



Coverage, field specialisation and the impact of scientific publishers indexed in the Book Citation Index

Daniel Torres-Salinas

Applied Medical Research Centre, University of Navarra, Pamplona, Spain

Nicolas Robinson-Garcia

Department of Information and Communication, University of Granada, Granada, Spain

Juan Miguel Campanario

Department of Physics, University of Alcalá, Alcalá de Henares, Spain, and

Emilio Delgado López-Cózar

Department of Information and Communication, University of Granada, Granada, Spain

Abstract

Purpose – The aim of this study is to analyse the disciplinary coverage of Thomson Reuters' Book Citation Index database focusing on publisher presence, impact and specialisation.

Design/methodology/approach – The authors conducted a descriptive study in which they examined coverage by discipline, publisher distribution by field and country of publication, and publisher impact. For this purpose the Thomson Reuters' subject categories were aggregated into 15 disciplines.

Findings – Humanities and social sciences comprise 30 per cent of the total share of this database. Most of the disciplines are covered by very few publishers mainly from the UK and USA (75.05 per cent of the books), in fact 33 publishers hold 90 per cent of the whole share. Regarding publisher impact, 80.5 per cent of the books and chapters remained uncited. Two serious errors were found in this database: the Book Citation Index does not retrieve all citations for books and chapters; and book citations do not include citations to their chapters.

Originality/value – There are currently no studies analysing in depth the coverage of this novel database which covers monographs.

Keywords Databases, Citation analysis, Publishers, Book Citation Index, Monographs, Thomson Reuters

Paper type Research paper



Introduction

The launch in October 2011 of Thomson Reuters' newest product for scientific literature – the Book Citation Index (hereafter BKCI) – may well be a historic milestone in the field of scientometrics or just another failure in the race to develop tools capable of evaluating monographs. Books have long been neglected by bibliometricians (Giménez-Toledo and Román-Román, 2009) mainly due to the lack of data sources and the ways in which they differ from scientific journal literature in terms of their ageing (Glänzel and Schoepflin, 1999; Hicks, 2004; Nederhof, 2006). Although many approaches have been considered before, especially focusing on the fields of the arts and humanities as well as the social sciences (Torres-Salinas and Moed, 2009; White *et al.*, 2009; Linmans, 2010; Zuccala and van Leeuwen, 2011) where this type of document plays a greater role than in other disciplines (Williams *et al.*, 2009), no definite solution has fully convinced the research community.

In this sense the launch of this product has an even higher impact in the field, as Thomson Reuters' Web of Science, which is highly valued by the scientific community, not only was the first provider of citation indexes but, in their own words, covers only “top tier international and regional journal literature” (Testa, 2010). Furthermore this database has become the main pillar on which many countries have built their whole research policy system. Fully developed by Eugene Garfield (2007) in 1972, the Science Citation Index and the subsequent Social Science Citation Index along with their Journal Citation Reports in which journals are ranked according to the widely known Journal Impact Factor (hereafter JIF) (Garfield and Sher, 1963) have become pivotal within the scientific communication system. Not only has this database dominated the field as the main source of multidisciplinary scientific literature for nearly 50 years, but its JIF has been established as the legitimate yardstick for measuring prestige among scientific journals. However it has long been criticised for its incapacity to fully develop a source capable of overcoming the lack of coverage of the social sciences and humanities due to the important role books play in these disciplines (Hicks, 2004). This weakness was in fact, already acknowledged by Garfield (1996) himself when he stated:

From the perspective of the social scientist or humanities scholar, the failure to include monographs as sources in the ISI citation indexes may be a drawback in drawing conclusions about the impact of certain work [...] Undoubtedly, the creation of a Book Citation Index is a major challenge for the future and would be an expected by-product of the new electronic media with hypertext capability!

However since 2004, the launch of Scopus – a multidisciplinary database which includes bibliometric data owned by Elsevier, the largest scientific publisher in the world – as well as the emergence of Google Scholar in the same year, signified a great threat to Thomson Reuters' Web of Science's long monopoly. Soon enough many studies emerged comparing Web of Science's coverage with that of these new scientific sources (Meho and Yang, 2007) with a special focus on those fields less well-covered by this database (Moed, 2005). In this line of work, Sivertsen and Larsen (2012) analysed the complete publication output data of Norway, finding that, while journals in the humanities and social sciences are quite diverse and influenced by national factors, book publishing is confined to a few publishers. Lacking fully reliable resources to assess publication types other than journals, Bar-Ilan (2010) performed a microscopic study comparing citation counts for a single book in Scopus, Web of Science and

Google Scholar, concluding that the latter offered better results. In a similar paper Baneyx (2008) compared citation metrics in the humanities and social sciences according to the Web of Science and Google Scholar, arguing that the former is unsuitable for these fields due to the lack of indexed monographs.

Here again, the importance of considering other publication types appears as an unsolved issue. Hicks (2004) comments on the importance of monographs for assessment in the social sciences and humanities, arguing that a restriction to journals offers a biased view, as demonstrated by Glänzel and Schoepflin (1999). In this sense it is important to mention the interesting project (Martin *et al.*, 2010) developed by different European countries which aims at exploring the possible development of a database that captures the many idiosyncrasies of these fields. Yet these fields are not the only ones substantially affected by this restriction, even those such as medicine where journals are the main channel of communication may also be influenced when other publication types are excluded (Lewison, 2004).

Literature review

Due to its youth, few studies can be found in the literature referring to the use of the BKCI for evaluation purposes or describing its internal characteristics; coverage, limitations, etc. To date, only two studies have been found. Leydesdorff and Felt (2012) analyse the citation rates of books, book chapters and edited volumes and compare the results offered by the BKCI with those of the other citation indexes. In the other study Torres-Salinas *et al.* (2012) propose the development of “Book Publishers Citation Reports” and analyse the strengths and weaknesses of such attempts in the social sciences and humanities fields. These types of seminal studies examining coverage, caveats and limitations are considered very useful as they serve to validate the accuracy and reliability of sources. The launch of Scopus (Elsevier) and Google Scholar and Books have led to the emergence of many studies analysing these alternative databases and their advantages and weaknesses when compared with the Thomson Reuters citation indexes (e.g. Kulkarni *et al.*, 2009; Kousha *et al.*, 2011). Yet books and book chapters still seem to lack widely accepted databases by funding agencies from which to retrieve bibliometric data (Edwards, 2012). Thus the importance of monographs in the social sciences and the arts and humanities has lead bibliometricians and research policy managers to search for information resources which would provide the necessary tools to assess researchers from these areas.

