words, they resort to bibliometric indicators.

In Spain, it is standard practice to turn to IFs when offering incentives for productivity. In fact, they are practically the only criterion used. In the late 1980s, the Ministry of Education's National Research Activity Assessment Commission (*Comisión Nacional Evaluadora de la Actividad Investigadora*, hereafter CNEAI)

FIGURE1
ANEP assessment criteria scores. 2009 funding round



established 6-yearly productivity bonuses to provide one such incentive. The essential criterion for awarding bonuses were standards applied in the pure sciences—together with factors such as a legal framework favoring state sector universities—that led to the increased internationalization of Spanish science in the 1990s (Jiménez-Contreras et al., 2003). In the last decade this has continued largely because researchers have adapted to the pattern of international publication. Moreover, more recently, the citation level of Spanish scientists has converged on international figures following years of under-performing by comparison with their international counterparts (Delgado-López-Cózar et al., 2009). However, the use of these criteria has generated widespread controversy (Camí, 1997; Bordons et al., 2002; Lawrence, 2002) because journal IFs do not represent the impact of the individual studies or that of their authors (Seglen, 1997). This is due to specific methodological and conceptual limitations of the indicator: e.g., the citation window is only two years; an asymmetric relation exists between the number of citation studies received and the IFs of the journals that publish them (van Leeuwen and Moed, 2005). However, despite these limitations, agencies, reviewers, and scientists themselves base a substantial proportion of their scientific decisions on this indicator. Ultimately, the reason for this success (not forgetting the saving in costs) is the assumption that the IF, although it may not predict the impact a specific article will achieve, *does* represent the visibility, prestige and total sum of obstacles authors have to overcome to finally see their work in print. In other

words, the IF spotlights study quality in a way that is reasonably proportionate to the journal's impact (Garfield, 2003).

Given this is the case, it has become fashionable to use IFs as a research and assessment tool in empirical studies (e.g., Alonso-Arroyo et al., 2006; Buela-Casal et al., 2004; González-de-Dios et al., 2009). Notwithstanding, some authors have made proposals or reflected on the methodology of using the IF or real or observed citations in attempts to resolve some of their limitations, especially those related to inter-category comparison (Schubert and Braun, 1996). A further issue is that posed by the need for a normalization procedure to permit cross-disciplinary comparisons, for which alternatives exist. One of the simplest procedures uses direct rankings and applies one or various criteria: production, funding, impact or whatever (Torres-Salinas et al., 2009b; Buela-Casal et al., 2010), although special care must be taken due to the many methodological factors involved in constructing the rankings (van Raan, 2005). Another alternative is percentile range, which involves rating the aggregates being assessed on a scale that permits inter-category comparisons (Lewison et al., 1999; Costas et al., 2010). Finally, another option is to develop methods based on production-related reference indicators—such as impact or collaboration—which are normalized to facilitate comparisons between disciplines. The Centre for Science and Technology Studies at the University of Leiden, The Netherlands (hereafter CWTS) proposed just such a set of indicators, which constitutes an appropriate means of assessing institutional compliance with scientific policy and has been successfully applied in different institutions, geographic areas and thematic fields (Moed et al., 1985; van Raan and van Leeuwen, 2002).

Beyond the methods used, our review of international publications revealed several studies aimed at drawing up bibliometric profiles of different constructs. These include, at the level of mesoanalysis, topic-based profiles of academic institutions and researchers in Great Britain (Carpenter et al., 1988), Mexico (Macías-Chapula et al., 2004), or Spain (Moya-Anegón et al., 2005; Torres-Salinas et al., 2009a), and at European research institutions (Thijs and Glänzel, 2009). Much has been written about the individual assessment of researchers by establishing discipline-based rankings. In the field of Information Science, authors have used criteria such as the h index (Cronin and Meho, 2006; Oppenheim, 2007) or production and citations received (Meho and Spurgin, 2005; Jiménez-Contreras et al., 2006). At a micro level, more complex approaches to the development of bibliometric profiles have recently been reported by Abramo and D'Angelo (2011), in the context of national assessment, and by Costas and his coauthors (Costas and Bordons, 2005; Costas, 2008; Costas et al., 2010). These studies establish profiles or levels of excellence that distinguish between three levels of analysis (production, impact and visibility) in order to place researchers on scales or within ranges of excellence or research quality. Typically, this bottom-up method (van Leeuwen, 2007) requires highly careful data collection and indicator-processing and development because search and processing errors can substantially distort the final results. In general, we can say that the closer the

analysis is to the authors, the wider-ranging the battery of indicators should be and, the more cautious we should be both in data-collection and indicator-processing and development, which should be in consonance with the final use the data will be put to.

As subject matter experts, reviewers doubtless know the standard parameters of productivity and quality in their discipline but they lack of any reference frame or accurate threshold to enable them to place the researchers being assessed in a productivity context constructed from empirical data and bounded by national research standards. We encounter the same scenario when looking at the impact of the studies these scientists publish. Frames of reference that permit comparative analysis or assessment of production and citation would provide reviewers with benchmarks against which to «measure» the CVs of researchers under assessment, thus enhancing the objectivity of the process. Hence, the peer review assessment process would gain objectivity in response to some of its current defects (Gordillo et al., 2004), favoring what in the literature is known as «informed peer review» (Lewison et al., 1999; van Raan, 1996).

Whether resorting to standard indicators has modified scientific practice and favored deviant or fraudulent behavior has been discussed (Butler, 2003). To overcome these problems and avoid the use of «formulaic» researcher assessment (Moed, 2005), the combined use of bibliometric indicators and expert reviewers has been proposed (Weingart 2005). In the context of reference frame construction for scientific disciplines, Thomson Reuters' Essential Science Indicators (ESI) represent a tool that —although not widely used— proves of inestimable value when establishing thresholds and citation means as a function of the category studies belong to and the date when they were published, thus neutralizing the two main variables that can influence the measurement of an article's impact. The ESI provide baselines: means and percentiles that enable us to establish the relative position of a research article within its field according to the number of citations received and the time since its publication. These indicators give us the mean citations of an article, and the percentile range, by year and category in both cases. The ESI tables, divided into 22 different-sized disciplines and excluding the Humanities, are constructed from the total number of WoS journals. They undergo bimonthly updates (Ruiz-Pérez et al., 2008) and, therefore, are dynamic time frames.

2. Objectives

The primary objective of the present study is to construct bibliometric profiles of applicants to Spain's National R & D Plan for project funding and, on the basis of this data, construct reference thresholds to aid reviewers when taking decisions on the basis of bibliometric evidence. In other words, we aim to establish the standards of publication and impact of those Spanish researchers who actively seek funding.

Specifically, we aim to:

- Construct a reference table similar to the ESI but adapted to the Spanish context, and related to performance in both production and impact.
- Define the main bibliometric indicators of production, productivity, impact and visibility by discipline.
- Test the validity of our proposed method for further use —with or without any relevant modifications— to ensure maximum objectivity in assessing researchers' CVs.

3. Materials & method

A retrospective descriptive analysis of the bibliometric performance of the lead Spanish researchers who applied for research project funding through the Spanish PN funding round for 2007, is conducted. We included 3356 researchers from the ANEP disciplines, excluding Law, Philology and Philosophy, and History and Art. All data were provided by the ANEP. The disciplines that received most applications were Fundamental Biology (12%), followed by Chemistry and Economics. To calculate the indicators and validate our study design, we calculated the percentage of researchers in each discipline with at least one article published during the study period. Data was collected by searching for these researchers' scientific production on the Thomson Reuters on-line WoS databases: the Science Citation Index-Expanded (SCI), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (A&HCI), Conference Proceedings Citation Index-Science (CPCI).

Information Science experts conducted a manual search for each researcher, retrieving all records dated between 2000 and 2006. Special care was taken to allow for the different forms of Spanish surnames appearing in ISI databases (Ruiz-Pérez et al., 2002) and the affiliations a researcher may have had during the study period (e.g., due to study leave spent in international centers). Researcher production was initially stored in text format and later exported to *ProCite 5* bibliographic management software. Once all production data had been collected, it was transferred to a purpose-built *MS Access 2003* relational database together with the corresponding biographical data (researcher name, surname, project code number, researcher type, institution and center). The document types analyzed were article, review, letter, editorial material and proceedings paper. Information relating to the journal of publication's IF was downloaded from Thomson Reuters *Journal Citation Reports (JCR)* and added to the database. We also included the number of citations received by each article retrieved using the *Create Citation Report* function on the online WoS results page. Once any given author's production had been identified, this option enabled us to download the citations received for each record as a function of the year of the citing article. The time frame for citations was 2000-2008. Finally, we connected citing documents with documents cited in the database. Data collection took place between February 2009 and May 2010.

We subsequently calculated each ANEP discipline indicator with *Access 2003*. The ANEP discipline categories have remained relatively stable over time although the Life Sciences have recently been modified. The categorization by disciplines presented here (table I) is that which was in use at the time of the 2007 PN.

TABLE I
ANEP disciplines and acronyms

Acronym	Discipline
AGR	Agriculture
BMED	Biomedicine
CEA	Civil engineering and architecture
CHE	Chemistry
CHT	Chemical technology
CLIM	Clinical medicine and epidemiology
CSI	Computer science and information technology
ECO	Economics
ECT	Electronic and communication technology
EDU	Education science
EEC	Electrical, electronic and control engineering
ESC	Earth sciences
FSB	Fundamental and system biology
FST	Food science and technology
LFF	Livestock farming and fisheries
MNA	Mechanical, naval and aeronautic engineering
MST	Materials science and technology
MTM	Mathematics
PHY	Physics and space sciences
PPH	Physiology and pharmacology
PSY	Psychology
SSC	Social sciences
VAB	Vegetable and animal biology, and ecology
<i>Not analyzed</i>	Law
<i>Not analyzed</i>	Philology and philosophy
<i>Not analyzed</i>	History and art

Statistical analysis was conducted with *SPSS 15.0 for Windows*. The indicators calculated were:

• **Production**

- **Productive researchers.** Researchers with at least one document published during the study period.
- **Ndoc/res.** Mean documents per researcher.
- **RV Rate of variation.** Percentage difference between production recorded in the first (2000) and last (2006) years of the study.
- **Baselines (production thresholds).** Cumulative frequency of production per researcher by percentile. This indicator informs us of the number of articles needed for a researcher to be placed in the 90-, 75- and 50-percentiles of the distribution of scientific production by year of publication.

• **Visibility**

- **IFAve.** Mean Impact Factor.
- **Top3.** Publications in the top 3 journals by JCR category.
- **Q1.** Publications in journals in the first quartile by JCR category.

• **Impact**

- **CitAve.** Mean citations per document.
- **% Ndoc cited.** Percentage of documents cited.
- **Baselines (citation thresholds).** Cumulative frequency of citation by percentile. This indicator informs us of the number of citations needed for an article to be placed in the 90, 75 and 50 percentiles of the distribution of scientific production by year of publication.

4. Results

4.1. Production

Some 3356 researchers (excluding those in the Humanities) participated in the PN funding round for 2007, grouped by discipline as shown in table II.

