ESTUDIOS / *RESEARCH STUDIES*

Assessing the teaching value of non-English academic books: The case of Spain

Amalia Mas-Bleda*, Mike Thelwall*

* University of Wolverhampton, School of Mathematics and Computing
e-mail: amalia.mas@wlv.ac.uk | ORCID iD: https://orcid.org/0000-0001-5927-424X
e-mail: M.Thelwall@wlv.ac.uk | ORCID iD: https://orcid.org/0000-0001-6065-205X

Abstract: This study examines the educational value of 15,117 Spanish-language books published by Spanish publishers in social sciences and humanities fields in the period 2002-2011, based on mentions of them extracted automatically from online course syllabi. A method was developed to collect syllabus mentions and filter out false matches. Manual checks of the 52,716 syllabus mentions found estimated an accuracy of 99.5% for filtering out false mentions and 74.7% for identifying correct mentions. A fifth of the sampled books (2,849; 19%) were mentioned at least once in online syllabi and almost all (95%) were from a third of the publishers included in the study. An in-depth analysis of the 23 books recommended most often in online syllabi showed that they are mostly single-authored humanities monographs that were originally written in Spanish. The syllabus mentions originated from 379 domains, but mostly from Spanish university websites. In conclusion, it is possible to make indicators from online syllabus mentions to assess the teaching value of Spanish-language books, although manual checks are needed if the values are to be used for assessing individual books.

Keywords: books; monographs; course syllabus; evaluation; teaching impact.

Estimación del valor educativo de los libros académicos que no están en inglés: el caso de España

Resumen: Este estudio examina el valor educativo de 15.117 libros escritos en español y publicados por editoriales españolas en disciplinas de ciencias sociales y humanidades en el período 2002-2011, en base a las menciones que reciben desde guías docentes en línea. Se implementó un método para identificar automáticamente las menciones y filtrar los resultados. El chequeo manual de las 52.716 menciones encontradas estimó una precision del 99,5% para filtrar las menciones falsas y del 74,7% para identificar las menciones correctas. Una quinta parte de los libros (2.849; 19%) se mencionaron al menos una vez en guías académicas en línea, y casi todos ellos (95%) han sido publicados por un tercio de las editoriales incluídas en el estudio. Un análisis detallado de los 23 libros más recomendados en guías docentes mostró que la mayoría son monografías de humanidades con un solo autor, escritos originalmente en español. Las menciones procedieron de 379 dominios web, mayoritariamente de sitios web de universidades españolas. En conclusión, es posible crear indicadores a partir de las menciones en guías docentes para evaluar el valor educativo de los libros en español, aunque se requieren chequeos manuales si los valores se usan para evaluar libros individuales.

Palabras clave: libros; monografías; guía docente; guía académica; evaluación; impacto educativo.

1. INTRODUCTION

Scholarly books are an important communication channel in the humanities and many social sciences fields (e.g. Engels et al., 2012; Huang and Chang, 2008; Nederhof, 2006), where they are also critical for tenure and promotion (Cordón-García et al., 2017; Cronin and La Barre, 2004). Attempts have been made to assess the quality and impact of books using a diverse range of indicators, such as citations, library holding statistics, publisher prestige or book reviews; however, the teaching value of scholarly books has rarely been examined, and never systematically for Spanish.

Scholarly course syllabi, as the set of documents with information related to the teaching planning of university courses, seem to be a suitable source of evidence for the educational value of scholarly publications, since they tend to include a list of bibliographic references of relevant materials for the course recommended by educators for their students to read.

Whilst previous research has examined the teaching utility of journal articles (Kousha and Thelwall, 2008) and monographs (Kousha and Thelwall, 2016a) by examining mentions of them in online academic syllabi, they focus on English-language publications and no study has systematically evaluated the outputs of selected academic publishers. Using books from academic Spanish publishers that produce scholarly books in social sciences and humanities fields, the current study introduces a method to search automatically for mentions in Spanish-language syllabi.

Scholarly books in Spain

The Spanish ISBN Agency has been responsible for assigning ISBNs to books since 1972. This agency also manages a database of Spanish book publishers. In 2016, the Spanish ISBN Agency assigned 86,000 ISBNs, mostly to books published for the first time (97.7%) by private publishers (90%) in Spanish official languages (91.4%; 73% in Spanish) and in paper format (70%) (MECD, 2017). These books were mostly from the social sciences and humanities (31%), were literary works (21%), such as literature, novels, and poetry, and scientific-technical books (14%), typically about hard sciences, engineering, medicine, communications and agriculture (MECD, 2017).

Spanish university presses published in 2015 mostly Spanish-language publications (81%) in humanities (41%) and social sciences (34%) (UNE, 2016). The economic situation of Spanish university presses has deteriorated in the last decade, with their budgets decreasing by 35% from 2005 to 2015 (UNE, 2016).

Scholarly books are important for Spanish scholars in the humanities and some social sciences fields (Osca-Lluch et al., 2013; Sorli Rojo et al., 2011). Law, history, education, economics and linguistics, literature and philology are the fields with the most book publishing, according to the number of ISBNs (Giménez-Toledo, 2017).

Publisher prestige is one of the aspects considered in the book assessment by ANECA, the Spanish National Agency for Quality Assessment and Accreditation. It uses the Scholarly Publishers Indicators portal (http://ilia.cchs.csic.es/SPI/rankings.html) for evidence of the prestige of a book publisher (ANECA, 2017).

Book impact assessment

The impact of scholarly publications has been traditionally assessed using the two most widely used citation-based databases, the Web of Science (WoS) and Scopus, which cover mainly mainstream English-language journals (Archambault et al., 2006). This coverage bias marginalises social sciences and humanities research (Caerols Mateo et al., 2017), which usually has a stronger national and regional interest (Hicks, 1999), tends to be published in the language of the nation or culture that is studied rather than in English (Engels et al., 2012; López-Navarro et al., 2015; Moed et al., 2002), and in national journals rather than in international journals (Larivière and Macaluso, 2011). For some social scientists, and many humanities scholars, monographs and book chapters play an important role in scholarly communication (Huang and Chang, 2008; Leydesdorff and Felt, 2012; Nederhof, 2006; Romanos de Tiratel, 2000; Sivertsen, 2016; Sivertsen and Larsen, 2012; Thompson, 2002).

Recognition of the importance of books in some fields has led to books being indexed in WoS and Scopus and the creation of citation data sources for monographs. Currently, the Thomson Reuters Book Citation Index (BKCI) indexes over 60,000 books published from 2005 onwards (Book Citation Index, 2017) and Elsevier's Scopus Books Expansion Project records more than 150,000 books (Elsevier, 2018). However, their coverage is biased towards English and publishers from English-speaking countries, especially the United States and the United Kingdom (Gorraiz et al., 2013; Torres-Salinas et al., 2014b). Some countries have also developed initiatives for the assessment of scholarly books (Giménez-Toledo et al., 2016).