Many studies have been conducted analysing the role of books in research monitoring and alternative citation-enhanced databases which could be used to retrieve bibliometric data. For instance Kousha and Thelwall (2009) use Google Books Search to compare citations from books with citations from journal articles, concluding that the former could complement and even replace the latter in the fields of the social sciences and the arts and humanities. In a further study Kousha *et al.* (2011) compare the citations to a sample of books submitted to the UK Research Assessment Exercise retrieved from Google Scholar and Books with those from Scopus. Moreover Chen (2012) compares the coverage of Google Books with that of OCLC WorldCat. These studies position Google Scholar and Books as possible resources for assessing monographs. However the launch of the Book Citation Index by Thomson Reuters adds a new alternative to the current landscape, and as such must be analysed as a first

step to any consideration as an option for bibliometric and evaluative purposes as well as a new information resource to add to library collections.

Research questions and target audience

Since the emergence of the BKCI has been received by the community with great expectations, we here conduct an in-depth study of its coverage and briefly explore the disciplinary distribution by publisher, output and impact. Our main aim is to analyse BKCI coverage by focusing on identifying the main publishers represented. More specifically we address the following research questions:

- (1) How is the BKCI distributed by discipline and country? Which are the main publishers at each level of analysis?
- (2) What is the profile of the main BKCI publishers in terms of their output and disciplinary focus?
- (3) What impact do the mainstream publishers make in terms of their citation rates and which publishers can therefore be considered more prestigious and influential?

This type of study in which a new information source is evaluated and analysed is quite common in the literature because they provide valuable information not just for evaluation or bibliometric purposes, but also for researchers when choosing tools for information retrieval or even academic librarians who must be especially careful and selective when acquiring and subscribing to new databases. Therefore the target audiences of this study are:

- *Bibliometricians*. This type of study allows them to consider the data sources they will use when conducting their studies, considering their coverage, reliability and functionalities. Moreover it can support their findings when understanding some of the phenomena they analyse in order to interpret the results they obtain.
- *Librarians and researchers*. The results of this particular study may help them to decide on the adequacy of the BKCI as an answer to their users' demands or as a data source. Indeed they often ignore the exact content and scope of the databases they are offered. This study may be a good starting point to understand the potential use of this database.
- *Scientific publishers*. This study particularly focuses on assessing how they are represented in the BKCI. How they appear and how they perform when compared with competitors would be of special interest to them.
- *Prospective authors*. The present paper offers a general view which will provide them with information on the coverage of fields of the BKCI and the main publishers represented in it.
- *Evaluation agencies*. If considered as an evaluation tool, as other Thomson Reuters' citation indexes already are, evaluation agencies should be informed of the strengths and weaknesses of the BKCI before including it as a new tool to be used in their research assessment exercises. The BKCI may also provide a useful framework from which to learn when is a book highly cited, the citation threshold by area and so on.

Methods

In this paper we conduct an analysis of Thomson Reuters' Book Citation Index for 2005-2012. In May 2012 we downloaded all BKCI records and created a relational database for data processing and to calculate the indicators. Publisher names were normalised as many had variants that differed as a function of the location of their head offices in each country. For instance, Springer uses variants such as Springer-Verlag Wien, Springer-Verlag Tokyo, Springer Publishing Co., among others. The 249 subject categories, to which records from BKCI were assigned, were also restructured into the 15 disciplines proposed by Moed (2005) in a study of Web of Science journal coverage. However 2 per cent of the total share could not be included as the records were not assigned to any subject category. In Table I we describe the indicators used in our study.

Indicator	Name	Definition
No. or percent of books	Book	Records indexed as document type "book" in the BKCI
No. or percent of book chapters	Book chapter	Records indexed as document type "book chapter" in the BKCI
No. of publishers	ScPub	Scientific publishers assigned to books and book chapters
Concentration Index	GINI	The Gini Index is usually used in the field of economics. Here it is defined as an indicator for measuring the concentration of publishers covering each discipline and also for measuring publishers' concentration in the 15 fields defined, in order to interpret the level of specialisation. It gives values between 0 and 1: 0 indicates no concentration in a publisher or discipline; 1 indicates concentration in one publisher or discipline. For this indicator, we only considered the document type "book"
Field Normalised Citation Score	FNCI	The indicator FNCI or Crown is an oriented-field normalised citation score average developed at the CWTS at Leiden University. According to the definition given by Lundberg (2007), it is calculated dividing the average number of received citations for a group of publications (CPP in CWTS terminology) with the average number that could be expected for publications of the same type, from the same year, published within the same field (FCSm in CWTS terminology). Dividing the actual citation rate with the expected citation rate CPP/FCSm we obtain the FNCI. In this study FNCI has been used to measure publisher's impact. We have used the 249 BKCI subject categories, allowing us to compare publishers from different areas. A value above 1 is a value above the world's average
Activity Index	Activity Index	The share of a publisher's papers in a specific discipline, divided by the database share of papers in that discipline. A value above 1 means greater activity in the discipline than that expected from database distribution. This indicator shows how much publishers specialise in a certain field

Table I.
Set of bibliometrics indicators calculated to characterise scientific publishers in the Book Citation Index

Results

This section is structured as follows. First we briefly introduce a general perspective of the BKCI. Then we present a subsection for each of the research questions. We analyse the disciplinary coverage of the database and each country's representation by publisher affiliation. Then we focus on publisher presence, analysing who contributes most to the BKCI and their disciplinary profile. Finally we present a brief descriptive analysis of publisher impact in terms of citation data provided by Thomson Reuters. We must point out that the results shown are related to the actual coverage of the BKCI and not to the book market sector. Therefore the results of this section reflect the main coverage characteristics of this database.