The data obtained range from 99.58 % of productive researchers in Chemistry to 29.45 % in Social Sciences. Most categories have a ratio of productive researchers of >90%. The exceptions are the four Social Science disciplines, which present clearly-defined patterns. Economics and Psychology have higher year-on-year productivity (with spectacular growth in Economics); Education (also with very high relative growth) and Social Sciences remain at <35%. Civil Engineering sharply contrasts with other fields of Engineering having one third of researchers with no production. Disciplines like Physics, Fundamental Biology or Mathematics,

TABLE II
*Total number of researchers, productive researchers, and rate of variation
2000-2006 by ANEP discipline*

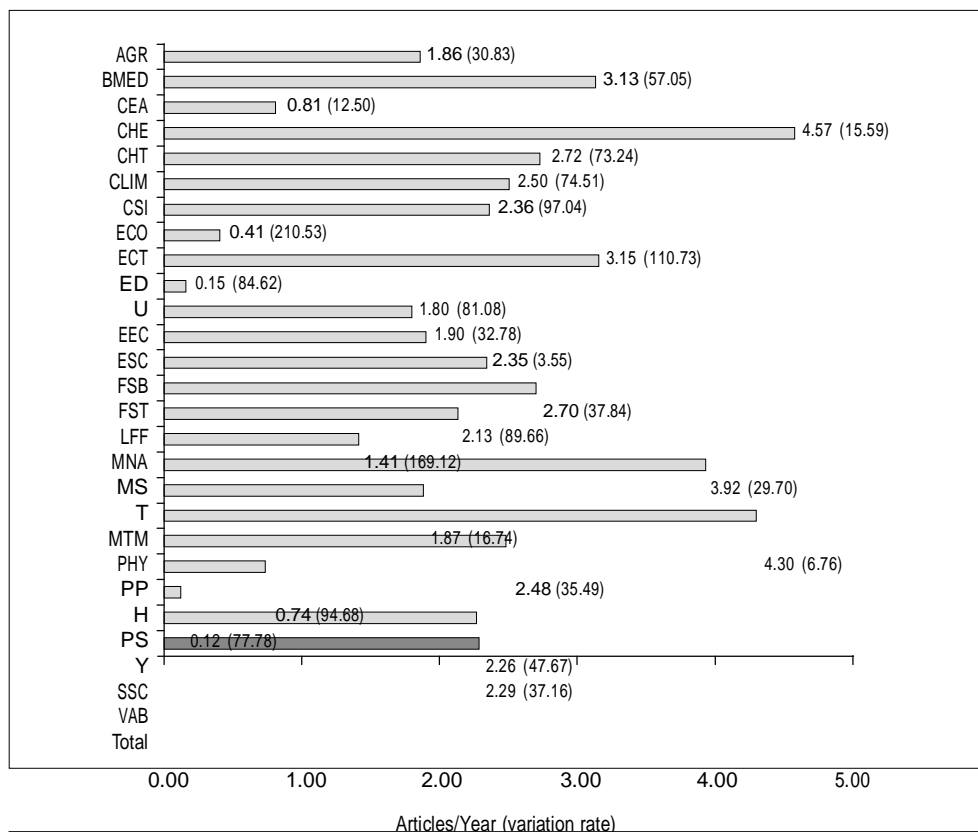
Discipline	Resear chers	Pro ductive	% Productive	RV 00-06
AGR	132	127	96.21	24.39
BMED	124	123	99.19	16.13
CEA	61	40	65.57	33.33
CHE	237	236	99.58	9.00
CHT	106	100	94.34	31.75
CLIM	51	49	96.08	11.43
CSI	152	143	94.08	45.68
ECO	182	131	71.98	157.14
ECT	139	131	94.24	40.00
EDU	119	41	34.45	112.50
EEC	86	84	97.67	36.17
ESC	141	130	92.20	17.44
FSB	405	400	98.77	0.64
FST	128	126	98.44	9.38
LFF	96	92	95.83	34.43
MNA	91	85	93.41	60.00
MST	171	167	97.66	17.74
MTM	131	130	99.24	5.38
PHY	180	178	98.89	1.31
PPH	145	144	99.31	10.53
PSY	169	136	80.47	56.86
SSC	146	43	29.45	35.71
VAB	164	160	97.56	25.00
Total	3356	2996	89.27	19.94

which have adopted a solid, international pattern of publication, show minor rates of variation. However, some technologies and fields of Engineering show high growth, probably due to the wider coverage of conference papers in the reference database.

We calculated the mean number of documents produced per researcher and the annual mean by ANEP category. To observe the chronological tendency, we calculated the rate of variation and found researchers increasingly adopt the international publication pattern. This makes inter-disciplinary differences in productivity clear, ranging from 4.3-4.6 documents per researcher per year in Physics and Chemistry to barely 0.1 documents in Social Sciences and Education. The extremely low starting figures in these disciplines explain the high relative increase and simultaneously low ISI productivity of these researchers. The greatest rates of variation occur in some fields of Engineering and Technology and in Economics (figure 2). One factor that influences the different discipline-related rates of productivity is the rate of collaboration in the studies published. The coauthorship index —i.e. mean signatories per article— ranges from 8.2 in Physics (excluding articles with >250 signatories) to 2.5 in Economics. The highest mode is 6 authors in Biomedicine, Clinical Medicine and Epidemiology, and Livestock farming and Fisheries; the lowest, 1 author, is in Social Sciences.

FIGURE2

Mean production per researcher per year, and rate of variation for 2000-2006 (in parentheses) by ANEP discipline



Given that productivity means may be biased by the asymmetric distribution of research productivity, we also calculated the production thresholds per year of the study period and for the whole period. We selected 50-, 75-, 80-, 90- and 99-percentiles and found two very different trends: on the one hand, the results for most disciplines are consistent, and growth between percentiles is exponential—i.e., the productive effort needed for a researcher to enter the higher distribution percentiles is substantially greater, hence these benchmarks are thresholds only surpassed through greater effort, in terms of reviewer criteria. Figures tend to hold in the long-term, with productivity growing towards the end of the study period. In disciplines like Chemistry or Physics, >20 articles (in 7 years) are needed to reach the category median. In contrast, in the Social Sciences (and Civil Engineering) the high number of non-productive researchers conditions results and performance is abnormal. For example, a researcher with no production can be placed in the three-year 75-percentile for distribution in disciplines like Education. Table III shows percentile data for researcher CV assessment at 3.5 and 7 years.

TABLE III

Production thresholds (baselines) by ANEP discipline for the 50-, 75- and 90-percentiles at 3.5 and 7 years

Discipline/ Percentile	3 years			5 years			7 years		
	P50	P75	P90	P50	P75	P90	P50	P75	P90
AGR	4	7	12	7	11	21	10	16	29
BMED	7	12	18	12	20	29	17	28	41
CEA	1	3	8	1	4	13	2	6	18
CHE	10	17	27	17	29	46	24	40	64
CHT	6	9	16	9	16	27	13	22	38
CLIM	5	9	17	8	15	29	11	21	40
CSI	6	9	15	9	15	24	13	21	34
ECO	1	2	3	1	3	5	2	4	7
ECT	7	14	21	11	24	35	16	33	49
EDU	0	0	2	0	1	3	0	1	4
EEC	4	8	12	7	13	20	10	18	28
ESC	3	6	11	6	11	19	8	15	26
FSB	6	9	14	9	15	24	13	21	33
FST	7	11	17	11	19	28	16	26	39
LFF	5	9	10	9	16	17	12	22	24

TABLE III (cont.)

Discipline/ Per centile	3 years			5 years			7 years		
	P50	P75	P90	P50	P75	P90	P50	P75	P90
MNA	3	6	10	6	10	17	8	14	24
MST	9	16	26	14	27	43	20	38	60
MTM	5	7	11	8	12	19	11	17	26
PHY	10	18	25	16	31	42	23	43	59
PPH	6	10	14	10	16	24	14	23	33
PSY	1	3	6	2	5	9	3	7	13
SSC	0	0	1	0	1	1	0	1	2
VAB	5	9	13	9	14	22	12	20	31
Total	5	9	16	8	15	26	11	21	37

* *Example:* A researcher in the field of Agriculture with a production of 12 articles in 5 years would be placed among the 10-25% most productive researchers of the discipline (in the P75-P90 range).

4.2. Visibility and impact

Impact and visibility indicators in relation to articles in journals with an IF and articles cited parallel the ratio of productive to non-productive researchers. The greatest imbalance occurs in the Social Sciences where 11% of researchers published in journals with no IF for the year of publication, which reflects publications recently incorporated into the WoS database. In Education, this affects 8.4% of researchers. In Civil Engineering and, to a lesser extent, in other engineering, there is also a certain imbalance due to publications in conference papers, a document type with no IF. The same occurs if we compare those productive researchers who are cited and those who are not. In this case, Economics and Psychology are also included in the group. Overall, 4% of productive researchers have non-cited production. Food Technology and Earth Sciences are the only disciplines in which all productive researchers have achieved at least one citation for one of their publications (table IV).

The indicators measuring performance as visibility in top3 and 1st quartile (Q1) journals differ greatly between categories. In Chemistry, Physiology or Fundamental Biology practically all researchers have published one or more articles in Q1 journals. The Social Sciences are ranked at the lower end of this indicator. Overall, 3 out of 4 researchers published at least one article in a Q1 journal during the study period. Nearly half of the researchers also published in one of the top3 journals in their discipline. Values range from 77% in Biomedicine to 0% in Education, where no applicant published in a top3 journal.

TABLE IV
Percentage of researchers by visibility and impact indicators by ANEP discipline

Discipline	% pr oductive	% IF	% Q1	% Top3	% cited
AGR	96.21	93.18	79.55	50.00	92.42
BMED	99.19	99.19	95.16	77.42	97.58
CEA	65.57	55.74	40.98	19.67	52.46
CHE	99.58	99.58	97.89	62.03	98.73
CHT	94.34	93.40	88.68	51.89	93.40
CLIM	96.08	96.08	82.35	70.59	92.16
CSI	94.08	90.13	58.55	21.05	88.16
ECO	71.98	66.48	29.67	13.74	62.64
ECT	94.24	91.37	69.06	33.81	90.65
EDU	34.45	26.05	10.08	0.00	24.37
EEC	97.67	91.86	58.14	25.58	89.53
ESC	92.20	91.49	82.98	49.65	92.20
FSB	98.77	98.77	97.78	68.15	96.05
FST	98.44	98.44	97.66	72.66	98.44
LFF	95.83	95.83	93.75	57.29	93.75
MNA	93.41	89.01	67.03	31.87	86.81
MST	97.66	97.66	92.40	69.01	95.91
MTM	99.24	99.24	71.76	16.03	96.95
PHY	98.89	98.33	95.00	71.11	97.22
PPH	99.31	99.31	97.93	72.41	98.62
PSY	80.47	77.51	46.15	20.12	73.37
SSC	29.45	18.49	8.90	3.42	17.12
VAB	97.56	97.56	90.85	54.27	96.95
Total	89.27	87.10	74.82	46.51	85.40

Article performance was >50% in Q1 journals, falling to 12% in top3 journals (table V). In Q1 journals, the performance of Fundamental Biology and Physics researchers was excellent; in top3 journals, Biomedicine and Food technology stood out with approximately 20% ratio. The highest percentages of articles cited are found in Chemistry, Physiology and Fundamental Biology. These disciplines, together with Biomedicine, also attain the highest mean of citations per document. Electrical Engineering and the two fields encompassed by Information and Communication Technology had the highest levels of non-cited documents,

possibly due to the importance given to conference proceedings. With a mean 2 citations per document, Education is the least cited discipline.

TABLE V
Visibility and impact indicators by ANEP discipline

Discipline	Mean IF	% Q1	% Top3	Mean citation	% cited
AGR	1.70	45.69	12.53	7.60	74.56
BMED	3.86	59.49	20.06	17.14	87.21
CEA	0.95	47.78	12.96	5.75	66.57
CHE	2.89	58.95	8.33	14.16	93.56
CHT	1.90	59.76	14.31	10.58	90.05
CLIM	2.82	48.56	16.07	11.13	84.66
CSI	0.72	14.62	2.99	3.07	49.70
ECO	0.71	21.97	6.78	3.47	70.04
ECT	1.31	37.83	9.74	4.12	51.58
EDU	0.79	23.53	0.00	2.21	60.94
EEC	0.93	30.95	9.89	2.97	47.65
ESC	1.94	55.63	14.59	11.05	86.38
FSB	4.82	67.44	14.62	21.43	92.14
FST	1.82	63.11	19.36	10.61	90.10
LFF	1.87	57.69	11.57	9.41	87.10
MNA	1.09	44.30	8.19	5.83	73.70
MST	1.98	56.16	13.69	9.94	85.75
MTM	0.67	23.52	1.97	4.02	71.79
PHY	3.09	65.29	14.49	12.31	79.62
PPH	3.72	58.29	10.69	15.39	92.70
PSY	1.57	26.29	7.99	6.93	81.54
SSC	2.11	38.82	12.94	3.94	61.60
VAB	2.22	45.59	9.02	9.43	87.07
Total	2.63	53.46	11.96	11.59	82.02

We calculated 50-, 75-, and 90-percentile figures for citation. Given that time is a factor that influences citation, we calculated the indicators as a function of years since publication. Table VI shows 50-, 75- and 90-percentiles corresponding to a 9-year time frame for articles recorded in 2000; a 3-year time frame has been used for articles published in 2006 (citations were recorded until 2008). If we

take the 9 years of the longest period calculated as our reference, an article usually needed 9 citations to reach the distribution median. This rose to 16 citations in Fundamental Biology; however 0 citations were needed in Social Sciences and Electrical Engineering. The 9-year results (production was lowest in 2000) are based on samples that included abnormal cases. For example, in some categories articles (published in 2002) that had 7 years to be cited show citation thresholds above those of articles that had a wider time frame to be cited (9 years).