Google Books (GB) has much greater coverage than WoS or Scopus. Launched in 2004, it includes the text of over thirteen million books digitalised by Google from publisher and library partners (Barron, 2011). It does not have a citation index, but a method has been developed to extract citations from it (Kousha and Thelwall, 2015).

Statistics derived from libraries have been also suggested for assessing the cultural impact of books, such as a count of the number of libraries holding a book (White et al., 2009) or the presence of a book in online catalogues (Torres-Salinas and Moed, 2009).

Publisher prestige is frequently used as an indicator of the value of a book since better books are likely to be edited by more prestigious publishers. There have been several attempts to rank scholarly book publishers according to their perceived editorial prestige (Giménez-Toledo et al., 2013), publication and reading preferences (Garand and Giles, 2011) and citations (Torres-Salinas et al., 2012; Zuccala et al., 2015a).

Books reviews are important to scholars, especially in the humanities (Hartley, 2006; Spink et al., 1998) and they might be used in book assessments. History and literature seem to be the humanities fields that publish the most book reviews (Gorraiz et al., 2014; Sorli Rojo et al., 2011; Zuccala and van Leeuwen, 2011). Besides book reviews written by scholars, sources that include book reviews from any readers, not necessarily scholars, include Goodreads (Kousha et al., 2017; Zuccala et al., 2015b) and Amazon (Dimitrov et al., 2015; Kousha and Thelwall, 2016b).

Other studies have compared potential sources of impact evidence for journal articles (Mas-Bleda and Thelwall, 2016) and books (Halevi et al., 2016). For instance, Halevi et al. (2016) analysed a set of alternative indicators for assessing the impact of books based on a sample of 70,000 ebooks from Ebrary. Focusing on the top-ten books for each indicator, this study showed that each source highlights different books, concluding that different types of books should be assessed using different indicators.

Impact assessment with syllabi

A scholarly syllabus is a document with pertinent information about a course, created by the professor who gives that course. It is a reference document in the teaching process for both professors and students. A course syllabus has many different functions, components and characteristics (Slattery and Carlson, 2005). Syllabi also reveal aspects of teaching strategies, such as assessment strategies (San Martín Gutiérrez et al., 2016), interactivity (Cummings et al., 2002) and the university services exploited (Williams et al., 2004).

An important component of many syllabi is a list of bibliographic references for books and other resources that students should read. Scholarly publications recommended in a course syllabus, as either required or supplementary reading, represent an endorsement by the instructor of their teaching utility.

The most frequently recommended resources in Spanish scholarly syllabi are in Spanish and are books. A study analysing syllabi directly related to Media Literacy in Communication and Education degrees in the Spanish context found that 90% of the references recommended were in Spanish (Marta-Lazo et al., 2014). Another study analysing syllabi related to pedagogy and didactics in sciences of physical activity and sport degrees found that 89.5% of the recommended publications were in Spanish and 73.5% were monographs (Gutiérrez García et al., 2016). A dissertation analysing syllabi from a Spanish university revealed that the monograph is the source type most recommended in history (69%) and sociology (51%) (Prieto-Paíno, 2013).

A few large-scale studies have examined whether course syllabi might be used as a source to identify the educational value of scholarly publications. A study analysed syllabus mentions to over 70,000 journal articles published in 2003, showing that they might be useful for some social sciences research (Kousha and Thelwall, 2008). Another study analysed syllabus mentions to 14,000 English-language monographs from 2005 to 2010 indexed in Scopus, showing that 56% of arts and humanities monographs had at least one syllabus mention in 2014 and about a third of them had at least one syllabus mention but no citations (Kousha and Thelwall, 2016a).

The proportion of books mentioned shows that course syllabi are a useful source to reflect the educational utility of English monographs from the prestigious publishers selected by Scopus, especially in the humanities. However, there is a need to investigate whether online syllabus mentions are common enough to be useful for impact assessment purposes for books that are not in English and for books that are not from the most prestigious international English-language publishers. This is important because many countries have academic presses, often based in universities, that publish at least partly in the local language. Academics that write such books may need evidence of their impact and may not find this in citation counts from WoS and Scopus, especially if the books are designed for a national audience.

Research questions

This study addresses the need to assess the educational impact of non-English books for the case of Spain. It assesses whether it is possible to demonstrate that a significant proportion of Spanish-language books from publishers that are prestigious or productive in Spain have impact on teaching, based on mentions that they receive from online syllabi. Spain was selected as a large developed nation with a national publishing tradition and a relatively strong research base. It is logical to assess such a country before examining smaller countries or less developed nations, for which the evidence would presumably be weaker. The following research questions are addressed.

1. Can mentions of Spanish-language books in online scholarly syllabi be automatically identified accurately?

2. Do enough Spanish-language books from successful Spanish publishers have at least one online syllabus mention to make online syllabus mentions a useful educational impact indicator for them?

3. What are the main characteristics of course syllabi that recommend Spanish books?

2. METHOD

Data collection

The National Library of Spain receives and keeps copies of all books published in Spain in addition to a valuable collection of incunabula, manuscripts, drawings, photographs, scores, etc. Its general catalog gives bibliographic references for all the documents kept in the library. Datos.bne.es (http://datos.bne.es), the bibliographic data portal of the National Library of Spain, was used as the data source. It makes available the bibliographic and authority catalogs of the library in RDF format in accordance with the principles of Linked Data to make them more visible on the web. Specifically, it offers around seven million records (Vila-Suero et al., 2012) through three files: authority records, bibliographic records and subject headings. For this research, its file of 1.5 million bibliographic records of all its holdings was downloaded on 17 November 2016. Each record included the book title, author names, publisher, publication year, publication place, document type, language and ISBN.

The study focuses on Spanish-language texts (Castillian), so records in other official languages in Spain (Catalan, Basque and Galician) were excluded. Records were restricted to those with 'texto impreso'

(printed text) as the document type. Most seemed to be monographs, but there was no filter to distinguish between monographs, edited books and book serials. Other types of document, such as electronic resources, maps, periodical publications, printed music or photographs were excluded. Records that did not include author, publisher or publication year were excluded as well as records that had strange characters in the author field (e.g. A.+B.) or year field (e.g. 3 or 5 digits). This filtering left 726,529 Spanish-language books. This should include nearly all books published in Spain.

The data source did not have a classification scheme for academic subject areas, so it was not possible to select books by discipline. It also did not differentiate between academic and non-academic publishers and it contained records for books published by the end of 2011. Therefore, the Scholarly Publishers Indicators (SPI) website (http://ilia.cchs.csic.es/SPI) was used to identify Spanish academic publishers.