The BKCI has a total of 408,700 records comprising 29,618 books (162 with no ISBN) and 379,082 book chapters (2,807 with no ISBN), averaging 13.78 book chapters per book. From the total number of books, 8,221 have ISSNs. This means that some of the records considered to be books by Thomson Reuters are actually serials. Regarding their publication date, books seem to be homogeneously distributed with ranges between 4 and 6 per cent each year, except for 2011 when the document share was largest (19 per cent of the database). In fact, in 2009, 2010, and 2011 similar numbers of books have been indexed (5,720, 5,798 and 5,517 respectively), which means numbers seem to have stabilised. Regarding book chapters, the distribution per year varies from 4 per cent in 2009 to 9 per cent in 2005 with 2012 being the year when the greatest share of book chapters can be found (21 per cent of the total).

Coverage of scientific publishers by discipline, country and language

Data was distributed among the 249 Web of Science subject categories. In Table II we present records from 15 research fields corresponding with those analysed by Moed (2005). An effort has been made in the selection process to cover those fields considered by Moed to have moderate coverage, that is, the humanities and arts and other social sciences. Each of these disciplines has around 30 per cent of the total share, taking into account books or book chapters. In fact they are the fields where more publishers have been indexed (143 in humanities and 114 in other social sciences). The field with the third highest share of the total database is engineering, with around 12 per cent and 83 different publishers.

When focusing on the distribution of publishers by discipline through their concentration index – which shows the distribution of books within fields for a certain publisher – we observe that most of these are covered by few publishers. “Other social sciences – medicine and health” presents a higher distribution of records per publisher with a Gini Index of 0.707, while most records in engineering and humanities and arts are distributed over very few publishers presenting a Gini Index of 0.901 and 0.900, respectively. If we observe the three top publishers per area, we find that only for “biological sciences – animals and plants”, and “other social sciences – medicine and health” do the top three publishers cover less than half of the content. For the rest of the areas, most records are concentrated in only three publishers. Applied physics and chemistry followed by engineering are the two fields where these publishers concentrate a higher share of the total with 82 per cent and 80 per cent, respectively. Springer seems to be the most multidisciplinary publisher as it is present in all fields, followed by Nova Science, which is one of the top three in eight of the fifteen disciplines.

Table II.
Book Citation Index
coverage and main
publishers per scientific
fields in 2005-2012

Scientific field	No. books	No. book chapters	% books	% book chapters	No. ScPub	GINI	Top three publishers in the field according to no. of books contributing to the field
Other social sciences	8,045	96,744	27.9	26.2	114	0.882	Palgrave 25%, Routledge 25%, Springer 12%
Humanities & arts	8,333	96,007	28.9	26.0	143	0.900	Palgrave 26%, Routledge 15%, Springer 10%
Engineering	3,446	42,895	11.9	11.6	83	0.901	Springer 68%, Nova Science 7%, CRC 5%
Economics	2,843	36,124	9.8	9.8	64	0.859	Springer 22%, Palgrave 20%, Routledge 17%
Clinical medicine	1,633	32,325	5.7	8.8	59	0.851	Springer 47%, Nova Science 13%, Humana Press 10%
Biological sciences – humans	1,544	22,515	5.3	6.1	68	0.826	Springer 44%, Elsevier 10%, Nova Science 7%
Molecular biology and biochemistry	1,119	19,098	3.9	5.2	38	0.815	Elsevier 25%, Springer 25%, Humana Press 24%
Biological sciences – animals	1,089	16,904	3.8	4.6	70	0.783	Springer 26%, Cabi Pub.13%, Nova Science 9%
Mathematics	1,522	16,718	5.3	4.5	55	0.858	Springer 58%, Birkhauser 7%, Siam 4%
Geosciences	1,189	16,139	4.1	4.4	78	0.810	Springer 38%, Nova Science 10%, Geological Soc 7%
Applied physics and chemistry	1,087	13,853	3.8	3.8	31	0.823	Springer 52%, Woodhead 15%, Transtech 15%
Other social sciences – medicine & health	671	9,074	2.3	2.5	51	0.707	Springer 18%, Nova Science 16%, Palgrave 12%
Chemistry	766	8,016	2.7	2.2	29	0.821	Springer 47%, Royal Society Chem. 17%, Nova Science 8%
Physics & astronomy	501	5,740	1.7	1.6	32	0.788	Springer 60%, Nova Science 8%, Annual Reviews 6%

Regarding publisher affiliation (see Figure 1), we must emphasise that countries are assigned according to where the book was published. That means, for instance, that books from Springer can be published in several countries. We find a strong bias towards English-speaking countries. In fact the US and England represent 56 per cent of publishers (194) and 75.05 per cent of the whole database taking into account books as publication type. However, when analysing the distribution by region, we observe that European countries represent 62.4 per cent of records while North America represents 36.42 per cent, Asia and Oceania 1 per cent and South America is only represented by Brazil with 0.03 per cent. When only focusing on publishers and not records, we see that The USA is by far the country with more BKCI indexed publishers, with three times more than second-placed England, with 50. The third-ranked country is Germany with 25 publishers, half as many as England. Regarding the distribution by languages (see Figure 2), the database is greatly biased towards the English language as many publishers affiliated to The Netherlands or Germany publish in English and not in the language of the publication country (e.g. Elsevier). This can be confirmed when observing the distribution by languages where around 96 per cent of the whole database are books in English followed by German books (2.44 per cent). The rest of the languages do not even reach 1 per cent of the total share of the BKCI.