TABLE VI

Citation thresholds (baselines) with 3-, 5-, 7- and 9-year time frames for 50-, 75- and 90-percentiles, by ANEP discipline

Discipline/ Percentile	3 years			5 years			7 years			9 years		
	P50	P75	P90	P50	P75	P90	P50	P75	P90	P50	P75	P90
AGR	1	4	9	4	9	15	7	16	28	7	15	30
BMED	4	10	18	8	18	38	11	24	44	10	28	57
CEA	0	2	6	3	6	11	5	9	21	5	20	33
CHE	4	8	15	8	15	28	11	20	34	13	24	42
CHT	3	6	12	6	13	25	9	17	31	10	18	40
CLIM	3	6	10	7	17	24	6	15	30	9	23	48
CSI	0	1	3	0	2	7	1	3	10	2	7	17
ECO	0	1	2	2	4	7	2	9	14	4	7	13
ECT	0	2	6	1	5	11	1	6	15	1	6	15
EDU	1	2	3	2	4	7	1	7	11	1	2	7
EEC	0	1	4	1	5	11	0	6	13	0	5	10
ESC	3	7	11	7	13	34	7	15	30	10	19	35
FSB	5	11	20	10	21	41	14	30	60	16	36	75
FST	3	6	10	7	14	22	9	18	30	11	22	42
LFF	3	5	8	7	11	18	5	11	22	9	18	33
MNA	1	3	6	3	8	17	4	11	17	5	10	20
MST	2	6	12	5	11	20	6	13	28	7	16	30
MTM	1	2	4	2	5	9	3	7	14	3	8	16
PHY	3	7	15	5	14	27	7	17	35	10	21	40
PPH	4	8	14	9	18	37	13	25	43	12	27	55
PSY	2	4	6	5	9	17	5	11	23	4	12	23
SSC	1	2	4	3	6	12	4	8	17	0	2	8
VAB	2	4	8	5	11	21	7	14	22	8	23	43
Total	2	6	12	5	13	24	7	17	33	9	21	42

* *Example:* An article in the field of Mathematics with 4 citations received in 3 years would be situated among the 10% most cited articles in the discipline (P90). The same article (with 4 citations) in Molecular Biology would not enter the 50% most cited articles.

5. Discussion

On the basis of the data analyzed we can confirm the existence of three types of disciplines: the first consists of those that from the outset have been fully integrated into international research. These are Science disciplines (AGR, FSB, BMED, VAB, ESC, PPH, PHY, LFF, CLIM, MTM, CHE, FST, MST, CHT). All match the pattern of international production and the vast majority of funding applications were from productive researchers. In disciplines like FSB or PHY, the percentage of productive researchers held steady during the study period. In this group, disciplines like Agriculture, Animal and Plant Biology and Ecology, or Livestock Farming and Fisheries, have become fully internationalized during the study period. These disciplines stand out for productivity levels of 1.9-4.6 documents per researcher per year, and (except in Mathematics) for publishing >45% of articles in Q1 journals. Moreover, production and citation thresholds held relatively stable year-on-year, although they appear to have been more demanding towards the end of the study period, probably due to less experienced applicants joining the funding round and the general increase in productivity in Spain.

Secondly, we have the group comprising the Engineering and Information and Communication Technology disciplines (ECT, CEA, MNA, EEC, CSI), which have clearly distinguishable profiles. The study period saw the incorporation of more researchers into the system, with productivity ranging from 0.8 (CEA) to 3.1 (ECT) documents per researcher per year. Visibility indicators describe 14%-48% of Q1 documents and 3-5.7 mean citations per document. The results for Civil Engineering are unusual in that within the second group this discipline has the best visibility and impact indicators despite the fact that one third of applicants were non-productive during the study period. This may indicate the existence of highly applied researchers, for whom more academic bibliometric indicators like citations may be inappropriate tools to assess activity and performance. Moreover, in Engineering the dissemination of results through conference proceedings is very important (63% in EEC; 60% in ECT; 31% in MNA)—in contrast with other disciplines where this document type accounts for approximately 4.5% of production—so any study that excludes them cannot assess these highly applied disciplines adequately.

Thirdly, the Social Sciences group (SSC, ECO, EDU, PSY) comprises two clearly distinctive trends. Economics and Psychology have 72% and 80% ratios of productive researchers, respectively; in Economics the rate of variation was 157%, rising from 15% of active researchers in the first year of the study to 40% in the last. In 2000, Economics had figures similar to those of Education or Social Science; at the end of the study period its profile looked more like that of Psychology. Economics and Psychology had production rates of 0.4 and 0.7, respectively. However, if we only include productive researchers, mean production for 2006 was 1.6 and 2.3, respectively. If this trend continues, in production both disciplines would become fully integrated into the experimental sciences' pattern of publication, although in terms of visibility indicators, they remain far behind the levels of the

fully established disciplines. While we might expect the Social Sciences and Education to follow the pattern demonstrated by Economics, the fact is that the indicators do not show substantial improvements over the study period. The heterogeneous thematic classification of the Social Sciences, which includes researchers in departments of Anthropology, Sociology, Geography or Library and Information studies, makes it difficult to reach conclusions applicable to all the disciplines. Less than a third of researchers have had one international publication during the study period and, moreover, the rates of variation are limited (from 9.6 % of productive researchers in the first year to 13 % in 2006). In Education, the productive researchers do increase substantially (112 %), although the very low initial figures facilitate this, showing a pattern of behavior that practically parallels that of the Social Sciences. In this category, the impact and visibility indicators present somewhat better figures. It is highly indicative that not one single article was recorded in the top3 journals in Education during the study period, showing the low international visibility of research in the discipline. These data lead us to reflect on the need to find alternative formulas and data sources with which to construct thresholds in these two disciplines from data gathered prior to 2006, at least, and based on methods proposed elsewhere (van Leeuwen, 2005; Torres-Salinas et al. 2009a). The broader WoS journal coverage of these disciplines in recent years (Thomson Reuters, 2008) could bring up differences in future studies.

The applied nature of the present study —aimed at facilitating the review process— led us to take the methodological decision to expressly reject the use of measures requiring additional calculations. Our final objective is to enable the reviewer to quickly and efficiently determine which productivity and impact thresholds the researcher attains in relation to the other applicants for project funding in any given discipline, and with a satisfactory degree of confidence. Establishing tools that facilitate the reviewers' task and improve the agility and efficiency of the Spanish R&D system does not in any way at all mean that bibliometric measures can or should replace the experts' analytical judgment. The ANEP itself recommends *not* penalizing young or inexperienced researchers, so the thresholds or data presented here should not be interpreted as «formulas» through the application of which x publications or citations would be worth y assessment points. Rather, we hope to provide guidelines that facilitate the fair assessment of the research merits of Spanish scientists. To this end, establishing dynamic time windows as a function of specific discipline-based characteristics or researcher types (junior or senior) is also intended to contribute to the fair assessment of candidates' CVs.

On the other hand, and despite our aforementioned main objective, the thresholds or reference tables we have constructed can also serve as benchmarks in academic and research institutions, providing frames that are more or less demanding according to specific needs and criteria. We are aware that the sample analyzed here is not necessarily representative of the research conducted in national centers as it is based on applications for project funding—which presupposes a relatively high, or at least above average standard of research for Spanish

institutions which are hampered by the high percentage of lecturers not actively involved in research. This is demonstrated by CNEAI data on 6-year research bonuses that highlights the fact that 10% of professors and 30% of tenured lecturers —posts that entail participation in research tasks— have never applied for, or never been awarded a single research bonus (CNEAI, 2005). However, institutions wishing to establish research-oriented careers on the basis of bibliometric parameters can use these data to construct their own reference parameters to meet specific internal needs.

To obtain more statistically robust results for observed production and citation, we suggest enrolling a wider-ranging study population that would include researchers applying in future funding rounds. In those categories with little production, or with a high number of non-productive researchers, this is practically obligatory, although alternative data sources —e.g., the National Research Council (CSIC) ISOC database, or the University of Granada IN-RECS database— could be used to evaluate the Social Sciences more fairly. An update of the present study would also provide more recent information on the internationalization of these disciplines as the WoS is including more Spanish journals, especially in the Social Sciences, perhaps making the use of alternative sources unnecessary. We suggest future studies should involve the coverage of document types such as monographs, which are especially important in some branches of the Social Sciences and the Humanities. In the latter, a count of international articles is clearly insufficient to assess researchers' merits.

6. Acknowledgements

The present research was conducted as part of the *Parametrización de los indicadores de citación a nivel nacional de acuerdo con las áreas ANEP (TIN2008-03180-E)* study, funded by Spain's National R&D Plan (Plan Nacional de I + D) 2008. The authors would like to thank Bryan J. Robinson (Department of Translation and Interpreting, University of Granada) for the translation of this text.

7. Refer ences

- Abramo, G., and D'Angelo, C. A. (2011). National-scale research performance assessment at the individual level. *Scientometrics*, vol. 86 (2), 347-364.
- Alonso-Arroyo, A.; Pulgarín, A., and Gil-Leiva, I. (2006). Análisis bibliométrico de la producción científica de la Universidad Politécnica de Valencia 1973-2001. *Revista Española de Documentación Científica*, vol. 29 (3), 345-363.
- Bordons, M.; Fernández, M., and Gómez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics*, vol. 53 (2), 195-206.

- Buela-Casal, G.; Bermúdez, M. P.; Sierra, J. C.; Quevedo-Blasco, R., and Castro, Á. (2010). Ranking de 2009 en investigación de las universidades públicas españolas. *Psicothema*, vol. 22, 171-179.
- Buela-Casal, G.; Medina, A.; Viedma, M.; Godoy, V.; Lozano, S., and Torres, G. (2004). Factor de impacto de tres revistas españolas de Psicología. *Psicothema*, vol. 16 (4), 680-688.
- Butler, L. (2003). Explaining Australia's increased share of ISI publications – the effects of a funding formula based on publication counts. *Research Policy*, vol. 32 (1), 143-155.
- CNEAI. (2005). Memoria sobre la situación del profesorado numerario con respecto a los sexenios de investigación. Año 2005.
- Camí, J. (1997). Impactolatría: diagnóstico y tratamiento. *Medicina Clínica*, vol. 109 (13), 515-524.
- Carpenter, M. P.; Gibb, F.; Harris, M.; Irvine, J.; Martin, B. R., and Narin, F. (1988). Bibliometric profiles for British academic institutions: An experiment to develop research output indicators. *Scientometrics*, vol. 14 (3-4), 213-233.
- Costas, R. (2008). *Análisis bibliométrico de la actividad científica de los investigadores del CSIC en tres áreas: Biología y Biomedicina, Ciencia de Materiales y Recursos Naturales. Una aproximación metodológica a nivel micro (Web of Science, 1994-2004)*. [Tesis doctoral] Universidad Carlos III.
- Costas, R., and Bordons, M. (2005). Bibliometric indicators at the micro-level: some results in the area of natural resources at the Spanish CSIC. *Research Evaluation*, vol. 14 (2), 110-120.
- Costas, R.; van Leeuwen, T. N., and Bordons, M. (2010). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: the effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology*, vol. 61 (8), 1564–1581.
- Cronin, B., and Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*, vol. 57 (9), 1275-1278.
- Cruz-Castro, L., and Sanz-Menéndez, L. (2007). Research Evaluation in transition: individual versus organisational assessment in Spain. En: Whitley, R., Gläser, J. (editors). *The Sociology of Sciences Yearbook: The Changing Governance of the Sciences. The Advent of the Research Evaluation Systems*. Springer, Dordrecht.
- Delgado-López-Cózar, E.; Jiménez-Contreras, E., and Ruiz-Pérez, R. (2009). España y los 25 grandes de la ciencia mundial en cifras (1992-2008). *El Profesional de la Información*, vol. 18 (1), 81-86.
- Garfield, E. (2003). The meaning of the Impact Factor. *International Journal of Clinical and Health Psychology*, vol. 3 (2), 363-369.
- González-De-Dios, J.; Valderrama-Zurian, J. C.; Gonzalez-Alcaide, G.; Sempere, A. P.; Bolanos-Pizarro, M., and Aleixandre-Benavent, R. (2009). Approximation to the impact of biomedical journals in neurological sciences: study of bibliometric indicators in the Journal Citation Reports-Science Citation Index 2006. *Revista de Neurología*, vol. 48 (3), 117-128.
- Gordillo, V.; González, J., and Muñoz, J. (2004). La evaluación de Proyectos de Investigación por la Agencia Nacional de Evaluación y Prospectiva. *Psicothema*, vol. 16 (3), 343-349.