The SPI website ranks academic publishers relevant to the humanities and social sciences by editorial prestige (both Spanish and some non-Spanish publishers) and thematic specialization (Spanish publishers only). This publisher prestige ranking is based on the opinions of Spanish scholars (Giménez-Toledo et al., 2013). The thematic specialization ranking classifies the most productive publishers in each discipline (Anthropology, Archeology & Prehistory, Fine arts, Library & Information Sciences, Political Sciences, Communication, Law, Economy, Education, Arab and Hebrew Studies, Philosophy, Geography, History, Linguistics, Literature & Philology, Psychology, and Sociology). For each discipline, it ranks publishers according to the number of books published in that discipline, based on an analysis of data from the Dilve database (which contains Spanish book marketing information).

The steps for the selection of the sample were the following:

a) Selection of both the most prestigious and the most productive Spanish academic publishers according to the SPI portal. Two features were noted: 1) the ten most prestigious publishers were not the same in every discipline and 2) in some disciplines, some productive publishers were not prestigious. For instance, the most productive publisher in Fine Arts (Lungwerg), the second most productive publisher in History (Nowtilus), the two most productive publishers in Geography (Lungwerg and Ediciones B) and Education (Adams and Editorial MAD) and the third most productive publisher in Arab and Hebrew Studies (Obelisco)

and Sociology (Vicens Vives) were not perceived as prestigious in those disciplines. Because this study aimed at covering publishers from every discipline, we selected both the most prestigious and the most productive ones. This approach lets us identify differences among these two groups regarding teaching utility. For instance: which type of publisher publishes the books with the highest number of syllabus mentions?

- Prestigious publishers were selected based on the editorial prestige ranking by discipline. Because the ten most prestigious publishers were not the same in every discipline, publishers that appeared among the top ten in at least 6 out of 16 disciplines were selected. This selects the publishers perceived as the most prestigious in most of social sciences and humanities fields.

- Productive publishers were selected based on the thematic specialization ranking. The three most productive publishers in each discipline were chosen.

b) Identification of books published by those publishers included in the 'Datos.bne.es' database.

c) Restriction of publishers to those that had at least 200 Spanish-language books in the period 2002-2011 (to focus on major publishers), after removing duplicates and books with single and two-word titles.

This produced 15,117 books from 27 Spanish publishers (Table I). All publishers meeting the criteria are profit-oriented, except for one university press (Universitat Oberta de Catalunya - UOC) and one publisher belonging to a research institution (CSIC). This is unsurprising since 90% of publishers in Spain in 2016 were private (MECD, 2017).

Automatic syllabus mentions searching

Mentions of the 15,117 books from 27 Spanish scholarly book publishers were searched for in online academic syllabi or course reading lists. Queries were submitted to Bing Search API, via the free Webometric Analyst program (http://lexiurl.wlv.ac.uk) on Friday 13 December 2016, as described below. The API is the only approved source of Bing data for automatic queries (Thelwall and Sud, 2012), but Bing seems to have a smaller index than Google (Van den Bosch et al., 2016), which has a much more limited API (100 queries per day, with a maximum of 10 results per query).

Table I. Prestigious or productive Spanish academic book publishers with at least 200 Spanish-language books in the Datos.bne.es database for the period 2002-2011 (n= 15,117)

Publisher	Books in Datos.bne.es	Prestigious	Productive	Publisher	Books in Datos.bne.es	Prestigious	Productive
Ediciones B	1521	No	Yes	Desclée de Brouwer	396	No	Yes
Alianza	1299	Yes	Yes	Cátedra	393	Yes	Yes
Tirant Lo Blanch	1206	Yes	No	UOC	366	No	Yes
Espasa Calpe	980	No	Yes	Tecnos	364	Yes	No
Paidós	904	Yes	No	Trotta	347	No	Yes
Aranzadi	781	No	Yes	Herder	341	No	Yes
Dykinson	708	No	Yes	Trea	304	No	Yes
Obelisco	626	No	Yes	CSIC	269	Yes	Yes
Síntesis	575	Yes	No	Nowtilus	250	No	Yes
Akal	573	Yes	Yes	Vicens Vives	235	No	Yes
Civitas	561	No	Yes	ESIC	234	No	Yes
Marcial Pons	504	Yes	No	Díaz de Santos	220	No	Yes
Ariel	485	Yes	No	Lunwerg	213	No	Yes
Crítica	462	Yes	No				

This method adapts a method previously used for English books by translating the syllabus terms to Spanish, adding new terms related to the Spanish educational system and adding new rules to exclude non-academic sites. This uses the assumption that courses recommending Spanish language books are likely to have syllabi in Spanish. This is a simplification since, for example, Spanish literature courses in English speaking nations may recommend Spanish-language analytical monographs.

Webometric Analyst was used to generate automatic searches, using the first author surname, the first (up to) 7 terms of the monograph title as a phrase search and the terms "guía docente" and "guía académica" to limit the results to Spanish-language course syllabi, since these were the most widely used terms, according to a previous manual check made by authors of this study through university websites. The vertical bar '|' (an OR operator internal to Webometric Analyst) was used to run both queries separately in Bing and combine the results after removing duplicates. Books with single or two-word titles were removed because these are likely to generate many false matches. The publisher name was added to the queries for books with three words in their titles to reduce the number of false matches. The following are examples of the queries used.

> Antón "Manual de técnica policial" "guía docente"|Antón "Manual de técnica policial" "guía académica"

> Bellori "Vidas de pintores" "Akal" "guía docente"|Bellori "Vidas de pintores" "Akal" "guía académica"

Standardisation of publisher names

Some publishers appeared in the 'Datos.bne.es' source with variants of their names (e.g. 'Alianza' and 'Alianza Editorial'; 'Civitas' and 'Thomson Civitas'). They all were included in the study, but publishers' names were standardised and automatic queries contained the terms most likely to give the maximum recall. The publisher 'Consejo Superior de Investigaciones Científicas' appeared with many different names. As most of them contained the word 'CSIC' (its acronym), this term was selected.

Identification of mentions from online scholarly syllabi

The automatic searches retrieved 52,716 syllabus mention matches for the 15,117 sampled books, but not all were correct. They were filtered via Webometric Analyst software to remove false matches. However, its default version did not work correctly when filtering false matches, since its set of rules (see Kousha and Thelwall, 2016a) was designed to work on English-language syllabi and the sample in this study was based on a set of books published in Spanish and likely to be mainly cited in Spanish-language academic syllabi. Consequently, a new version of the Webometric Analyst software was created, defining a new set of rules to filter out correct and false syllabus mentions on Spanish-language text and its accuracy was tested with manual checks.