Publisher production and specialisation

In Figure 3 we present five BKCI publishers according to their concentration index (GINI), showing whether or not they have a specialised or a multidisciplinary focus. These publishers account for more than 60 per cent of the BKCI in terms of books or book chapters. Therefore they can be considered the mainstream publishers in the database. As suggested by Table I, we find that Springer and Nova Science have



Figure 1.
Distribution of the Book
Citation Index according
to each publisher's
affiliation country
2005-2012

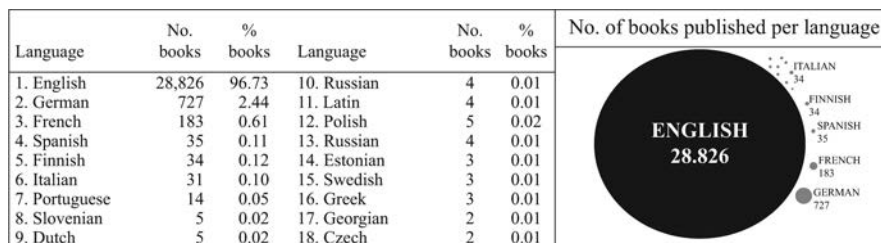


Figure 2.
Distribution of the Book
Citation Index according
to language 2005-2012

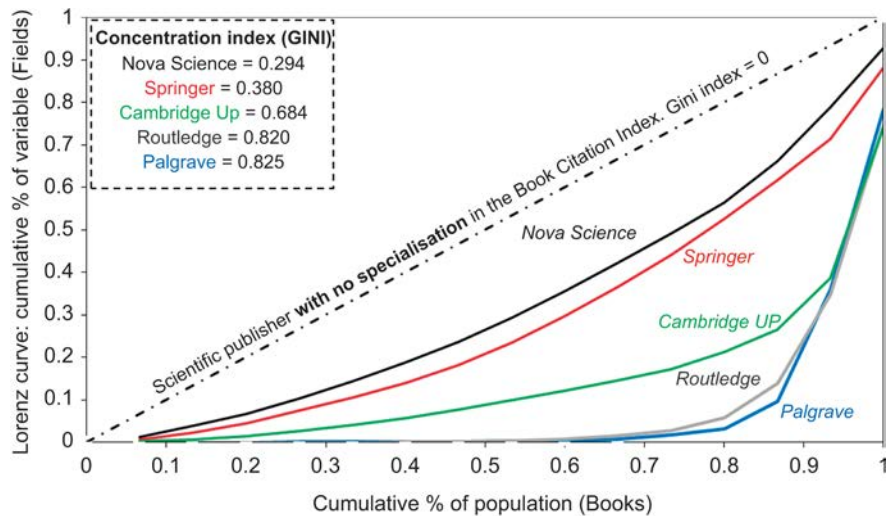


Figure 3.
Level of specialisation for
the top five publishers in
the Book Citation Index
2005-2012

Notes: GINI INDEX provides values between 0 and 1; 0 indicates no concentration in any specific discipline, i.e. a multidisciplinary publisher, and it is represented with a soft curve (e.g. Nova Science); 1 indicates concentration in any specific discipline, that is, a specialised publisher and it is represented with a sharp curve (e.g. Palgrave)

multidisciplinary profiles with values of 0.380 and 0.294, respectively. On the other hand Palgrave with a GINI score of 0.825 and Routledge with 0.820 are publishers that specialise in highly specific disciplines as their acute curves show. Finally Cambridge University Press is somewhere in between with a GINI of 0.684.

In Table III we present the top 33 BKCI publishers ranked by the number of books published. They represent 90 per cent of the total coverage and each has at least 100 books indexed in the database. By far the most prolific publishers in terms of books are Springer (7,783), Palgrave (4,164) and Routledge (3,369). However, if we analyse the disciplines they mainly specialise in, we observe different patterns. While Nova Science, Springer, Blackwell Science and Annual Reviews have multidisciplinary profiles that never surpass 0.4 of their concentration index, 18 of the top 33 publishers have highly specialised profiles which all pass the 0.8 concentration index barrier. Of these, only six present a profile focused on health sciences while the other 12 specialise in social sciences and humanities. Humana Press for instance has an activity index of 14.02 for molecular biology and chemistry, but the highest rate is that of Woodhead Publishing with an activity index of 30.64. Palgrave and Routledge, the two second-placed mainstream publishers in BKCI, are highly specialised in the social sciences and humanities.

Impact

The 408,700 records indexed by the BKCI have received a total of 680,680 citations; this means an average of 1.67 citations per record. Citation counts for books or book chapters include citations from BKCI but also from all the rest of the Web of Science databases, for example AH&CI, SSCI or SCI. However, as described in Table IV, 80.5 per cent of the total publications have received no citations and 1 per cent of the

Publisher *	Main country	No. books	% Book	No. book chapters	% book chapters	G:INI	Main specialisation fields – Activity Index **
Springer	The Netherlands	7,783	26.28%	101,975	26.90%	0.380	ENG 2.45
Palgrave	England	4,164	14.06%	45,350	11.96%	0.825	SOC 1.81
Routledge	England	3,399	11.48%	42,017	11.08%	0.820	SOC – 2.15
Nova Science	The USA	1,318	4.45%	14,168	3.74%	0.294	PSY – 1.68
							PHY AP
							MED CL 2.78
							PSY 2.18
Cambridge Univ Press	England	1,095	3.70%	14,566	3.84%	0.684	PSY 1.64
							SOC 0.74
							SOC MED – 0.71
Walter De Gruyter & Co.	Germany	962	3.25%	13,429	3.54%	0.872	MAT 0.69
							SOC 0.38
Elsevier	The Netherlands	733	2.47%	8,247	2.18%	0.621	PHY AP
							BIO HU 4.13
							CHE 2.83
Princeton Univ Press	The USA	715	2.41%	7,857	2.07%	0.763	HUM 1.85
Edward Elgar Publishing	England	616	2.08%	8,055	2.12%	0.818	SOC 1.62
							SOC 1.20
Univ California Press	The USA	602	2.03%	7,288	1.92%	0.801	HUM 1.85
Humana Press Inc.	The USA	543	1.83%	11,883	3.13%	0.821	SOC 1.53
							BIO HU 2.96
							ME CL 5.81
							SO ME 1.20
							SO ME 1.38
Univ Pennsylvania Press	The USA	465	1.57%	4,025	1.06%	0.869	ECO 0.37
							SO ME 0.28
Crc Press-Taylor & Francis	The USA	373	1.26%	5,801	1.53%	0.506	ENG 3.71
							BIO AP 2.44
Emerald Group	England	319	1.08%	4,152	1.10%	0.824	SO ME
							HUM 0.28
Woodhead Publ Ltd	England	309	1.04%	5,566	1.47%	0.790	ENG 1.64
							BIO MO 0.26
Information Age	England	287	0.97%	3,641	0.96%	0.848	PSY 1.68
Blackwell Science Publ	England	281	0.95%	4,813	1.27%	0.397	ECO 1.32
							PHY AP 2.72
Annual Reviews	The USA	252	0.85%	5,738	1.51%	0.368	SO ME
							PHY AP 3.42
							BIO MO 2.30
Cabi Publishing-C A B Int	England	236	0.80%	4,075	1.07%	0.715	PHY AP 3.60
							BIO MO 2.33
							SOC 1.02

(continued)

Table III.