- Jiménez-Contreras, E.; Delgado-López-Cózar, E., and Ruiz-Pérez, R. (2006). Producción española en biblioteconomía y documentación con visibilidad internacional a través del Web of Science (1995-2004). *El Profesional de la Información*, vol. 15 (5), 373-383.
- Jiménez-Contreras, E.; Moya-Anegón, F. D., and Delgado-López-Cózar, E. (2003). The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy*, vol. 32 (1), 123-142.
- Lawrence, P. (2002). Rank injustice. *Nature*, vol. 415 (6874), 835-836.
- Lewison, G.; Cottrell, R., and Dixon, D. (1999). Bibliometric indicators to assist the peer review process in grant decisions. *Research Evaluation*, vol. 8 (1), 47-52.
- Macías-Chapula, C.; Rodea-Castro, I.; Gutierrez-Carrasco, A., and Mendoza-Guerrero, J. (2004). Producción científica institucional y posicionamiento nacional: el caso del Hospital General de México. *Revista Española de Documentación Científica*, vol. 27 (4), 482-497.
- Meho, L. I., and Spurgin, K. M. (2005). Ranking the research productivity of library and information science faculty and schools: An evaluation of data sources and research methods. *Journal of the American Society for Information Science and Technology*, vol. 56 (12), 1314-1331.
- Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht; Springer.
- Moed, H.; Burger, W.; Frankfort, J., and van Raan, A. F. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, vol. 14 (3), 131-149.
- Moya Anegón, F.; Chinchilla Rodríguez, Z.; Corera Álvarez, E.; Vargas Quesada, B.; Muñoz Fernández, F., and Herrero Solana, V. (2005). Análisis de dominio institucional: la producción científica de la Universidad de Granada (SCI 1991-99). *Revista Española de Documentación Científica*, vol. 28 (2), 170-195.
- Oppenheim, C. (2007). Using the h-index to rank influential British researchers in information science and librarianship. *Journal of the American Society for Information Science and Technology*, vol. 58 (2), 297-301.
- Ruiz-Pérez, R.; Delgado-López-Cózar, E., and Jiménez-Contreras, E. (2002). Spanish name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *Journal of the Medical Library Association*, vol. 90 (4), 411-430.
- Ruiz-Pérez, R.; Jiménez-Contreras, E., and Delgado-López-Cózar, E. (2008). Complementos bibliométricos de Thomson Scientific en la Web: buenos, bonitos y gratuitos. *El Profesional de la Información*, vol. 17 (5), 559-563.
- Sanz-Menéndez, L. (1997). *Estado, ciencia y tecnología en España: 1939-1997*. Madrid; Alianza.
- Schubert, A., and Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics*, vol. 36 (3), 311-324.
- Seglen, P. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, vol. 314 (7079), 498-502.
- Thijs, B., and Glänzel, W. (2009). A structural analysis of benchmarks on different bibliometrical indicators for European research institutes based on their research profile. *Scientometrics*, vol. 79 (2), 377-388.

- Thomson Reuters. (2008). Thomson Scientific Begins Expansion of Web of Science. <http://science.thomsonreuters.com/press/2008/8445762/> [consulta: 02/07/2010].
- Torres-Salinas, D.; Delgado-López-Cózar, E., and Jiménez-Contreras, E. (2009a). Análisis de la producción de la Universidad de Navarra en revistas de Ciencias Sociales y Humanidades empleando rankings de revistas españolas y la Web of Science. *Revista Española de Documentación Científica*, vol. 32 (1), 22-39.
- Torres-Salinas, D.; Delgado-López-Cózar, E., and Jiménez-Contreras, E. (2009b). Rankings for departments and researchers within a university using two different databases: Web of Science versus Scopus. *Scientometrics*, vol. 80 (3), 761-774.
- Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, vol. 62 (1), 117-131.
- van Leeuwen, T. (2005). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible. *Scientometrics*, vol. 66 (1), 133-154.
- van Leeuwen, T. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation*, vol. 16 (2), 93-105.
- van Leeuwen, T., and Moed, H. (2005). Characteristics of Journal Impact Factors: The effect of uncitedness and citation distribution on the understanding of journal impact factors. *Scientometrics*, vol. 63 (2), 357-371.
- van Raan, A. F. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics*, vol. 36 (3), 397-420.
- van Raan, A. F. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, vol. 62 (1), 133-143.
- van Raan A. F., and van Leeuwen, T. (2002). Assessment of the scientific basis of interdisciplinary, applied research: Application of bibliometric methods in Nutrition and Food Research. *Research Policy*, vol. 31 (4), 611-632.

ESTUDIOS / RESEARCH STUDIES

Productividad e impacto de los investigadores españoles: umbrales de referencia por áreas científicas

Evaristo Jiménez-Contreras*, Nicolás Robinson-García*,
Álvaro Cabezas-Clavijo*

Resumen: Se presentan umbrales de referencia de producción e impacto científico de la investigación española con visibilidad internacional para las áreas definidas por la Agencia Nacional de Evaluación y Prospectiva (ANEP) en sus convocatorias. Tomando como población los solicitantes de proyectos del Plan Nacional de I+D 2007 ($n = 3.356$) se construyen tablas de referencia por percentiles que funcionan a modo de *benchmarks*, permitiendo efectuar comparaciones entre el comportamiento bibliométrico de un investigador y los registros de referencia en su área científica. Igualmente se ofrecen los datos de producción, impacto y visibilidad promedios para las áreas ANEP, y se discute el proceso de internacionalización de dichas áreas en el período 2000-2006, con una especial atención a las Ciencias Sociales. Finalmente, se sugiere el uso de umbrales de referencia como método de evaluación tanto para agencias financiadoras como para instituciones de investigación en sus procesos selectivos.

Palabras clave: Indicadores bibliométricos, umbrales de referencia, benchmarking, ANEP, ciencia, investigación científica, España.

Productivity and impact of Spanish researchers: reference thresholds within scientific areas

Abstract: Reference thresholds for the scientific production and impact of internationally visible Spanish research within the areas defined by the Spanish National Agency for Evaluation and Prospective (ANEP) are presented. These percentile reference tables are constructed from the population of researchers who applied for a project within the 2007 National R&D Plan ($n = 3.356$) and are to be used as benchmarks, permitting comparisons between researchers' bibliometric behavior and mean performance in their respective scientific disciplines. Data relating to mean production, impact and visibility

* EC3: Evaluación de la Ciencia y la Comunicación Científica, Departamento de Biblioteconomía y Documentación, Universidad de Granada. Correo-e: evaristo@ugr.es; elrobin@ugr.es; acabezasclavijo@gmail.com.

Recibido: 05-01-2011; 2.ª versión: 03-02-2011; aceptado: 04-02-2011.

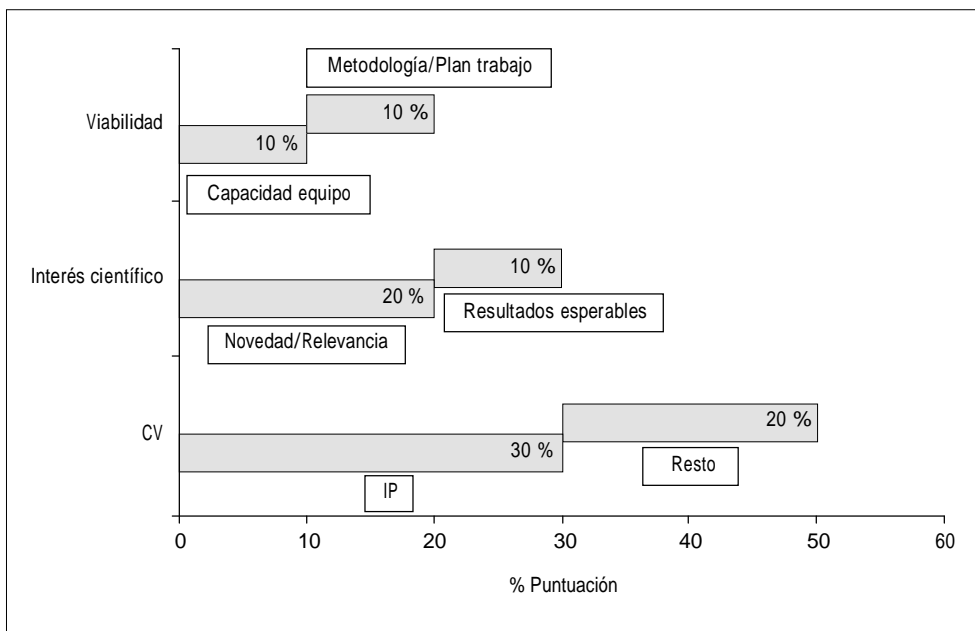
for each ANEP area is also presented. The internationalization process of these areas between 2000 and 2006 is discussed, with special emphasis on the Social Sciences. Finally, we recommend these reference thresholds be used by funding agencies and research institutions as instruments of evaluation in their selection processes.

Keywords: *Bibliometric indicators, reference thresholds, benchmarking, ANEP, science, scientific research, Spain.*

1. Introducción

La creciente importancia de los parámetros bibliométricos en los procesos de promoción universitaria, de asignación de becas o en la concesión de proyectos de investigación han generado en las instituciones la necesidad de contar con indicadores sólidos, ajustados a las diferentes áreas de evaluación, y que permitan la valoración justa de los candidatos en las concurrencias competitivas. A la sombra de la profesionalización de la actividad científica en España, en los últimos 25 años (Sanz-Menéndez, 1997; Cruz-Castro y Sanz-Menéndez, 2007), han aparecido una serie de instituciones encargadas de la gestión de la actividad científica y como parte de ellas, las agencias evaluadoras. Una de las más significativas, la Agencia Nacional de Evaluación y Prospectiva (ANEP) ha ido desarrollando un sistema de evaluación en el que se deslindan y valoran separadamente los aspectos de contenido, presupuestarios y curriculares. Desde su creación en 1986 el método usado para la evaluación de proyectos es el *peer review*; cada proyecto a evaluar se envía a dos revisores expertos, elegidos por el coordinador del área científica, que deben de realizar un informe con la correspondiente puntuación en cada apartado y una calificación final (Gordillo y otros, 2004). Si bien la ANEP realiza evaluaciones para diversas convocatorias tanto del Ministerio de Ciencia e Innovación, como de otros organismos, su actividad más importante es la que concierne las solicitudes del Plan Nacional de I + D (PN). Dicho Plan es la convocatoria más relevante para la mayoría de científicos españoles, que tienen la posibilidad de conseguir financiación por un período de tres años para la ejecución de propuestas científicas. Los criterios en que se basa dicha asignación de proyectos descansan fundamentalmente en la evaluación del currículum vitae del investigador principal (30 %) y del resto del equipo investigador (20 %) (figura 1). Una sólida experiencia, junto a un buen equipo investigador o una propuesta técnicamente aceptable asegura el éxito en dichas convocatorias selectivas. Para valorar la calidad del curriculum vitae (c.v.), es decir, el valor y repercusión de lo publicado, es muy frecuente recurrir, especialmente en las ciencias exactas y experimentales y cada vez más en las sociales al número de publicaciones internacionales (esto es, las recogidas en ISI Web of Science – WOS) y a los factores de impacto de las mismas, dicho de otra manera, recurrir a los indicadores bibliométricos.