Rules to filter out false matches from Spanish-language pages

Manual checks of over fifty thousand syllabus mentions were carried out to help define a set of rules to filter out false matches in Spanish-language texts. Different sets of rules were tested and the most accurate set was selected. Mentions in course syllabi hosted in a webpage of an academic teaching centre (university, secondary school, college of music, art school, higher education centre assigned to a university), excluding primary and secondary education academies were classed as correct matches.

A set of rules was defined for the titles, descriptions and URLs of the syllabus search results from scholarly websites, partly motivated by the study carried out by Kousha and Thelwall (2016a).

- Search results title matches. The title field in the Bing API results was used to limit the results to pages containing terms with a high likelihood of showing that the webpage was a syllabus. Wildcard matching (*) was used to cover multiple term endings. Some of the terms included were: _asignatura*, guía docente, guía académica, guía didáctica, bibliografía, lecturas recomendadas, material de consulta_.

- Search results description matches. Some syllabus webpages did not include in their titles syllabus-related terms, such as _asignatura_ or _guia docente_. Instead, they used a course name, the repository where the syllabus was or the faculty name, which made it difficult to automatically detect if the webpage was a syllabus from its title. Therefore, additional rules were added to search syllabus-related terms in the description field. Most of the title search terms were therefore also searched for in the description field, as at least one syllabus-related term was usually in one of the two fields. Rules for the description field can be case sensitive so terms were included in

different case formats (e.g. 'Guía docente', 'Guía Docente', 'GUÍA DOCENTE').

- Search results URL matches. A set of rules was added to locate syllabus terms in URLs to catch pages that did not include such terms in the title or description fields. A list of terms that should be found in URLs was therefore specified, mainly taken from the title terms. Some relevant instances of syllabus terms in URLs include the following.

  https://guiae.uclm.es/vistaGuia/345/038336/2014-15

  http://titulaciones.unizar.es/asignaturas/26111/actividades15.html

Mentions that came from non-scholarly sites were judged to be false matches – probably mentioning books in a context other than higher education teaching. These included document sharing websites (e.g. *docplayer.com*), book download websites (e.g. *descargalibros.org*), platforms hosting blogs (e.g. *blogs.com*), wiki hosts (e.g. *wikipedia.org*), publishers, online bookshops, scientific databases, scientific journals, academic social networks and a wide range of other sites.

False matches also came from scholarly teaching centre webpages, not from course syllabi but from other types of scholarly documents, such as reports, journal articles, conference articles or PhD dissertations. Most were hosted in institutional repositories, library catalogues, university journals and scientific databases. There was not an easy way to define the rules to exclude this type of false match, since universities do not make available their course syllabi in a standard way. For instance, some universities had a database for their course syllabi, but others used an institutional repository section or a syllabus catalogue. Interestingly, some universities offered a database focused on the recommended readings of their academic syllabi rather than on the syllabi themselves (e.g. https://biblioteca.unirioja.es/biba). Repositories and databases were excluded when their contents were solely scientific. In these cases, all results from their websites were removed.

To remove false matches from other digital libraries, book databases or book lists, results with a range of related terms in their title (e.g. *Catalogo de libros*) and URL (e.g. **catalogo-libros**, **catalogo-bibl**) were ignored. False matches from non-syllabus publications were removed by ignoring any results with a range of related terms in their title (e.g. *Libro de actas**, *Memoria de investigación**) and URL (e.g. **viewArticle**, **congreso**). Blogs, forums and CVs were also removed from results.

Academic syllabi were assumed to be in Word (*.doc* and *.docx*), PDF (*.pdf*) or HTML (*.htm* and *.html*) files, so results with other file extensions (e.g. *.ppt, .txt*., *.xls*, *.xml*) were excluded. It was found, for instance, that URLs ending with '.pdf.txt' were mainly research reports, PhD dissertations, conference proceedings and university resolutions and Excel files (*.xls* and *.xlsx*) were mainly library catalogues listing books and other publications and bulletins of new acquisitions. Not all the files with *.doc*, *.pdf* and *.html* extensions were course syllabi, though (see Table II).

Although most of the terms used in the automatic rules were in Spanish, a few English terms were included, when necessary, such as 'journal_content' and 'viewArticle' to exclude mentions from journals or 'ebook' to exclude electronic books.

The new version of the software to work in Spanish text has been integrated into the Webometric Analyst software (http://lexiurl.wlv.ac.uk/), which is publicly available. To access to this specific option as well as the full set of defined rules, specify Spanish as the syllabus filtering language.

Manual checks of syllabus mentions filtered by the new rules

Out of 52,716 syllabus mentions reported by the Bing API, 26,195 (49.7%) were automatically classified as initially correct matches and 26,521 (50.3%) as initially false matches by Webometric Analyst, based on the rules defined above. The first author of this study manually checked of all the syllabus mentions to verify whether the sampled books were mentioned in the context of academic course reading lists (during the period June-November 2017). The content of the citing course syllabi was accessed and checked to see whether the books cited in them matched the sampled books.

A mention was judged as valid when the sampled book matched the book cited in the course syllabus.

- If the book cited in a syllabus did not include its publisher, but the remainder information was the same as the sampled book, it was assumed to be edited by the same publisher and classed as a valid mention.

- If the book cited in a syllabus did not include the publication year, it was assumed that any of its editions was good for students and classed as a valid mention.

- If the book cited in a syllabus had a publication year different from the sampled book, the mention was also judged valid. It

Table II. File extensions of online documents mentioning the sampled books

File extension	Type of document	Examples
.pdf .txt	Research report	http://acceda.ulpgc.es/bitstream/10553/1127/2/5496.pdf.txt
	PhD dissertation	http://digibug.ugr.es/bitstream/10481/3488/2/1852574x.pdf.txt
	Conference proceeding	http://repositori.uji.es/xmlui/bitstream/handle/10234/116608/9788480219426.pdf.txt
	University resolution	http://rabida.uhu.es/dspace/bitstream/handle/10272/1421/BOUH%20 2008,%20n.46-30.pdf.txt
.txt	Conference proceeding	http://sid.usal.es/idocs/F8/FDO21556/amadis.txt
	Course syllabus	http://www.uprrp.edu/acoi/uploads/376abcf8ab041a9e73e1e9a30daa22e1.txt
.xls .xlsx	Library catalogue	http://iespgaza.educa.aragon.es/biblioteca/qrEjemplares.xls
	News bulletins of acquisitions	http://www.siglodelhombre.com/boletines/excel/siglodelhombre-boletin-novedades-30-agosto-2013.xls
.ppt	Conference presentation	http://portalvicfie.webs.uvigo.es/arquivos/021_lamela_pautas.ppt
	Course presentation	http://www.unizar.es/abarrio/Principios%20Derecho.ppt
.pdf .doc .docx	Course syllabus	http://web.unican.es/estudios/Documents/Guias/2016/es/M1211.pdf
	Report	http://indoc.uca.es/memorias/sol-201400047965-tra.pdf
	Curriculum vitae	http://www.um.es/agcorrea/curriculum.doc
	Library catalogue	http://www.farem.unan.edu.ni/docs/Bibliografia.docx
	University resolution/ announcement	http://eii.unex.es/general/gobierno/colegiados/actas/20090924%20JE%20 extraordinaria.pdf
	Other academic publications	http://ddd.uab.cat/pub/tesis/2016/hdl_10803_382466/kvpp1de2.pdf

was considered a good mention when any of the editions of the book (for instance, the last edition) was recommended in the course syllabus.