Publisher *	Main country	No. books	% Book	No. book chapters	% book chapters	GINI	Main specialisation fields – Activity Index **			
MIT Press Royal Soc Chemistry	England	210	0.71%	2,796	0.74%	0.641	PSY 3,22	GEO 2,61	SO MED 1,47	ECO 1,34
		207	0.70%	2,553	0.67%	0.669	CHE 23,98	BIO MO 3,63	ENG 1,67	BIO HU 1,54
Univ North Carolina Press	The USA	199	0.67%	2,104	0.56%	0.877	HUM 2,97	SOC 0,94	SO MED SO MED	xGEO 0,26
Chandos Publ E J Brill	England The Netherlands	193	0.65%	2,118	0.56%	0.884	SOC 2,70	ECO 0,96	HUM 0,62	ENG – 0,04
		188	0.63%	2,242	0.59%	0.867	HUM 2,75	BIO AP 1,23	SOC 0,78	GEO 0,32
Australian Natl Univ Birkhauser Verlag Ag	Australia Switzerland	183	0.62%	2,463	0.65%	0.792	GEO 2,29	SOC 1,95	HUM – 1,75	ECO 1,30
		143	0.48%	1,448	0.38%	0.807	MAT 13,14	MED CL 0,99	PHY AP 0,81	BIO HU 0,65
Artech House	The USA	124	0.42%	1,492	0.39%	0.847	PHY AP 10,23	ENG	7,50 GEO 1,18	MED CL 1,00
Maney Publishing Intellect Ltd Editions Rodopi B V Karger	England	117	0.40%	906	0.24%	0.932	HUM 3,47	SOC 0,03	–	–
		114	0.38%	1,401	0.37%	0.881	SOC 2,54	HUM 1,50	ECO 0,09	–
		104	0.35%	1,628	0.43%	0.926	HUM 3,27	SOC 0,21	–	–
		103	0.35%	1,567	0.41%	0.839	MED CL 14,11	BIO HU 3,98	PSY 2,98	SO MED 2,29
Ios Press	The Netherlands	102	0.34%	2,434	0.64%	0.737	ENG 5,34	SOC ME 2,53	ECO 1,49	SOC 1,30
World Scientific	Singapore	102	0.34%	1,304	0.34%	0.548	PHY AP 7,73	MAT 6,46	PHY AS – 2,47	ENG 2,42

Notes: * Publishers have at least 100 books indexed and they concentrate 90 of the BKCI total output regarding books as well as book chapters. **PHY AP: Applied physics and chemistry; BIO AP: Biological sciences – animals and plants; BIO HU: Biological sciences – humans; CHE: Chemistry; MED CL: Clinical medicine; ECO: Economics; ENG: Engineering; GEO: Geosciences; HUM: Humanities and arts; MAT: Mathematics; BIO MO: Molecular biology and biochemistry; SOC: Other social sciences; SOC ME: Other social sciences – medicine and health; PHY AS: Physics and astronomy; PSY: Psychology and psychiatry

total output has more than 16 citations per publication. Publications with a range between zero and five citations represent 95.1 per cent of the total database.

However, the most interesting thing is that one single publisher, Annual Reviews, represents 41 per cent of the total citation count. Actually records from this publisher have an average of 46.45 citations per document. This means that if we exclude this publisher, the average citation rate drops to 0.99 per document. When looking at the nature of the books published by Annual Reviews, we find most are actually periodicals as not only are they ordered by volumes but they also have an ISSN assigned, greatly distorting the database as the publisher with the second highest average of citations per document, Oxford University Press, has 12.35. In fact most of the records of this publisher are review articles. As pointed out by Archambault and Larivière (2009) journals with the highest Journal Impact Factor are usually review journals, simply because review articles are more frequently cited than regular articles, a factor that contributes to the anomalous performance of Annual Reviews. This distortion is so great that it forced us to exclude it in order to represent publishers according to their impact and concentration index as in Figure 4.

Figure 4 represents publishers according to their Field Normalised Citation Impact, Field Concentration Index (GINI) and number of books. We observe that most publishers are grouped on the right of the graph, showing that most of them display

Times cited	Records	% records
0	329,172	80.5
1-5	59,855	14.6
6-10	8,232	2.0
11-15	3,312	0.8
16-20	1,870	0.5
21-25	1,219	0.3
> 25	5,040	1.2
Total	408,700	98.8

Table IV.
Number of publications
per range of Book
Citation Index citations
for 2005-2012