FIGURA 1
*Puntuación para cada uno de los criterios de evaluación ANEP:
Convocatoria 2009*



El recurso a los factores de impacto es habitual en la ciencia española, que ha tomado esta medida prácticamente como criterio único a la hora de incentivar la producción. El establecimiento de un estímulo para la productividad como los sexenios de la CNEAI a finales de los ochenta, tomando como criterio básicamente los estándares de las ciencias puras (junto a otros factores como un marco legal más favorable a la investigación en las universidades públicas) generaron un incremento en la internacionalización de la ciencia española a lo largo de la década de los noventa (Jiménez-Contreras y otros, 2003). Dicho aumento se ha mantenido en esta década debido en gran parte, a la adecuación de los investigadores al patrón de publicación internacional, detectándose además una convergencia con los estándares de citación mundiales en la última etapa, tras años de registros sistemáticamente inferiores a las referencias internacionales (Delgado-López-Cózar y otros, 2009). Del mismo modo el uso de estos criterios ha generado amplias controversias (Camí, 1997; Bordons y otros, 2002; Lawrence, 2002), dado que los factores de impacto de las revistas de publicación no son medidas representativas del impacto individual de los trabajos o de los investigadores que los publican (Seglen, 1997). Esto se debe a las propias limitaciones metodológicas y conceptuales del indicador, como son por ejemplo; una ventana de citación de sólo dos años o una relación asimétrica entre el nú-

mero de citas que reciben los trabajos y el factor de impacto de las revistas donde éstos se publican (van Leeuwen y Moed, 2005). Sin embargo, a pesar de estas limitaciones, agencias, evaluadores y los propios científicos, toman una parte importante de sus decisiones científicas empleando este indicador. La razón última de este éxito (sin olvidar el ahorro de costes) es la asunción de que el factor de impacto, aunque no sea predictivo del impacto concreto que va a alcanzar un artículo, sí que representa la visibilidad, el prestigio y, en suma, las dificultades y controles que ha tenido que superar el autor para ver publicado su trabajo. Dicho de otro modo, el factor de impacto otorga un marchamo de calidad al trabajo que es bastante proporcional al propio impacto de la revista (Garfield, 2003).

A partir de aquí se ha producido una corriente de trabajos de corte empírico que han utilizado los factores de impacto como herramienta de trabajo y valoración (por ejemplo, Alonso-Arroyo y otros, 2006; Buela-Casal y otros, 2004; González-de-Dios y otros, 2009). Paralelamente algunas aportaciones han presentado propuestas o reflexiones de corte metodológico encaminadas, partiendo de los factores de impacto o de las citas reales u observadas, a salvar algunas de sus limitaciones especialmente las relacionadas con las comparaciones inter-categorías (Schubert y Braun, 1996). Por otro lado, se plantea igualmente el problema de resolver el procedimiento de normalización que permita las comparaciones cruzadas entre disciplinas, para lo que se presentan varias posibilidades. Uno de los procedimientos más sencillos son los rankings directos con cargo a uno o varios criterios: producción, financiación, impacto, o cualquier otro (Torres-Salinas y otros, 2009b; Buela-Casal y otros, 2010) con los que hay que tener especial cuidado debido a los múltiples factores metodológicos que inciden en su confección (van Raan, 2005). Otra posibilidad, es la elección de rangos percentiles, de modo que los agregados objeto de evaluación se sitúen en una escala que permite comparaciones inter-categorías (Lewison y otros, 1999; Costas y otros, 2010). Finalmente otra opción es el desarrollo de métodos basados en indicadores de referencia, en lo relativo tanto a producción, como impacto o colaboración, y convenientemente normalizados a fin de posibilitar las comparaciones entre áreas. En este grupo habría que situar el conjunto de indicadores propuesto por el CWTS de Leiden que constituyen una metodología propia para la evaluación de instituciones con propósito de servir a la política científica, y que ha sido aplicada con éxito a diversas instituciones, áreas geográficas e incluso campos temáticos (Moed y otros, 1985; van Raan y van Leeuwen, 2002).

Más allá de las metodologías usadas, en la literatura internacional se encuentran diversos trabajos encaminados a dibujar los perfiles bibliométricos de diferentes agregados. Por ejemplo, se han trazado los perfiles temáticos de instituciones académicas e investigadoras en Gran Bretaña (Carpenter y otros, 1988), México (Macías-Chapula y otros, 2004), o España (Moya-Anegón y otros, 2005; Torres-Salinas y otros, 2009a), así como de institutos de investigación en Europa (Thijs y Glanzel, 2009) en lo referente a análisis meso. Con respecto a la evalua-

ción individual de investigadores, es abundante la literatura que trata de establecer rankings disciplinares en función de criterios como el índice h (Cronin y Meho, 2006; Oppenheim, 2007), o la producción y citas recibidas (Meho y Spurgin, 2005; Jiménez-Contreras y otros, 2006) circunscribiéndonos únicamente al área de *Information Science*. El desarrollo de perfiles bibliométricos a nivel micro usando metodologías más complejas ha sido abordado recientemente en varios trabajos de Abramo y D'Angelo (2011), en el marco de las evaluaciones nacionales, y de Costas y colaboradores (Costas y Bordons, 2005; Costas, 2008; Costas y otros, 2010). En estos últimos trabajos se establecen perfiles o niveles de excelencia diferenciando tres niveles de análisis (producción, impacto y visibilidad) con el objetivo de ubicar a los investigadores en escalas o rango de excelencia o calidad investigadora. Esta metodología «bottom-up» (van Leeuwen, 2007) se caracteriza por la necesidad de tomar mayores precauciones tanto en la recogida de datos como en el procesamiento y desarrollo de los indicadores a usar, dado la alta distorsión que errores en la búsqueda y procesamiento pueden originar en los resultados finales. En general se puede decir que cuanto más cercano a los autores sea el análisis tanto más amplia tiene que ser la batería de indicadores empleados, así como mayores deben ser las prevenciones y precauciones a tomar, tanto en la recogida de datos como en el procesamiento y desarrollo de los indicadores a usar, que deben estar en consonancia con el uso último que se hará de esos datos.

Los evaluadores, expertos en la materia, sin duda conocen cuáles son los parámetros habituales de productividad y de *calidad* en sus áreas científicas, pero hasta el presente carecen de unos marcos de referencia o umbrales precisos que permitan situar a los investigadores objeto de evaluación en un contexto productivo construido a partir de datos empíricos y circunscrito a los estándares de la investigación nacional. El mismo escenario se presenta en lo que concierne al impacto de los trabajos publicados por dichos científicos. La construcción de unos marcos de referencia que permitan realizar análisis comparativos o evaluativos tanto para la producción como para la citación proporcionaría a los evaluadores unas guías o *benchmarks* en los que insertar el currículum vitae del investigador valorado, dotando así de mayor objetividad al proceso evaluador. De este modo se logra una mayor objetividad en el proceso de evaluación por pares, respondiendo a algunas de las lagunas que presenta este sistema (Gordillo y otros, 2004) y favoreciendo lo que se conoce en la literatura como *informed peer review* (Lewison y otros, 1999; van Raan, 1996).

Se ha debatido si el recurso a indicadores estándar ha modificado las prácticas de los propios científicos y propiciado prácticas desviadas o fraudulentas (Butler, 2003). Para solventar estos problemas se ha propuesto, además de evitar el uso de «fórmulas» para la evaluación de los investigadores (Moed, 2005), la utilización combinada de indicadores bibliométricos y de evaluadores expertos (Weingart 2005). Dentro de este contexto de construcción de marcos de referencia para las áreas científicas, no podemos dejar de citar el producto de Thomson Reuters, Essential Science Indicators (ESI), herramienta que aunque no muy em-

pleada, resulta de ayuda inestimable para establecer umbrales y promedios de citación en función de la categoría a la que pertenecen los trabajos y del tiempo en que fueron publicados, neutralizando así las dos variables principales que pueden afectar a la medición del impacto de un artículo. Esta herramienta son los llamados *baselines*; promedios y percentiles que permiten establecer la posición relativa de un artículo de investigación dentro de su especialidad según el número de citas recibido y el tiempo transcurrido desde su publicación. Dichos indicadores proporcionan el número de citas promedio por categoría y año de un artículo así como el rango percentil de un trabajo también por año y categoría. Estas tablas, desagregadas en 22 áreas de tamaños variables, y donde se excluyen las Humanidades se construyen a partir del conjunto de revistas incluidas en la *Web of Science* y se actualizan con carácter bimestral (Ruiz-Pérez y otros, 2008), originando unos marcos temporales dinámicos.

2. Objetivos

El objetivo global que se ha planteado en este estudio ha sido el de construir los perfiles bibliométricos de los solicitantes de proyectos dentro del marco del Plan Nacional de I + D, para elaborar a partir de esos datos unos umbrales de referencia que sirvan a los evaluadores para la toma de decisiones basadas en la evidencia bibliométrica. Dicho de otro modo, se pretende establecer cuáles son los estándares de publicación e impacto de los investigadores españoles o mejor dicho, de aquellos que están activos en términos de solicitud de proyectos.

Este objetivo genérico se puede desglosar en los siguientes objetivos específicos:

- Construir un conjunto de tablas de referencia, similar al ofrecido por los *Essentials Science Indicators*, adaptadas al caso español, y relativos tanto al rendimiento en producción como en impacto.
- Definir los principales indicadores bibliométricos de producción, productividad, impacto y visibilidad por especialidades.
- A nivel metodológico, testar la validez de la metodología propuesta, con el objeto de mantenerla en el futuro o introducir modificaciones a fin de hacer lo más objetiva posible la evaluación de los CV de los investigadores.

3. Material y métodos

Se plantea un análisis descriptivo retrospectivo del rendimiento bibliométrico de los investigadores principales españoles solicitantes de proyectos de investigación en las distintas convocatorias del Plan Nacional de I + D en el año 2007. La población analizada abarca un total de 3356 investigadores, pertenecientes a

todas las áreas establecidas por la ANEP, exceptuando Derecho, Filología y Filosofía e Historia y Arte. Dicha información fue facilitada por la propia agencia. Las áreas con mayor número de solicitudes fueron, la de Biología Fundamental (12 % del total), seguida por la Química y la Economía. Respecto al cálculo de indicadores, y con el objeto de validar el diseño metodológico de nuestro estudio, se procedió a determinar para cada una de las áreas el porcentaje de investigadores con al menos un artículo publicado en el período de estudio. Para este conjunto de investigadores, se buscó su producción científica en las bases de datos del *Web of Science (WoS)* producidas por la empresa *Thomson Reuters*, en su versión *on-line*. Dichas fuentes de datos son *Science Citation Index Expanded (SCI)*, *Social Sciences Citation Index (SSCI)*, *Arts and Humanities Citation Index (A&HCI)*, *Conference Proceedings Citation Index – Science (CPCI)*.

La búsqueda se ejecutó de forma manual para cada uno de los investigadores, recuperándose todos los registros comprendidos entre los años 2000 y 2006. Dicha búsqueda fue realizada por documentalistas expertos, teniendo en cuenta las diferentes formas que pueden adoptar los nombres españoles en las bases de datos del ISI (Ruiz-Pérez y otros, 2002) y contemplando las diferentes afiliaciones que un investigador ha podido tener a lo largo del período (por ejemplo, debido a estancias en centros internacionales). La producción de cada uno de los investigadores se almacenó en formato texto, exportándose posteriormente al gestor de referencias bibliográficas *ProCite 5*. Una vez almacenada de forma conjunta toda la producción, dicha producción se trasladó a una base de datos relacional en formato *MS Access 2003*, junto a los datos asociados a los investigadores (nombre, apellidos, código de proyecto, tipología de investigador, institución y centro). Las tipologías documentales analizadas fueron *article*, *review*, *letter*, *editorial material* y *proceedings paper*. Igualmente en dicha base de datos se introdujo la información relativa al factor de impacto de las revistas de publicación, descargada desde los *Journal Citation Reports (JCR)*, producto también de *Thomson Reuters*. Del mismo modo se incluyó el número de citas recibidas por cada uno de los artículos recuperados. Para dicha tarea, se utilizó la función *Create Citation Report* disponible desde la página de resultados de la versión online del *WoS*. Así, una vez identificada la producción de un autor, y a través de dicha opción, se descargaron las citas recibidas para cada uno de los registros en función del año de los artículos citantes. El marco temporal del conjunto de citantes abarca el período 2000-2008. Finalmente se pusieron en relación los documentos citantes con los citados, en la base de datos relacional creada al efecto. La recopilación de datos se inició en febrero de 2009, finalizando en mayo de 2010.

Posteriormente, y mediante consultas en el *software Access 2003*, se calcularon los diversos indicadores, para cada una de las áreas ANEP. Ésta categorización se ha mantenido relativamente estable en el tiempo, si bien en fechas recientes se han introducido algunas modificaciones en áreas de Ciencias de la Vida. La categorización por áreas que aquí presentamos (tabla I) es la que estaba vigente en la convocatoria de 2007.