A mention was judged as false when the sampled book did not match the book cited in the course syllabus, or the page was not a syllabus, mainly due to the following reasons:

- The mention came from a non-scholarly webpage.

- The mention came from a scholarly webpage, but not from a course syllabus.

- A different version of the book was cited. Both the sampled and the cited books had the same author/s and title, but they were edited by different publishers.

- The cited book had the same author as the sampled book, but their titles were slightly different.

- The book cited in the syllabus had a similar or identical title to the sampled book, but a different author/s, so that they were clearly different monographs.

  Sampled book: Bourdieu, P. (2008). *Cuestiones de sociología*. Madrid: Akal.

  Cited book: Alberoni, G. *Cuestiones de sociología*. Barcelona: Herder.

- The sampled book was not cited in the references, but its title was like the department name, academic field, course title or it was within the course description. This was common for books with general titles, such as "Historia del arte" (History of art).

It was impossible to verify if a mention was valid or false when there was a problem with the URL, such as: the link did not work, the link redirected to another webpage (usually the department homepage) or the syllabus file was password-protected.

3. RESULTS

The Bing API reported 52,716 syllabus mentions for the 15,117 sampled books. Based on the final set of filtering algorithms, the Webometric Analyst software classified 26,521 (50.3%) as false mentions and 26,195 (49.7%) as correct mentions (see Table III).

Accuracy of the Spanish rules for filtering out false syllabus mentions

The manual checks of mentions classified as false by Webometric Analyst confirmed that 26,400 (99.5%) were false mentions but 121 (0.5%) were correct mentions that had been incorrectly classified as false (Table III). A common reason for missing relevant mentions was that search result titles, descriptions or URLs did not include the syllabus-related terms specified in the rules.

Kousha and Thelwall (2016a) reported about 8% missing relevant English syllabus mentions. The reason for this difference probably is that in the present study, the rules to filter out false mentions were defined after several manual checks of the Bing results for this sample, so that the rules might remove a higher proportion of correct syllabus mentions from different samples if the new syllabus mentions match new types of pages.

Most of the false mentions came from docplayer site (43.5%), scribd.com (14%), myslide.es (3.5%), academia.edu (3.5%), issuu.com (3%), documents.mx (2.2%), calameo.com (2%), blogspot (2%) and wordpress.com (2%). Some of these non-scholarly sites had several domains, such as docplayer (e.g., docplayer.com, docplayer.es, docplayer.it), scribd.com, blogspot and docslide.

Accuracy of the Spanish rules for retaining correct syllabus mentions

The Spanish rules in Webometric Analyst classified 26,195 of the Bing API matches as initially correct syllabus mentions. A manual check verified that two thirds (65.3%) were correct mentions (i.e., syllabi citing the sampled monographs), whereas 22.1% were false mentions and 12.6% could not be checked (see Table IV). Ignoring the matches that could not be checked, the estimated accuracy of the method was 74.7%. Thus, a quarter of the matches judged by the rules to be correct were not.

Not all the manually verified correct mentions had identical citations. About 39% referred to another edition (published in a different year) and about 8% referred to a book without including its publication year.

The predominant reasons why syllabus mentions classified as correct by the Spanish rules were actually false mentions were the following: a book with a similar title was mentioned; the book title was like a department name, academic field, or course name; and another publisher had translated the same book.

It was sometimes not possible to check whether a syllabus mention was correct. The main reasons were that the web address (URL) did not work (Table IV). Some .pdf files seemed to be damaged as they could not be opened. Files with .zip or unusual format (application/octet-stream) were skipped as well as URLs identified as a virus threat. A few mentions came from blank syllabi, suggesting that they had been modified in the period between the submission of the automatic queries and the manual checking of the syllabus mentions.

Books with teaching value

Out of the 15,117 sampled books, 2,849 (19%) had at least one mention in an online course syllabus. Most had less than ten mentions (Table V). About 95% of the books with at least one syllabus mention were from a third of the publishers: Alianza, Paidós, Síntesis, Tirant Lo Blanch, Ariel, Akal, Crítica, Tecnos, Cátedra and Marcial Pons.

Table III. Results of manual checks of syllabus mentions found for the 15,117 sampled books

Webometric Analyst filtering result	Web pages	Manual check results
False mention	26,521 (50.3%)	26,400 (99.5%) false mentions
		121 (0.5%) correct mentions
Correct mention	26,195 (49.7%)	17,104 (65.3%) correct mentions
		5,782 (22.1%) false mentions
		3,309 (12.6%) unverifiable
Total	52,716 (100%)	

Table IV. Characteristics found from manual checks of the 26,195 initially correct mentions

Type of mention	Syllabus mentions	Characteristics	
Verified correct	17,104 (65.3%)	6,680 (39%)	referred to another publication year
		1,337 (8%)	no publication year
Verified false	5,782 (22.1%)	2,142 (37%)	cited a publication with a similar title
		1,648 (29%)	mentioned a similar text within course description
		1,141 (20%)	referred to a version from another publisher
		272 (5%)	linked to an empty syllabus or an updated version that did not cite the sampled book
		450 (8%)	from an academic site, but not from a syllabus
		129 (2%)	not from an academic site
Unverified	3,309 (12.6%)	2,781 (84%)	URL did not work
		224 (7%)	webpages were performing maintenance tasks
		173 (5%)	URL redirected to another webpage
		79 (2%)	damaged .pdf file, .zip file or unusual file (application/octet-stream)
		32 (1%)	blank syllabus
		12 (0.4%)	unavailable (non-active, password-protected, restricted access)
		8 (0.2%)	URL identified as a threat by antivirus program

Table V. Number of books with at least one syllabus mention (n=2,849)

Nº of syllabus mentions	Nº of books (%)
40-49 mentions	3 (0.1%)
30-39 mentions	9 (0.3%)
20-29 mentions	52 (1.8%)
10-19 mentions	264 (9.3%)
1-9 mentions	2521 (88.5%)

Table VI displays the 23 books mentioned most often in online course syllabi. A database of books published in Spain (https://www.mecd.gob.es/cultura-mecd/areas-cultura/libro/bases-de-datos-del-isbn/base-de-datos-de-libros.html) was used to identify the topic of the books and check whether they were translations. According to this database, all the books are monographs, but it does not clearly distinguish monographs from edited volumes or textbooks, so additional web searches were made to know if they targeted students. The abstracts and/or reviews of the books from Amazon and publisher websites were also checked. The principal characteristics of the 23 most mentioned books in online syllabi are:

The 23 most mentioned books in online syllabi are single-authored (83%), originally written in Spanish (76%), from humanities fields, particularly law (39%), and are published by publishers perceived as prestigious (96%), especially by Alianza (26%), Tecnos (22%) and Ariel (17%). Six (26%) books might be considered textbooks, as they seem to target students.