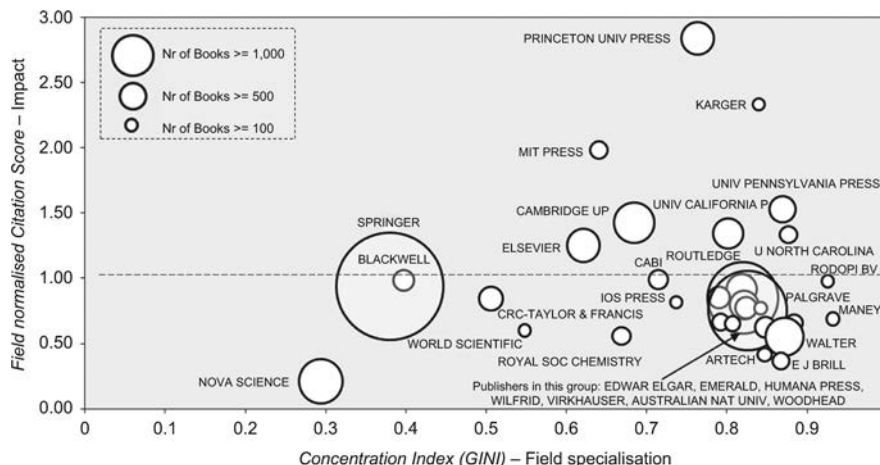


Figure 4.
Field Normalised Citation
Score and Concentration
Index for publishers
according to the Book
Citation Index 2005-2012

a disciplinary focus. Prestigious university publishers have higher impact (e.g. Princeton University Press, MIT Press, and Cambridge University Press), while multidisciplinary publishers tend to have lower impact (e.g. Nova Science, Blackwell, and Springer). We also observe that Springer, Palgrave, and Routledge, which can be considered three of the main scientific publishers in terms of number of books indexed in the BKCI, show lower values than expected, probably due to the high volume of books they generate, but also because their book chapters are little cited. Finally it is interesting to note that some clusters emerge from this representation. First is that of university publishers, highly specialised and with a high impact (e.g. Princeton, MIT Press, Cambridge University Press and University of California Press); second is a group of highly specialised publishers with low impact (Humana Press, Palgrave, Routledge and Emerald, for instance).

But we must call for caution when interpreting these results. As with every new, on-going project, we have detected errors on the citation count which may put into question some of the BKCI impact data. In Table V we present some examples of the two main mistakes we have found: a data processing problem and a conceptual problem. The first has to do with the data processing as some publications have considerably fewer citations in BKCI than can be seen when using the Cited Reference Search option. The case of Henk Moed's *Citation Analysis in Research Evaluation* is of special significance. While the BKCI retrieves seven citations to the book and 12 to its chapters, we have retrieved nearly 300 citations from the Cited Reference Search Option. In fact if we look further at possible causes we see that the BKCI identifies the book title as the source and therefore does not match most of the citations correctly. There are also some problems when processing different editions of a single book, for instance, Griewank and Walther's book for which only citations to the second edition are counted, from 2005. The second error has to do with a conceptual issue. As the BKCI indexes books and their chapters in separate records, it also considers their citations as separate. Therefore a book may have fewer citations than the sum of the citations received by its chapters, an especially worrying issue when analysing single-author books. Although this decision may be understandable for edited volumes it seems problematic, particularly since users would usually be searching for the overall impact of a monograph. Indeed, although it is a multi-authored book, we find that the *Handbook of Photosensory Receptors* by Briggs and Spudich has received six citations according to BKCI, but its chapters have received a total of 260 citations, giving a misleading idea of the book's impact.

Discussion

In this paper we analyse BKCI coverage in terms of publishers indexed by country and discipline as well as their impact. For this purpose we posed four research questions regarding the main publishers present in this database, their distribution by country and by discipline, the mainstream publishers' disciplinary focus, and their impact.

We observe that in the BKCI 33 scientific publishers take up 90 per cent of the whole share of the database. Indeed three publishers – Springer, Palgrave and Routledge – represent more than half of the books indexed in the BKCI. Interestingly Springer and Nova Science, the first and fourth most prolific publishers, are also those with the most multidisciplinary focus, while Palgrave and Routledge are highly specialised in the social sciences and humanities. This shows Thomson Reuters' efforts to overcome their

Problem A) BKCI does not retrieve all citations for books and book chapters. When using the “Cited Reference Search” option from Web of Science, more citations are identified rather than those shown by BKCI. This is a data processing problem

Examples	No. of citations in the BKCI (Books)	No. of citations in the BKCI (Book Chapters)	No. of citations using “Cited Reference Search”
Moed, H.F. (2005), <i>Citation Analysis in Research Evaluation</i> , Springer	7	12	298
Griewank, A. and Walther, A. (2008), <i>Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation</i> , 2nd ed., Siam	51	0	305
Dill, K.A. and Bromberg, S. (2011), <i>Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience</i> , Garland Science	4	0	186
Forrester, P.J. (2010), <i>Log-Gases and Random Matrices</i> , Princeton University Press	52	0	174

Problem B) Book citations do not include their chapters’ citations. This is a conceptual problem

Examples	No. of citations in the BKCI (Books)	No. of citations in the BKCI (Book Chapters)
Hancock, G.R. and Mueller, R.O. (2006), <i>Quantitative Methods in Education and the Behavioral Sciences-Issues Research and Teaching</i> , Information Age Publishing, Charlotte	25	267
Wood, J. and Dupont, B. (2006), <i>Democracy, Society and the Governance of Security</i> , Cambridge University Press	25	94
Briggs, W.R. and Spudich, J.L. (2005), <i>Handbook of Photosensory Receptors</i> , Blackwell Science, Oxford	6	260
Lau, W.K.M. and Waliser, D.E. (2005), <i>Intraseasonal Variability in the Atmosphere-Ocean Climate System</i> , Springer-Praxis Books In Environmental Sciences	69	278

Table V.
Two problems detected in the Book Citation Index citation count for 2005-2012

well-known weakness in failing to adequately cover these two research fields in the citation indexes. However, along with publication type, Thomson Reuters has been also criticised for its heavy bias towards English-speaking publications (Van Leeuwen *et al.*, 2001), an issue that especially affects the fields of the humanities and social sciences where the literature is more dispersed and there are strong regional influences (Archambault and Larivière, 2010; Hicks and Wang, 2011). This is has not been achieved by this product, as England and The USA alone represent 75.05 per cent of the total share by publication country and the English language represents 96 per cent of the database, showing a strong bias towards the Anglo-Saxon world and neglecting the research output of other countries and linguistic communities such as German, French or Spanish languages. In fact, although they acknowledge the importance of languages other than English, they emphasise their preference for this language

stating that “English language full text is highly desirable, but books with full text in a language other than English are also considered” (Testa, 2010). Moreover the great concentration of publishers in these areas (0.882 in other social sciences and 0.900 in humanities and arts) can be problematic, as many of these disciplines are characterised by their controversies between different trends of thought and paradigms.