TABLA I
Áreas ANEP y acrónimos empleados

Acrónimo	Área
AGR	Agricultura
BMC	Biología fundamental y de sistemas
BMED	Biomedicina
BVA	Biología vegetal, animal y ecología
COM	Tecnología electrónica y de las comunicaciones
CS	Ciencias sociales
CT	Ciencias de la Tierra
ECO	Economía
EDU	Ciencias de la educación
FFA	Fisiología y farmacología
FI	Física y ciencias del espacio
GAN	Ganadería y pesca
ICI	Ingeniería civil y arquitectura
IEL	Ingeniería eléctrica, electrónica y automática
IME	Ingeniería mecánica, naval y aeronáutica
INF	Ciencias de la computación y tecnología informática
MCLI	Medicina clínica y epidemiología
MTM	Matemáticas
PS	Psicología
QMC	Química
TA	Ciencia y tecnología de alimentos
TM	Ciencia y tecnología de materiales
TQ	Tecnología química
<i>no analizada</i>	Derecho
<i>no analizada</i>	Filología y filosofía
<i>no analizada</i>	Historia y arte

La explotación estadística de los datos analizados se ha realizado con el software *SPSS 15.0 para Windows*. Los indicadores calculados son:

• **Producción**

- **Investigadores productivos.** Investigadores con al menos un documento publicado en el período.

- **Ndoc/invest.** Promedio de documentos por investigador.
- **TV. Tasa de variación.** Diferencia porcentual entre la producción registrada en el primer año de estudio (2000) y el último (2006).
- **Baselines (umbrales de producción).** Frecuencias acumuladas de producción por investigador por percentiles. Este indicador nos da información sobre el número de artículos necesarios para que un investigador se sitúe en el percentil 90, 75 y 50 de la distribución de la producción científica según años de publicación.

• **Visibilidad**

- **PromFI.** Promedio de Factor de Impacto.
- **Top3.** Producción en revistas top 3 por categorías JCR.
- **Q1.** Producción en revistas del primer cuartil por categorías JCR.

• **Impacto**

- **PromCit.** Promedio de citas por documento.
- **%Ndoc citados.** Porcentaje de documentos citados.
- **Baselines (umbrales de citación):** Frecuencias acumuladas de citación por percentiles. Este indicador nos da información sobre el número de citas necesarias para que un artículo se sitúe en el percentil 90, 75 y 50 de la distribución de la producción científica según años de publicación.

4. Resultados

4.1. Producción

En total 3356 investigadores (excluidos de este cómputo los de Humanidades) participaron en la convocatoria del PN 2007, repartidos como queda recogido en la tabla II.

Los datos obtenidos se mueven en un rango entre el 99,58 de investigadores productivos en Química, y el 29,45 % en Ciencias Sociales. La mayoría de las categorías alcanzan una ratio de investigadores productivos por encima del 90 %. La excepción la encontramos en las cuatro áreas de Ciencias Sociales, en los que se perciben patrones bien diferenciados. Mientras Economía y Psicología presentan valores productivos mayores año a año (con una espectacular tasa de crecimiento en el primer caso), Educación (también con tasas relativas de crecimiento muy altas) y Ciencias Sociales permanecen por debajo del 35 %. Ingeniería Civil, por su parte, se aparta del patrón del resto de las ingenierías, con un tercio de investigadores sin producción. Áreas con un patrón de publicación internacional consolidado como Física, Biología Fundamental o Matemáticas experimentan leves tasas de variación. Por otro lado, algunas ramas tecnológicas e ingenieriles apuntan crecimientos altos, probablemente debido a la ampliación de la cobertura de actas de congreso en la base de datos de referencia.

TABLA II
*Investigadores totales, productivos y tasa de variación 2000-2006
 según áreas ANEP*

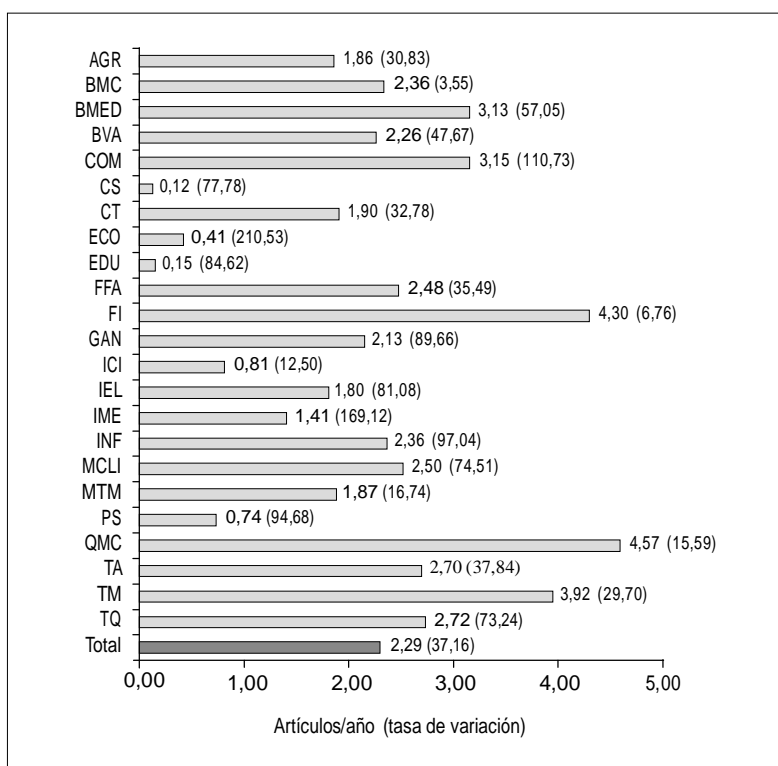
Área	Investigadores	Productivos	% Productivos	TV 00-06
AGR	132	127	96,21	24,39
BMC	405	400	98,77	0,64
BMED	124	123	99,19	16,13
BVA	164	160	97,56	25,00
COM	139	131	94,24	40,00
CS	146	43	29,45	35,71
CT	141	130	92,20	17,44
ECO	182	131	71,98	157,14
EDU	119	41	34,45	112,50
FFA	145	144	99,31	10,53
FI	180	178	98,89	1,31
GAN	96	92	95,83	34,43
ICI	61	40	65,57	33,33
IEL	86	84	97,67	36,17
IME	91	85	93,41	60,00
INF	152	143	94,08	45,68
MCLI	51	49	96,08	11,43
MTM	131	130	99,24	5,38
PS	169	136	80,47	56,86
QMC	237	236	99,58	9,00
TA	128	126	98,44	9,38
TM	171	167	97,66	17,74
TQ	106	100	94,34	31,75
Total	3.356	2.996	89,27	19,94

Respecto a los datos de producción, se ha determinado el total de documentos producidos en término medio por investigador así como el promedio anual según las diferentes categorías ANEP. Con el objeto de observar la tendencia cronológica se ha calculado la tasa de variación, que da idea de la progresiva incorporación de cada vez más investigadores al patrón de publicación internacional. En este caso, se observan claramente las diferencias productivas entre áreas, desde los 4,3-4,6 documentos por investigador y año que publican físicos y químicos, a los apenas 0,1 artículos anuales en Ciencias Sociales y Educación. Las bajísimas cifras de partida en estas áreas explican el alto incremento relativo a la par que la baja productividad ISI de sus investigadores. Las mayores tasas de variación se producen en algunas ingenierías y áreas tecnológicas así como

en la Economía (figura 2). Uno de los factores que influyen en las diferentes productividades por áreas es la tasa de colaboración de los trabajos. Así, el índice de coautoría, o número medio de firmantes por artículo se situó en un rango de entre 8,2 para el área de Física (excluyendo del cómputo artículos con más de 250 autores) y 2,5 para Economía. En cuanto a la moda, la más alta es de 6 autores para Biomedicina, Medicina Clínica y Epidemiología y Ganadería y Pesca mientras que la menor, de un autor, corresponde a las Ciencias Sociales.

FIGURA 2

Promedio de producción por investigador y año, y tasa de variación 2000-2006 (entre paréntesis) según áreas ANEP



Dado que los promedios productivos pueden estar sesgados por la distribución asimétrica de las productividades investigadoras se calculó de manera complementaria los umbrales de producción para cada año del período de estudio así como para el período completo. Se seleccionaron los percentiles 50, 75, 80, 90 y 99 detectándose dos tendencias bien diferenciadas: por un lado, la mayoría de áreas arrojan resultados consistentes, y crecientes de manera exponencial entre percentiles; es decir el esfuerzo productivo que hay que realizar para situarse en los

percentiles superiores de la distribución es cada vez mayor, de manera que estas *benchmarks* proporcionan umbrales de exigencia, en función de los criterios de los evaluadores. Los valores tienden a mantenerse a lo largo del tiempo, con crecimientos productivos en los años finales del período. En áreas como Química o Física es necesaria una producción por encima de los 20 artículos (en 7 años) para situarse en la mediana de la categoría. Por otro lado, en las áreas de Ciencias Sociales (e Ingeniería Civil) el alto número de investigadores no productivos condiciona los resultados obtenidos, generando comportamientos anómalos, como que un investigador sin producción se sitúe incluso en el P75 de la distribución, para un cálculo a tres años, en áreas como Educación. La tabla III muestra los datos por percentiles para una evaluación de los currícula investigadores a 3, 5 y 7 años.

TABLA III

Umbral (baseline) de producción a 3,5 y 7 años para los percentiles 50, 75 y 90 según áreas ANEP

Área/ Percentil	3 años			5 años			7 años		
	P50	P75	P90	P50	P75	P90	P50	P75	P90
AGR	4	7	12	7	11	21	10	16	29
BMC	6	9	14	9	15	24	13	21	33
BMED	7	12	18	12	20	29	17	28	41
BVA	5	9	13	9	14	22	12	20	31
COM	7	14	21	11	24	35	16	33	49
CS	0	0	1	0	1	1	0	1	2
CT	3	6	11	6	11	19	8	15	26
ECO	1	2	3	1	3	5	2	4	7
EDU	0	0	2	0	1	3	0	1	4
FFA	6	10	14	10	16	24	14	23	33
FI	10	18	25	16	31	42	23	43	59
GAN	5	9	10	9	16	17	12	22	24
ICI	1	3	8	1	4	13	2	6	18
IEL	4	8	12	7	13	20	10	18	28
IME	3	6	10	6	10	17	8	14	24
INF	6	9	15	9	15	24	13	21	34
MCLI	5	9	17	8	15	29	11	21	40
MTM	5	7	11	8	12	19	11	17	26
PS	1	3	6	2	5	9	3	7	13
QMC	10	17	27	17	29	46	24	40	64
TA	7	11	17	11	19	28	16	26	39
TM	9	16	26	14	27	43	20	38	60
TQ	6	9	16	9	16	27	13	22	38
Total	5	9	16	8	15	26	11	21	37

* *Ejemplo:* Un investigador en el área de Agricultura con una producción de 12 artículos en 5 años se situaría entre el 10 y el 25% de investigadores más productivos del área (entre P75 y P90).

4.2. Visibilidad e impacto

Los indicadores de impacto y visibilidad relativos a artículos en revistas con factor de impacto, y artículos citados corren paralelos a la ratio de investigadores productivos. El mayor desfase se produce en la categoría de Ciencias Sociales, donde hay un 11 % de investigadores que han publicado en revistas que aún no disponían de factor de impacto para el año de publicación, reflejo éste de publicaciones recién incorporadas a la base de datos WoS. En Educación hay un 8,4 % de investigadores en la misma situación. En el área de Ingeniería Civil, y en menor medida en las otras ingenierías se da también un cierto desfase, debido a la publicación en actas de congreso, tipología ésta que no dispone de factor de impacto. El mismo fenómeno se produce si efectuamos la comparación entre los investigadores productivos citados y los no citados. En este caso la Economía y la Psicología se incorporan a este grupo. Globalmente hay un 4 % de investigadores productivos cuya producción no ha sido citada. Tecnología de Alimentos, es por su parte la única área, junto a Ciencias de la Tierra, donde todos los investigadores productivos han logrado al menos una cita para alguno de sus trabajos (tabla IV).