Online academic syllabi citing the sampled books

This section analyses the manually checked 17,104 correct syllabus mentions to gain insights into their origins. They came from 379 domains, 160 sites and 12 top and second level domains. A domain is the part of an URL after "//" (e.g. *guiae.uclm.es* in the URL 'https://guiae.uclm.es/vistaGuia/314/39325/2012-13'). A site is the distinguishing end of the domain names of a URL (e.g. *uclm.es*). The top-level domain is the text after the last dot (e.g. *.es*, *.com*). A second level domain also sometimes exists (e.g. *ac.uk*).

Table VI. The 23 sampled books mentioned most often in online Spanish course syllabi

Authors	Title	Year	Publisher	Syllabus mentions	Type	Broad topic
Broseta Pont, M., Martínez Sanz, F.	Manual de derecho mercantil	2010	Tecnos	48	Monograph (2 vol.)	Law
Arnheim, R. (tr. M. L. Balseiro)	Arte y percepción visual	2002	Alianza	47	Monograph	Art & Psychology
Mangas Martín, A., Liñán Nogueras, D.	Instituciones y derecho de la Unión Europea	2010	Tecnos	42	Monograph	Law
Harris, M. (tr. F. Muñoz de Bustillo)	Introducción a la antropología general	2004	Alianza	38	Textbook	Antropology
Lasarte Álvarez, C.	Principios de derecho civil	2011	Marcial Pons	36	Monograph (several vol.)	Law
Vicent Chuliá, F.	Introducción al derecho mercantil	2010	Tirant lo Blanch	36	Monograph (2 vol.)	Law
Escandell Vidal, M. V.	Introducción a la pragmática	2006	Ariel	35	Monograph	Linguistics
Calsamiglia, H., Tusón Valls, A.	Las cosas del decir	2007	Ariel	34	Monograph	Linguistics
De Micheli, M. (tr. A. Sánchez-Gijón)	Las vanguardias artísticas del siglo XX	2002	Alianza	33	Monograph	Art, History
Davis, F. (tr. L. Mourglier)	La comunicación no verbal	2010	Alianza	31	Monograph	Psychology
Martín Valverde, A., Rodríguez-Sañudo, F., García Murcia, J.	Derecho del trabajo	2011	Tecnos	31	Monograph	Law
Gardner, H. (tr. J. Meler Ortí)	Educación artística y desarrollo humano	2011	Paidós	30	Monograph	Society, Psychology, Pedagogy
Maravall, J.A.	La cultura del Barroco	2008	Ariel	29	Monograph	Society, Art
Lasarte Álvarez, C.	Curso de derecho civil patrimonial	2011	Tecnos	29	Textbook	Law
Peña, D.	Fundamentos de estadística	2008	Alianza	28	Textbook	Statistics
Savater, F.	El valor de educar	2008	Ariel	28	Monograph	Society, Philosophy
Pérez Royo, J.	Curso de derecho constitucional	2010	Marcial Pons	28	Textbook	Law
Querol Fernández, M. A.	Manual de gestión del patrimonio cultural	2010	Akal	27	Textbook	Archaeology
Castells, M. (tr. M. Hernández Díaz)	Comunicación y poder	2009	Alianza	27	Monograph	Sociology
Bercovitz Rodríguez-Cano, A.	Apuntes de derecho mercantil	2002	Aranzadi	27	Monograph	Law
Ruiz Ramón, F.	Historia del teatro español	2005	Cátedra	27	Monograph	Literary studies
Torrojat, E.	Razón y ser de los tipos estructurales	2007	CSIC	27	Monograph	Engineering
Pastor Ridruejo, J. A.	Curso de derecho internacional público y organizaciones internacionales	2008	Tecnos	27	Textbook	Law

The top 15 sites and domains hosted about two thirds (67.9%) of the syllabus mentions (Table VII). The sampled books were mainly cited in online syllabi from Universidad de la Laguna, Universidad de Castilla-La Mancha, Universidad de Jaén, Universidad de Granada, Universidad de Valladolid, Universidad de Zaragoza and Universidad de Alicante. The first two universities have created a specific domain for hosting their academic course syllabi, resulting to be the most citing domains.

The top and second level domains that mentioned the sample were almost identical (Table VIII). Over half (58%) of syllabus mentions came from ".es" (Spain), a fourth (21%) from ".com" (general) and 7% from ".edu". The ".com" domains tend to belong to private higher education centres linked to Catholicism or focusing on Theology, as well as music conservatoires, art schools and research centres assigned to state universities. Although the ".edu" domain is usually related to U.S. higher education institutions, none of the eleven institutions mentioning the sampled books is American. They all are Spanish universities, except one that is from Mexico.

4. DISCUSSION

The objective of this study was to examine whether online syllabus mentions could be useful impact evidence for a set of Spanish-language books published by Spanish scholarly publishers that were prestigious or productive in social sciences and humanities fields. The methods have some limitations. Course syllabi tend to distinguish between required and supplementary readings. Materials included as basic reading have a higher educational value for that course but this study did not distinguish between the two types. The results might be influenced by the selection of sample, and may not apply to books from less prestigious or smaller presses. Whilst there is an increasing trend to post syllabi online, the proportion of syllabi available online is unknown. Moreover, some of them are likely to have restricted access, hiding their references, or they cannot be found by search engines. Finally, the reasons that lead instructors to suggest reading material are not well understood and may produce selection biases.