An important issue that must be addressed after analysing the BKCI is whether the coverage of the BKCI represents the book market sector. Unlike journals where we can check Ulrich’s Periodicals Directory, one of the main problems we encounter when analysing this issue is the lack of exhaustive information resources which would allow us to know with precision whether the BKCI is an accurate representation. However, if we consult the statistics of specific countries we observe that the coverage by languages and countries is still very poor. For instance in Table VI we show data for the US and Spain retrieved from the New Book Titles and Editions, 2002-2011 and Estadística de la Producción Editorial 2011. As observed, the US book market sector is three times bigger than the Spanish one, proportions which are not reflected in the BKCI and which lead us to think that the database is strongly biased towards certain countries and languages.

Probably the most interesting issue to discuss from our study has to do with analysing publisher impact and how the BKCI processes citation counts. The low citation rates found are quite surprising, especially when the company emphasises its commitment to indexing the most relevant literature (Testa, 2010). However the uncitedness rate does not seem to differ much from that described by Schwartz (1997) who reports levels of uncitedness of 75 per cent and 92 per cent for social sciences and humanities respectively. Uncitedness issue is a long-discussed issue in bibliometrics, first mentioned by Price (1965) who stated that 10 per cent of scientific literature never gets cited. Because the BKCI only covers books dating from 2005 and since books and monographs are cited at a slower rate than papers, this high level should probably not be considered suspicious. A wider citation window would probably be needed to analyse this phenomenon. Moreover the citation rate of monographs differs from that of journals (Hammarfelt, 2011), calling for caution when interpreting citation counts for each publication type.

The problems indicated in the citation count should be taken into account and solved. Those derived from technical issues will probably be corrected soon but others of a more conceptual nature such as the distinction between book and chapter citations, or those regarding the processing of different editions of a single book should be analysed in depth. Indeed the inclusion of 40 journals published by Annual Reviews (Palo Alto, USA) distorts the nature of the database in terms of citations as its publications should not strictly be considered books especially if we take into account

Table VI.
Number of books
published in the USA and
Spain in 2011

Discipline	No. of books	
	USA	Spain
Social sciences	31,633	13,595
Applied sciences	24,692	10,194
Science	18,499	2,911
Total	74,824	26,700

that 35 of the journals published by Annual Reviews are indexed in the Journal Citation Reports and have Impact Factors.

Conclusion

The launch of the BKCI by Thomson Reuters is great news for information scientists and bibliometricians, as it provides a new resource for the analysis and development of tools to complete the research evaluation picture. Long neglected by citation indexes, books have become an important flaw in bibliometric studies because they are an important channel of communication in the fields of humanities and the social sciences. Their inclusion may improve the coverage of these fields in the Thomson Reuters' Web of Science. In this study we have analysed the product in order to assess its validity from the viewpoint of publisher presence and impact. Indeed the humanities and social sciences represent more than half of the total database. But the strong presence of English-speaking countries in the BKCI shows a heavy bias towards these countries, demonstrating the same limitations that have already been criticised in the other citation indexes (Van Leeuwen *et al.*, 2001).

Another problem with the coverage of the BKCI is that it includes serials that are actually part of the journal citation indexes (Science Citation Index, Social Science Citation Index and Arts and Humanities Citation Index). We have previously mentioned the case of Annual Reviews that clearly should not be part of the BKCI. In relation to this case if Thomson-Reuters considers Annual Reviews to be books, then other similar serials should have been included such as the Annual Review of Information Science and Technology that also has both an ISBN and ISSN. Thomson-Reuters appear to lack a clear policy to distinguish a journal from a book.

With regard to impact many limitations have been found in the citation count. Although some of them are data-processing errors and should be easily corrected, others are of a conceptual nature. As previously discussed the issue regarding different editions of a single book, or the differences in citation counts between books and their chapters may well be serious flaws in this database. However, as we have only included records from 2005 in our analysis, more time may well be needed before analysis of impact data and book citation curves is satisfactory. Further analyses are also needed in order to assess the content of the BKCI and the accuracy of its reflection of the publishing sector, similar to that performed by Chen (2012) regarding Google Books.

References

- Archambault, E. and Larivière, V. (2009), "History of the journal impact factor: contingencies and consequences", *Scientometrics*, Vol. 79 No. 3, pp. 639-653.
- Archambault, E. and Larivière, V. (2010), "The limits of bibliometrics for the analysis of the social sciences and humanities literature", *World Social Science Report 2010: Knowledge Divides*, UNESCO, Paris, pp. 251-254.
- Baneyx, A. (2008), "'Publish or perish' as citation metrics used to analyze scientific output in the humanities: international case studies in economics, geography, social sciences, philosophy, and history", *Archivum Immunologie et Therapie Experimentalis*, Vol. 56 No. 6, pp. 363-371.