TABLA IV
Porcentaje de investigadores según indicadores de visibilidad e impacto por áreas ANEP

Área	% productivos	% FI	% Q1	% Top3	% citados
AGR	96,21	93,18	79,55	50,00	92,42
BMC	98,77	98,77	97,78	68,15	96,05
BMED	99,19	99,19	95,16	77,42	97,58
BVA	97,56	97,56	90,85	54,27	96,95
COM	94,24	91,37	69,06	33,81	90,65
CS	29,45	18,49	8,90	3,42	17,12
CT	92,20	91,49	82,98	49,65	92,20
ECO	71,98	66,48	29,67	13,74	62,64
EDU	34,45	26,05	10,08	0,00	24,37
FFA	99,31	99,31	97,93	72,41	98,62
FI	98,89	98,33	95,00	71,11	97,22
GAN	95,83	95,83	93,75	57,29	93,75
ICI	65,57	55,74	40,98	19,67	52,46
IEL	97,67	91,86	58,14	25,58	89,53
IME	93,41	89,01	67,03	31,87	86,81
INF	94,08	90,13	58,55	21,05	88,16

TABLA IV (continuación)

Área	% productivos	% FI	% Q1	% Top3	% citados
MCLI	96,08	96,08	82,35	70,59	92,16
MTM	99,24	99,24	71,76	16,03	96,95
PS	80,47	77,51	46,15	20,12	73,37
QMC	99,58	99,58	97,89	62,03	98,73
TA	98,44	98,44	97,66	72,66	98,44
TM	97,66	97,66	92,40	69,01	95,91
TQ	94,34	93,40	88,68	51,89	93,40
Total	89,27	87,10	74,82	46,51	85,40

Respecto a los indicadores que miden el rendimiento en términos de visibilidad en revistas top3 y Q1, las diferencias entre categorías son amplias. En áreas como Química, Fisiología o Biología Fundamental prácticamente todos los investigadores han publicado algún artículo en revistas del 1er cuartil. En la parte inferior de este indicador se sitúan las áreas de Ciencias Sociales. Globalmente, 3 de cada 4 investigadores publicaron en revistas de 1er cuartil al menos uno de sus artículos en el período de estudio. Respecto al indicador top3, cerca de la mitad de investigadores también publicaron en alguna de las tres revistas con mayor impacto en sus áreas. Los valores abarcan desde el 77 % de Biomedicina hasta los valores nulos en Educación, donde ninguno de los solicitantes de proyectos publicó en revistas top3.

Cuando se trata de medir el rendimiento a nivel de artículo (tabla V), más de la mitad de estos se publican en revistas de primer cuartil, mientras que este porcentaje cae al 12 % cuando se mide el rendimiento en las revistas de máximo impacto (top3). Son destacables los registros alcanzados por Biología Fundamental y Física en Q1, mientras que en top3 destacan Biomedicina y Tecnología de Alimentos, con una ratio en torno al 20 %. El porcentaje de artículos citados alcanza sus mayores cotas en Química, Fisiología y Biología Fundamental. Estas áreas son también, junto a la Biomedicina las que alcanzan mayor promedio de citas por documento. Por su parte, Ingeniería Eléctrica y las dos áreas de Tecnologías de la Información y Comunicación alcanzan los mayores niveles de documentos no citados, posiblemente por la importancia de las actas de congreso en estas disciplinas. Educación, con 2 menciones por documento de término medio es el área que presenta parámetros de citación más modestos.

Al igual que con la producción, se han determinado los valores percentílicos (P50, P75, P90) para la citación. Dada la influencia que el factor tiempo tiene en este cómputo, se han realizado los cálculos en función de los años transcurridos desde la publicación. En la tabla VI se muestran los P50, 75 y 90, correspondiendo el cálculo con marco temporal de 9 años a los artículos recogidos en el año 2000, mientras que para el cálculo con marco temporal de 3 años se han tomado

TABLA V
Indicadores de visibilidad e impacto según áreas ANEP

Área	FI promedio	% Q1	% Top3	Prom Citas	% citados
AGR	1,70	45,69	12,53	7,60	74,56
BMC	4,82	67,44	14,62	21,43	92,14
BMED	3,86	59,49	20,06	17,14	87,21
BVA	2,22	45,59	9,02	9,43	87,07
COM	1,31	37,83	9,74	4,12	51,58
CS	2,11	38,82	12,94	3,94	61,60
CT	1,94	55,63	14,59	11,05	86,38
ECO	0,71	21,97	6,78	3,47	70,04
EDU	0,79	23,53	0,00	2,21	60,94
FFA	3,72	58,29	10,69	15,39	92,70
FI	3,09	65,29	14,49	12,31	79,62
GAN	1,87	57,69	11,57	9,41	87,10
ICI	0,95	47,78	12,96	5,75	66,57
IEL	0,93	30,95	9,89	2,97	47,65
IME	1,09	44,30	8,19	5,83	73,70
INF	0,72	14,62	2,99	3,07	49,70
MCLI	2,82	48,56	16,07	11,13	84,66
MTM	0,67	23,52	1,97	4,02	71,79
PS	1,57	26,29	7,99	6,93	81,54
QMC	2,89	58,95	8,33	14,16	93,56
TA	1,82	63,11	19,36	10,61	90,10
TM	1,98	56,16	13,69	9,94	85,75
TQ	1,90	59,76	14,31	10,58	90,05
Total	2,63	53,46	11,96	11,59	82,02

los artículos publicados en 2006 (en todos los casos, el período de recogida de citas abarca hasta 2008). A nivel general, y tomando como referencia los 9 años del período más largo calculado, un artículo necesitaría 9 citas para situarse en la mediana de la distribución. Esta cifra se eleva a 16 en el caso de la Biología Fundamental, mientras que es 0 para Ciencias Sociales e Ingeniería Eléctrica. Las muestras en que se basan los resultados a 9 años (el año 2000 es el que registra menor producción del período analizado) genera casos anómalos, como que en algunas categorías los artículos que han dispuesto de 7 años para ser citados (los producidos en 2002) marquen umbrales de citación más elevados que los que han gozado de 9 años para ser mencionados.

TABLA VI

Umbral (baselines) de citación según marcos temporales de 3,5,7 y 9 años para los percentiles 50, 75 y 90 por áreas ANEP

Área/ Percentil	3 años			5 años			7 años			9 años		
	P50	P75	P90	P50	P75	P90	P50	P75	P90	P50	P75	P90
AGR	1	4	9	4	9	15	7	16	28	7	15	30
BMC	5	11	20	10	21	41	14	30	60	16	36	75
BMED	4	10	18	8	18	38	11	24	44	10	28	57
BVA	2	4	8	5	11	21	7	14	22	8	23	43
COM	0	2	6	1	5	11	1	6	15	1	6	15
CS	1	2	4	3	6	12	4	8	17	0	2	8
CT	3	7	11	7	13	34	7	15	30	10	19	35
ECO	0	1	2	2	4	7	2	9	14	4	7	13
EDU	1	2	3	2	4	7	1	7	11	1	2	7
FFA	4	8	14	9	18	37	13	25	43	12	27	55
FI	3	7	15	5	14	27	7	17	35	10	21	40
GAN	3	5	8	7	11	18	5	11	22	9	18	33
ICI	0	2	6	3	6	11	5	9	21	5	20	33
IEL	0	1	4	1	5	11	0	6	13	0	5	10
IME	1	3	6	3	8	17	4	11	17	5	10	20
INF	0	1	3	0	2	7	1	3	10	2	7	17
MCLI	3	6	10	7	17	24	6	15	30	9	23	48
MTM	1	2	4	2	5	9	3	7	14	3	8	16
PS	2	4	6	5	9	17	5	11	23	4	12	23
QMC	4	8	15	8	15	28	11	20	34	13	24	42
TA	3	6	10	7	14	22	9	18	30	11	22	42
TM	2	6	12	5	11	20	6	13	28	7	16	30
TQ	3	6	12	6	13	25	9	17	31	10	18	40
Total	2	6	12	5	13	24	7	17	33	9	21	42

* *Ejemplo:* Un artículo en el área de Matemáticas con 4 citas recibidas en 3 años se situaría entre el 10 % de artículos más citados del área (P90). El mismo artículo (con 4 citas) en el área de Biología Molecular no llegaría a ubicarse entre el 50% más citado en dicha área.

5. Discusión

La primera constatación que se desprende de los datos analizados es la existencia de tres tipos de áreas: el primer grupo lo componen aquellas que desde el principio están plenamente incorporadas a la investigación internacional, y que corresponden a las áreas de Ciencias (AGR, BMC, BMED, BVA, CT, FFA, FI, GAN,

MCLI, MTM, QMC, TA, TM, TQ). Todas cumplen con un patrón de producción internacional, con una inmensa mayoría de solicitantes que son investigadores productivos. En el caso de áreas como BMC o FI el porcentaje de investigadores productivos se mantiene inalterable durante el período. Dentro de este primer grupo se significan áreas como Agricultura, Biología Vegetal, Animal y Ecología, o Ganadería y Pesca, que han culminado su proceso de internacionalización en el período analizado. Este primer grupo se destaca también por unas productividades entre 1,9 y 4,6 documentos por investigador y año, así como por publicar (excepto las Matemáticas) más de un 45 % de artículos en revistas del primer cuartil. Los umbrales de producción y citación se mantienen, además, relativamente estables año a año, aunque se aprecia en cuanto a producción unos umbrales más exigentes en los últimos años del período, probablemente debido a la incorporación al sistema de solicitantes con menor experiencia, junto al aumento general de la productividad española en el período.

En segundo lugar se sitúa el grupo de las Ingenierías y TIC (COM, ICI, IME, IEL, INF), que presentan perfiles diferenciados. Por un lado, hay una incorporación de investigadores al sistema, con una productividad en un rango de entre 0,8 (ICI) y 3,1 (COM) documentos por investigador y año. Los indicadores de visibilidad describen un porcentaje de documentos en el primer cuartil de entre un 14 y un 48 %, y un promedio de citas por documento de entre 3 y 5,7. En este caso es anómalo el comportamiento de Ingeniería Civil, ya que es el área de este segundo grupo que describe los mejores indicadores de visibilidad e impacto, a pesar de que un tercio de los solicitantes por esta vía no han sido productivos en el período. Esto puede indicarnos la existencia en este grupo de investigadores con una naturaleza muy aplicada, para los cuáles el uso de los indicadores bibliométricos de corte más académico como las citas pueden no ser la vía más adecuada para valorar su actividad y rendimiento. Además, se constata que la difusión de resultados en las ingenierías mediante la publicación en actas de congresos (63 % en IEL; 60 % en COM; 31 % en IME) es —al contrario que en el resto de áreas donde la media de publicación en actas de congresos se sitúa en torno al 4,5 % de toda la producción— muy importante por lo que un análisis que hubiera excluido estas tipologías documentales sería insuficiente para calibrar dichas áreas de naturaleza aplicada.

En tercer lugar, y dentro de las Ciencias Sociales (CS, ECO, EDU, PS) se perciben dos ritmos bien diferenciados. Por un lado, Economía y Psicología cuentan con una ratio de investigadores productivos de un 72 y un 80 % respectivamente, con una tasa de variación para la Economía del 157 %, pasando de apenas un 15 % de investigadores activos en el primer año de estudio a un 40 % en el último año analizado. Economía, que partía en 2000 con valores similares a los de Educación o Ciencias Sociales se asemeja más al final de esta etapa al perfil que describe Psicología. Las tasas de producción son de 0,4 y de 0,7 para Economía y Psicología, respectivamente, aunque si tomamos únicamente los investigadores productivos estos promedios se sitúan en 1,6 y 2,3 para el año 2006, lo que demuestra que, de seguir esta tendencia, tanto la Psicología como

la Economía se incorporarían de lleno al patrón de publicación de las ciencias experimentales, al menos en cuanto a producción, ya que en lo relativo a la visibilidad, se quedan lejos de los indicadores que muestran las áreas consolidadas. Respecto a las Ciencias Sociales y la Educación, de las cuales cabría pensar que pueden seguir el camino emprendido por la Economía en estos años, lo cierto es que los indicadores no muestran grandes avances a lo largo del período. Lo heterogéneo de la clasificación temática Ciencias Sociales, donde caben investigadores de departamentos de Antropología, Sociología, Geografía, Periodismo o Biblioteconomía, hace complicado expresar unas conclusiones comunes a estas áreas de conocimiento. No llegan a un tercio los investigadores con alguna publicación internacional en el período, y además las tasas de variación son escasas (de un 9,6 % de investigadores productivos el primer año a un 13 % en 2006). Respecto a Educación, los investigadores productivos sí experimentan importantes alzas (112 %), aunque las bajísimas cifras de partida facilitan este estadístico, mostrando un comportamiento prácticamente paralelo al descrito en Ciencias Sociales. Los indicadores de impacto y visibilidad presentan cifras algo mejores en esta última categoría. Cabe destacar como dato muy sintomático que no se registra ni un solo artículo en revistas top3 para el período en la categoría de Educación, lo que señala la baja visibilidad internacional de la investigación en el área. Estos datos nos hacen reflexionar acerca de la necesidad de buscar fórmulas y fuentes de datos alternativas para la confección de umbrales en estas dos áreas, al menos con los datos registrados hasta 2006, y en línea con las metodologías propuestas por otros autores (van Leeuwen, 2005; Torres-Salinas y otros, 2009a) si bien la ampliación de la cobertura de revistas en WoS para estas áreas en los últimos años (Thomson Reuters, 2008) podría arrojar diferencias en futuros estudios.