The *first research question* involved analysing the accuracy of a Spanish-language method developed

Table VII. Top-15 sites and domains from which syllabus mentions come

University	Site	URLs	%	University	Domain	URLs	%
U. de La Laguna	ull.es	1201	10,5%	U. de La Laguna	eguia.ull.es	754	6,6%
U. de Castilla-La Mancha	uclm.es	848	7,4%	U. de Castilla-La Mancha	guiae.uclm.es	686	6,0%
U. de Jaén	ujaen.es	729	6,4%	U. de Jaén	uvirtual.ujaen.es	524	4,6%
U. de Granada	ugr.es	670	5,9%	U. de Alicante	cv1.cpd.ua.es	501	4,4%
U. de Valladolid	uva.es	572	5,0%	U. de Zaragoza	titulaciones.unizar.es	486	4,3%
U. de Zaragoza	unizar.es	543	4,8%	U. de La Laguna	www.ull.es	388	3,4%
U. de Alicante	ua.es	515	4,5%	U. de Cantabria	web.unican.es	372	3,3%
U. de Valencia	uv.es	473	4,1%	U. de Valladolid	alojamientos.uva.es	331	2,9%
U. de Cantabria	unican.es	376	3,3%	U. Pablo de Olavide	www.upo.es	307	2,7%
U. Internacional de La Rioja	unir.net	321	2,8%	U. de Santiago de Compostela	www.usc.es	298	2,6%
U. Pablo de Olavide	upo.es	308	2,7%	U. Católica de Murcia	www.ucam.edu	292	2,6%
U. Complutense Madrid	ucm.es	305	2,7%	U. Internacional de La Rioja	gestor.unir.net	265	2,3%
U. de Santiago de Compostela	usc.es	302	2,6%	U. de Valencia	webges.uv.es	250	2,2%
U. Católica de Murcia	ucam.edu	292	2,6%	U. de Almería	cvirtual.ual.es	212	1,9%
U. de Alcalá	uah.es	292	2,6%	U. de Alcalá	www.uah.es	210	1,8%

Table VIII. Top and second level domains hosting syllabus mentions

Top level	Sites	%	Second level	Sites	%
es	93	58,1%	es	93	58,1%
com	34	21,3%	com	34	21,3%
edu	11	6,9%	edu	11	6,9%
org	8	5,0%	org	8	5,0%
cat	5	3,1%	cat	5	3,1%
eus	2	1,3%	eus	2	1,3%
net	2	1,3%	net	2	1,3%
eu	1	0,6%	eu	1	0,6%
mx	1	0,6%	edu.mx	1	0,6%
co	1	0,6%	co	1	0,6%
gal	1	0,6%	gal	1	0,6%
ec	1	0,6%	ec	1	0,6%

to identify correct syllabus mentions from Bing API search engine and, therefore, the results in this study are limited to the Bing API search results.

Manual checks of all 52,716 Bing API syllabus query results for 15,117 books found that the Spanish rules to filter our incorrect matches rejected very few correct matches since 99.5% of the rejected Bing results were false mentions. Nevertheless, ignoring matches that could not be assessed, 74.7% of the matches judged by the filtering rules to be correct syllabus mentions were genuinely correct matches. The method is therefore not accurate enough overall to be relied upon for assessing individual papers. For this purpose, manual checking should be used. Nevertheless, the Bing API matches judged to be incorrect by the filtering rules can be safely ignored, reducing by about half the amount of manual checking needed. Thus, when assessing the educational impact of a set of Spanish books, the following stages are recommended.

1. Identify potential syllabus matches using the Spanish-language Bing API queries in Webometric Analyst.

2. Use the Spanish-language rules in Webometric Analyst to filter out matches that can automatically be ruled out as incorrect.

3. Manually check the remaining matches for accuracy.

The manual filtering stage 3 above can be ignored if the syllabus mention data is to be used to compare the educational impact of groups of articles if there is no reason to believe that there would be more incorrect matches for one group

than for another. Thus, for example, if comparing the proportion of books (Thelwall, 2017) with at least one syllabus mention between Spanish departments or publishers, then only the first two stages above would be needed, making the method a practical one. For safety, a random sample of the matches judged to be correct by the filtering could be checked to assess whether there was likely to be differing proportions of incorrect matches between groups.

The failure of the filtering rules to remove many incorrect matches is a problem. To classify a mention as correct or false, the software checked if the terms used in the query (first author's last name, book title and publisher) were within the titles and description of webpages. Mentions sometimes referred to publications with the same or a very similar title to the sampled books, but authored by another person. It is difficult to define rules for an automatic method to detect such differences.

It was also difficult to detect false mentions when the sampled books had also been edited by another publisher, which frequently occurred for classic books, such as those by Charles Darwin, Hannah Arendt, and Sigmund Freud. The best practical solution in these cases is to avoid including this type of book in samples analysed. Such books would presumably not be of interest for most evaluation purposes so this should not be a problem in practice. Moreover, books with very general titles that coincide with departmental names, course titles or sentences within course descriptions should perhaps be singled out for more intensive checking than other books in a sample.

The *second research question* was to examine the proportion of books that had at least one syllabus mention to know if syllabus mentions are common enough to be used as evidence of the educational utility of books. Since about a fifth (19%) of the books had been recommended at least once in an online course syllabus, it seems that there is enough data to differentiate between books that have educational utility and those that don't (or are less likely to have educational value). An analysis of the 23 books recommended most often in online syllabi showed that they are mostly single-authored humanities monographs that are not translations into Spanish. These seem to represent genuine Spanish-language contributions with educational value.

Books with multiple editions are common in some fields, such as law (Moed et al., 2002). This brings conceptual issues, such as whether different editions of the same book should be considered as different publications (Gorraiz et al., 2013). The answer might depend on whether the revised edition has substantial modifications or whether the authorship has changed. In this study, syllabus mentions were judged to be valid if they referred to the sampled books, regardless of the edition. Since 39% of the verified syllabus mentions to the sampled books referred to another publication year, the method would be much less accurate if mentions of different editions should be ignored.

Another problem regarding the gathering of syllabus mentions is the lack of basic information (author(s), title, publisher, publication year, edition number) or mistakes in the bibliographic references in scholarly syllabi. If the edition and year are not included, it might be assumed that any edition could be consulted by students. Conversely, students might be expected to find the most recent version. Since 7.8% of the verified syllabus mentions excluded the publication year of the sampled books in the bibliographic reference, this is a small but important problem. The latest edition (without specifying the year) was recommended in about a fifth of the 7.8% of cases and exclusively for books about law.

A revised edition of a book includes changes in the content and (usually) a new ISBN. Revised editions might be released by a different publisher. For instance, a mention referred to the book "Escandell Vidal, M. V. *Introducción a la pragmática*. Barcelona: Anthropos, 1992", but the sampled book was "Escandell Vidal, M. V. *Introducción a la pragmática*. Barcelona: Ariel, 2006". This occurred for 296 (2%) books in the sample. It cannot be easily identified whether a specific book has changed publisher. This was common for classic books (e.g. René Descartes's 'El discurso del método' or Charles Darwin's 'El origen de las

especies'), which are mainly translations. These issues affected the accuracy of the method used.