- Bar-Ilan, J. (2010), "Citations to the 'Introduction to informetrics' indexed by WOS Scopus and Google Scholar", *Scientometrics*, Vol. 82 No. 3, pp. 495-506.
- Chen, X. (2012), "Google Books and WorldCat: a comparison of their content", *Online Information Review*, Vol. 36 No. 4, pp. 507-516.
- Edwards, L. (2012), "Editing academic books in the humanities and social sciences: maximizing impact for effort", *Journal of Scholarly Publishing*, Vol. 44 No. 1, pp. 61-74.
- Garfield, E. (1996), "Citation indexes for retrieval and research evaluation", paper presented at the Consensus Conference on the Theory and Practice of Research Assessment, 7 October, Capri, available at: www.garfield.library.upenn.edu/papers/ciretreseval-capri.html (accessed 27 July 2012).
- Garfield, E. (2007), "The evolution of the Science Citation Index", *Perspectives: International Microbiology*, Vol. 10 No. 1, pp. 65-69.
- Garfield, E. and Sher, I.H. (1963), "New factors in the evaluation of scientific literature through citation indexing", *American Documentation*, Vol. 14 No. 3, pp. 195-201.
- Giménez-Toledo, E. and Román-Román, A. (2009), "Assessment of humanities and social sciences monographs through their publishers: a review and a study towards a model of evaluation", *Research Evaluation*, Vol. 18 No. 3, pp. 201-213.
- Glänzel, W. and Schoepflin, U. (1999), "A bibliometric study of reference literature in the science and social sciences", *Information Processing & Management*, Vol. 35 No. 1, pp. 31-44.
- Hammarfelt, B. (2011), "Interdisciplinarity and the intellectual base of literature studies: citation analysis of highly cited monographs", *Scientometrics*, Vol. 86 No. 3, pp. 705-725.
- Hicks, D. (2004), "The four literatures of social science", in Moed, H.F., Glänzel, W. and Schmoch, U. (Eds), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*, Kluwer Academic Publishers, Dordrecht, pp. 473-496.
- Hicks, D. and Wang, J. (2011), "Coverage and overlap of the new Social Sciences and Humanities Journal Lists", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 2, pp. 284-294.
- Kousha, K. and Thelwall, M. (2009), "Google book search: citation analysis for social science and the humanities", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 8, pp. 1537-1549.
- Kousha, K., Thelwall, M. and Rezaie, S. (2011), "Assessing citation impact of books: the role of Google Books, Google Scholar, and Scopus", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 11, pp. 2147-2164.
- Kulkarni, A.V., Aziz, B., Shams, I. and Busse, J.W. (2009), "Comparisons of citations in Web of Science, Scopus and Google Scholar for articles published in general medical journals", *Journal of the American Medical Association*, Vol. 302 No. 10, pp. 1092-1096.
- Lewis, G. (2004), "Citations to papers from other documents", in Moed, H.F., Glänzel, W. and Schmoch, U. (Eds), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*, Kluwer Academic Publishers, Dordrecht, pp. 457-472.
- Leydesdorff, L. and Felt, U. (2012), "Edited volumes, monographs and book chapters in the Book Citation Index", *Journal of Scientometric Research*, Vol. 1 No. 1, pp. 28-34.
- Linmans, A.J.M. (2010), "Why with bibliometrics the humanities does not need to be the weakest link", *Scientometrics*, Vol. 83 No. 2, pp. 337-354.

-
- Lundberg, J. (2007), "Lifting the crown-citation z-score", *Journal of Informetrics*, Vol. 1 No. 2, pp. 145-154.
- Martin, B., Tang, P., Morgan, M., Glänzel, W., Hornbostel, S., Lauer, G., Lenclud, G., Lima, L., Oppenheim, C., vandenBesselaar, P. and Zic-Fuchs, M. (2010), "Towards a bibliometric database for the social sciences and humanities – a European scoping project", available at: www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/esf_report_final_100309.pdf (accessed 20 July 2012).
- Meho, L.I. and Yang, K. (2007), "Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 3, pp. 2105-2115.
- Moed, H.F. (2005), *Citation Analysis in Research Evaluation*, Springer, Dordrecht.
- Nederhof, A.J. (2006), "Bibliometric monitoring of research performance in the social sciences and the humanities. A review", *Scientometrics*, Vol. 66 No. 1, pp. 81-100.
- Price, D.S.D. (1965), "Networks of scientific papers", *Science*, Vol. 145 No. 3683, pp. 510-515.
- Schwartz, C.A. (1997), "The rise and fall of uncitedness", *College & Research Libraries*, Vol. 58 No. 1, pp. 19-29.
- Sivertsen, G. and Larsen, B. (2012), "Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential", *Scientometrics*, Vol. 91 No. 2, pp. 567-575.
- Testa, J. (2010), "The book selection process for the Book Citation Index in Web of Science", available at http://wokinfo.com/media/pdf/BKCI-SelectionEssay_web.pdf (accessed 25 July 2012).
- Torres-Salinas, D. and Moed, H.F. (2009), "Library Catalog Analysis as a tool in studies of social sciences and humanities: an exploratory study of published book titles in Economics", *Journal of Informetrics*, Vol. 3 No. 1, pp. 9-26.
- Torres-Salinas, D., Robinson-García, N., Jiménez-Contreras, E. and Delgado López-Cózar, E. (2012), "Towards a 'Book Publishers Citation Reports'. First approach using the 'Book Citation Index'", *Revista Española de Documentación Científica*, Vol. 35 No. 4, pp. 615-620.
- Van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S. and van Raan, A.F.J. (2001), "Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance", *Scientometrics*, Vol. 51 No. 1, pp. 335-346.
- White, H., Boell, S.K., Yu, H., Davis, M., Wilson, C.S. and Cole, F.T.H. (2009), "Libcitations: a measure for comparative assessment of book publications in the humanities and social sciences", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 6, pp. 1083-1096.
- Williams, P., Stevenson, I., Nicholas, D., Watkinson, A. and Rowlands, I. (2009), "The role and future of the monograph in arts and humanities research", *Aslib Proceedings*, Vol. 61 No. 1, pp. 67-82.
- Zuccala, A. and van Leeuwen, T. (2011), "Book reviews in humanities research evaluations", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 10, pp. 1979-1991.

About the authors

Daniel Torres-Salinas is CEO and co-founder of the University of Granada spin-off, EC3metrics, and works as a management technician at the University of Navarra. His current interests are research evaluation of universities, evaluation of scientific journals and the analysis of new evaluation tools and indicators. Dr Torres-Salinas is the corresponding author and can be contacted at: torressalinas@gmail.com

Nicolás Robinson-García is a PhD candidate at the University of Granada and member of the EC3metrics research group. His research is in the field of bibliometrics, where his interests include research evaluation of universities and the analysis of new evaluation tools and indicators.

Juan Miguel Campanario is a Professor in the Department of Physics at the University of Alcalá. His main research interests include bibliometrics, scientific communication, peer review and the resistance to revolutionary discoveries and science education.

Emilio Delgado López-Cózar is the Professor of Research Methods at the University of Granada and co-founder of the University of Granada spin-off, EC3metrics. His research interests include evaluation of scientific journals, research assessment at the institutional and individual levels and Google Scholar as a bibliometric tool.