En el marco metodológico, es preciso resaltar que dada la naturaleza aplicada del estudio, destinada a facilitar la labor de los evaluadores, se renunció expresamente al uso de medidas que requirieran de cálculos adicionales para su elaboración. El objetivo final es que el evaluador pueda determinar de forma rápida y eficaz, en qué umbrales de productividad e impacto se sitúa un investigador en relación al resto de solicitantes de proyectos en una determinada área, con un margen suficiente de confianza. El establecimiento de herramientas encaminadas a facilitar la labor de los evaluadores y a mejorar la agilidad y eficiencia del sistema español de I+D no supone en ningún momento que las medidas bibliométricas puedan o deban reemplazar el juicio de los expertos en sus análisis. La propia ANEP recomienda no penalizar a los investigadores jóvenes o con poca experiencia, por lo que los umbrales o datos aquí presentados no deben interpretarse como una «fórmula» por el cual a x número de publicaciones o citas le correspondería x puntos en la evaluación. Se trata más bien de proporcionar unas guías o directrices que puedan permitir la valoración justa de los méritos investigadores de los científicos españoles. El establecimiento en este contexto de ventanas temporales dinámicas en función de las características concretas de cada área o del tipo de investigador evaluado (investigadores junior

o senior) también pretende contribuir a la valoración justa de los CV de los candidatos.

Por otro lado, y pese a que el objetivo principal de este trabajo ya se ha expresado, los umbrales o tablas de referencia que se han construido también pueden servir de *benchmarks* en instituciones académicas y de investigación, proporcionando marcos más o menos exigentes en función de las necesidades y criterios de cada una. Somos conscientes de que la muestra aquí analizada no tiene por qué ser representativa de la investigación realizada en los centros de investigación nacionales, ya que se basa en solicitantes de proyectos de investigación, es decir, ya se presupone un nivel investigador relativamente alto, o al menos superior al promedio de las instituciones españolas, que se ven lastradas por un porcentaje importante de profesores que en la práctica no ejercen la investigación. Así lo demuestran los datos de sexenios de investigación facilitados por la CNEAI, donde llama poderosamente la atención que un 10 % de los catedráticos, y un 30 % de los profesores titulares, cuerpos que tienen encomendadas las labores investigadoras, no se hayan presentado o no hayan obtenido nunca un tramo de investigación (CNEAI, 2005). Las instituciones que deseen, sin embargo, establecer carreras investigadoras según parámetros bibliométricos pueden tomar estos datos para construir sus propios parámetros de referencia en función de las peculiaridades domésticas.

Cabe mencionar que, a fin de lograr resultados estadísticamente más robustos de producción y citación observados, sería conveniente ampliar la población del estudio con los investigadores solicitantes en nuevas convocatorias. Para las categorías donde hay escasa producción, o poseen un alto número de investigadores no productivos, esto es prácticamente obligado, aunque también cabría la posibilidad de proponer otras fuentes de datos alternativas como las bases de datos ISOC, del CSIC, o IN-RECS, de la Universidad de Granada para la evaluación más justa de las áreas de Ciencias Sociales. La actualización de este trabajo debe proporcionarnos también información acerca de la internacionalización de éstas áreas en las fechas más recientes ya que la entrada en la fuente de datos de referencia, Web of Science, de un mayor número de revistas españolas, especialmente en los ámbitos de Ciencias Sociales pueden tornar innecesario el empleo de fuentes alternativas. Por otra parte, también sería interesante abordar la cobertura de otros tipos documentales como lo son las monografías, pensando especialmente en algunas ramas de las Ciencias Sociales y abarcando también las Humanidades, área en la cual el cómputo de artículos internacionales es a todas luces insuficiente para la valoración de los méritos de los investigadores en este campo.

6. Agradecimientos

Esta investigación se ha desarrollado dentro de la acción complementaria *Parametrización de los indicadores de citación a nivel nacional de acuerdo con las áreas ANEP* (TIN2008-03180-E), financiada por el Plan Nacional de I + D 2008.

7. Referencias

- Abramo, G., y D'Angelo, C. A. (2011). National-scale research performance assessment at the individual level. *Scientometrics*, vol. 86 (2), 347-364.
- Alonso-Arroyo, A.; Pulgarín, A., y Gil-Leiva, I. (2006). Análisis bibliométrico de la producción científica de la Universidad Politécnica de Valencia 1973-2001. *Revista Española de Documentación Científica*, vol. 29 (3), 345-363.
- Bordons, M.; Fernández, M., y Gómez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics*, vol. 53 (2), 195-206.
- Buela-Casal, G.; Bermúdez, M. P.; Sierra, J. C.; Quevedo-Blasco, R., y Castro, Á. (2010). Ranking de 2009 en investigación de las universidades públicas españolas. *Psicothema*, vol. 22, 171-179.
- Buela-Casal, G.; Medina, A.; Viedma, M.; Godoy, V.; Lozano, S., y Torres, G. (2004). Factor de impacto de tres revistas españolas de Psicología. *Psicothema*, vol. 16 (4), 680-688.
- Butler, L. (2003). Explaining Australia's increased share of ISI publications – the effects of a funding formula based on publication counts. *Research Policy*, vol. 32 (1), 143-155.
- CNEAI. (2005). Memoria sobre la situación del profesorado numerario con respecto a los sexenios de investigación. Año 2005.
- Camí, J. (1997). Impactolatría: diagnóstico y tratamiento. *Medicina Clínica*, vol. 109 (13), 515-524.
- Carpenter, M. P.; Gibb, F.; Harris, M.; Irvine, J.; Martin, B. R., y Narin, F. (1988). Bibliometric profiles for British academic institutions: An experiment to develop research output indicators. *Scientometrics*, vol. 14 (3-4), 213-233.
- Costas, R. (2008). Análisis bibliométrico de la actividad científica de los investigadores del CSIC en tres áreas: Biología y Biomedicina, Ciencia de Materiales y Recursos Naturales. Una aproximación metodológica a nivel micro (*Web of Science, 1994-2004*). [Tesis doctoral] Universidad Carlos III.
- Costas, R. y Bordons, M. (2005). Bibliometric indicators at the micro-level: some results in the area of natural resources at the Spanish CSIC. *Research Evaluation*, vol. 14 (2), 110-120.
- Costas, R.; van Leeuwen, T. N., y Bordons, M. (2010). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: the effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology*, vol. 61 (8), 1.564-1.581.
- Cronin, B., y Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*, vol. 57 (9), 1.275-1.278.
- Cruz-Castro, L., y Sanz-Menéndez, L. (2007). Research Evaluation in transition: individual versus organisational assessment in Spain. En: Whitley, R. y Gläser, J. (editors). *The Sociology of Sciences Yearbook: The Changing Governance of the Sciences. The Advent of the Research Evaluation Systems*. Springer, Dordrecht.
- Delgado-López-Cózar, E.; Jiménez-Contreras, E., y Ruiz-Pérez, R. (2009). España y los 25 grandes de la ciencia mundial en cifras (1992-2008). *El Profesional de la Información*, vol. 18 (1), 81-86.

- Garfield, E. (2003). The meaning of the Impact Factor. *International Journal of Clinical and Health Psychology*, vol. 3 (2), 363-369.
- González-De-Dios, J.; Valderrama-Zurian, J. C.; Gonzalez-Alcaide, G.; Sempere, A. P.; Bolanos-Pizarro, M., y Aleixandre-Benavent, R. (2009). Approximation to the impact of biomedical journals in neurological sciences: study of bibliometric indicators in the Journal Citation Reports-Science Citation Index 2006. *Revista de Neurología*, vol. 48 (3), 117-128.
- Gordillo, V.; González, J., y Muñoz, J. (2004). La evaluación de Proyectos de Investigación por la Agencia Nacional de Evaluación y Prospectiva. *Psicothema*, vol. 16 (3), 343-349.
- Jiménez-Contreras, E.; Delgado-López-Cózar, E., y Ruiz-Pérez, R. (2006). Producción española en biblioteconomía y documentación con visibilidad internacional a través del Web of Science (1995-2004). *El Profesional de la Información*, vol. 15 (5), 373-383.
- Jiménez-Contreras, E.; Moya-Anegón, F. D., y Delgado-López-Cózar, E. (2003). The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy*, vol. 32 (1), 123-142.
- Lawrence, P. (2002). Rank injustice. *Nature*, vol. 415 (6874), 835-836.
- Lewis, G.; Cottrell, R., y Dixon, D. (1999). Bibliometric indicators to assist the peer review process in grant decisions. *Research Evaluation*, vol. 8 (1), 47-52.
- Macías-Chapula, C.; Rodea-Castro, I.; Gutierrez-Carrasco, A., y Mendoza-Guerrero, J. (2004). Producción científica institucional y posicionamiento nacional: el caso del Hospital General de México. *Revista Española de Documentación Científica*, vol. 27 (4), 482-497.
- Meho, L. I., y Spurgin, K. M. (2005). Ranking the research productivity of library and information science faculty and schools: An evaluation of data sources and research methods. *Journal of the American Society for Information Science and Technology*, vol. 56 (12), 1314-1331.
- Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht; Springer.
- Moed, H.; Burger, W.; Frankfort, J., y van Raan, A. F. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, vol. 14 (3), 131-149.
- Moya Anegón, F.; Chinchilla Rodríguez, Z.; Corera Álvarez, E.; Vargas Quesada, B.; Muñoz Fernández, F., y Herrero Solana, V. (2005). Análisis de dominio institucional: la producción científica de la Universidad de Granada (SCI 1991-99). *Revista Española de Documentación Científica*, vol. 28 (2), 170-195.
- Oppenheim, C. (2007). Using the h-index to rank influential British researchers in information science and librarianship. *Journal of the American Society for Information Science and Technology*, vol. 58 (2), 297-301.
- Ruiz-Pérez, R.; Delgado-López-Cózar, E.; Jiménez-Contreras, E. (2002). Spanish name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *Journal of the Medical Library Association*, vol. 90 (4), 411-430.
- Ruiz-Pérez, R.; Jiménez-Contreras, E., y Delgado-López-Cózar, E. (2008). Complementos bibliométricos de Thomson Scientific en la Web: buenos, bonitos y gratuitos. *El Profesional de la Información*, vol. 17 (5), 559-563.
- Sanz-Menéndez, L. (1997). *Estado, ciencia y tecnología en España: 1939-1997*. Madrid; Alianza.

- Schubert, A. y Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics*, vol. 36 (3), 311-324.
- Seglen, P. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, vol. 314 (7079), 498-502.
- Thijs, B., y Glänzel, W. (2009). A structural analysis of benchmarks on different bibliometrical indicators for European research institutes based on their research profile. *Scientometrics*, vol. 79 (2), 377-388.
- Thomson Reuters. (2008). Thomson Scientific Begins Expansion of Web of Science. <http://science.thomsonreuters.com/press/2008/8445762/> [consultado el 2 de julio de 2010].
- Torres-Salinas, D.; Delgado-López-Cózar, E., y Jiménez-Contreras, E. (2009a). Análisis de la producción de la Universidad de Navarra en revistas de Ciencias Sociales y Humanidades empleando rankings de revistas españolas y la Web of Science. *Revista Española de Documentación Científica*, vol. 32 (1), 22-39.
- Torres-Salinas, D.; Delgado-López-Cózar, E., y Jiménez-Contreras, E. (2009b). Rankings for departments and researchers within a university using two different databases: Web of Science versus Scopus. *Scientometrics*, vol. 80 (3), 761-774.
- Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, vol. 62 (1), 117-131.
- van Leeuwen, T. (2005). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible. *Scientometrics*, vol. 66 (1), 133-154.
- van Leeuwen, T. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation*, vol. 16 (2), 93-105.
- van Leeuwen, T. y Moed, H. (2005). Characteristics of Journal Impact Factors: The effect of uncitedness and citation distribution on the understanding of journal impact factors. *Scientometrics*, vol. 63 (2), 357-371.
- van Raan, A. F. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics*, vol. 36 (3), 397-420.
- van Raan, A. F. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, vol. 62 (1), 133-143.
- van Raan A. F., y van Leeuwen, T. (2002). Assessment of the scientific basis of interdisciplinary, applied research: Application of bibliometric methods in Nutrition and Food Research. *Research Policy*, vol. 31 (4), 611-632.