Book translation is another conceptual issue to be addressed in the book assessment (Torres-Salinas et al., 2014a). The original book and its translations should be considered when assessing its impact. The syllabus mentions of the translated books in the sample used here only reflect the educational utility of the Spanish version. From a research evaluation perspective, the syllabus mentions also conflate the value of the translation and that of the content of the original version. About 16% of books published in Spain in 2016 were translations, mainly from English (MECD, 2017).

Self-citations in course syllabi are a potential concern for research evaluation. This occurs when the instructor of a course recommends their own publications in the course syllabus.

The *third research question* related to the main characteristics of course syllabi mentioning Spanish texts. The syllabus mentions came from 379 web domains and 160 websites. In some cases, course syllabi had been posted to multiple domains within the same site, duplicating the syllabus mentions. For instance, syllabus mentions were found in twenty domains within the Universidad de La Laguna website ('.ull.es'). Although this university has a domain to host its course syllabi (eguia.ull.es), syllabi have also been uploaded to the domains of different faculties (e.g.: filosofi.webs.ull.es, fcfrmc.webs.ull.es). Duplicate course syllabi within a university website occurred for different academic years (e.g.: 2015-2016, 2016-2017), formats (e.g.: .html and .pdf) and languages (e.g.: Spanish-Galician, Spanish-Basque, Spanish-Catalan). Thus, the total number of syllabus mentions might be misleading for some books.

A solution to prevent duplication might be to count web sites citing the sampled books rather than the individual web pages. Nevertheless, a book might legitimately be in course syllabi from different subjects within the same university and this strategy would treat these as duplicates. Moreover, this would not solve completely the duplication problem, as results in this study show that some universities have websites hosted on multiple top-level domains, such as the Universidad de Castilla-La Mancha (uclm.es, uclm.edu, uclm.com) or the Universidad Católica de San Antonio de Murcia (ucam.edu, ucamonline.net), which would stop the website rule from working properly.

Some universities located in Spanish areas with two official languages used the national domain ('.es') to offer the content in Spanish and the regional domain ('.cat', '.eus' or '.gal') to make available the content in Catalan, Basque or Galician. Some

examples include the Universitat de Lleida (udl.es, udl.cat), Universitat de les Illes Balears (uib.es, uib.cat), Universidade da Coruña (udc.es, udc.gal) or Deustuko Unibertsitatea (deusto.es, deusto.eus). However, there were other universities with only the regional domain, such as the Universidad del País Vasco (ehu.eus) or the Universidade de Vigo (uvigo.gal) and other universities with a general top-level domain, such as the Universitat de Barcelona (ub.edu), Universitat Oberta de Catalunya (uoc.edu) and Universitat Pompeu Fabra (upf.edu).

Most of the syllabus mentions came from the .es top-level domain. This is reasonable since the books in the sample were in Spanish and the most common terms describing a syllabus in Spain were used for the creation of the automatic queries ("guia docente" and "guia académica) and for the development of the rules defined in the software. Several factors might explain the scarcity of domains from Latin American countries (e.g. '.ar', '.bo', '.mx'), including the lack of university policies for sharing course syllabi online, the use of different terms to refer to course syllabi in these countries or a trend to recommend books published by Latin American publishers.

5. CONCLUSIONS

This article introduced and assessed a method to estimate the teaching value of Spanish-language monographs from their mentions in online syllabi. The automatic method is imperfect because a quarter of its results are false matches. This can be remedied by manual checks if syllabus mention counts are needed for individual articles. The manual checks may not be necessary if comparing sets of books rather than individual books, however. Two stage and three stage approaches for this are described above. Whilst it would be impractical to manually check tens of thousands of results, as reported above, individual academics and departments would only need to check the results for their own books, which is a much smaller task.

Whilst the proportion of the sampled Spanish academic books with at least one syllabus mention is relatively small (19%), the percentage is high enough for syllabus mentions to be used to help distinguish between books that have some educational impact and books that may have none. The proportion cited is unsurprisingly lower than for English language books from prestigious publishers (Kousha and Thelwall, 2016a), presumably because there are fewer Spanish-language universities and probably some of them do not make available their course syllabi or they are not traceable. Thus, syllabus mentions might not be frequent enough to be useful for books in languages that are less frequently spoken than Spanish.

In conclusion, the results suggest that online syllabus mentions can be a useful source of educational impact evidence for languages other than English and published by presses that are less internationally prestigious. Thus, humanities researchers working in common non-English languages now have a new plausible source of evidence for the impact of their books.

6. REFERENCES

ANECA (2017). *Méritos evaluables para la acreditación nacional para el acceso a los cuerpos docentes universitarios de Profesor Titular de Universidad y Catedrático de Universidad. Criterios de Arte y Humanidades.* Available at: http://www.aneca.es/Programas-de-evaluacion/ACADEMIA/Criterios-de-evaluacion-noviembre-2017

Archambault, E.; Vignola-Gagné; É.; Côté, G.; Larivière, V.; Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68 (3), 329-342. https://doi.org/10.1007/s11192-006-0115-z

Barron, P. (2011). The library of the future: Google's vision for books. *Learned Publishing*, 24 (3), 197-201. https://doi.org/10.1087/20110307

Book Citation Index (2017). Putting books back into the library. Completing the research picture. The Book Citation Index. Available at: http://wokinfo.com/products_tools/multidisciplinary/bookcitationindex/

Bosch, A. van den; Bogers, T.; Kunder, M. de (2016). Estimating search engine index size variability: A 9-year longitudinal study. *Scientometrics*, 107 (2), 839-856. https://doi.org/10.1007/s11192-016-1863-z

Caerols Mateo, R.; Verdú Ruiz, S.; Viñarás Abad, M. (2017). Las artes en la educación superior: la evaluación en España de la actividad investigadora en Bellas Artes a través de los indicadores de calidad. *Revista Española de Documentación Científica*, 40 (4), e191. https://doi.org/10.3989/redc.2017.4.1394

Cordón-García, J.A.; Gómez-Díaz, R.; Rodríguez-García, A.; Merchán Sánchez-Jara, J.; Mangas-Vega, A. et al., (2017). Visibility of scientific monographs in the academic field: the institutional assessment of research. In: *5th International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM*. Spain, Cádiz. https://doi.org/10.1145/3144826.3145380

Cronin, B.; Barre, K. La (2004). Mickey Mouse and Milton: Book publishing in the humanities. *Learned Publishing*, 17 (2), 85-98. https://doi.org/10.1087/095315104322958463

Cummings, J.A.; Bonk, C.J.; Jacobs, F.R. (2002). Twenty-first century college syllabi: Options for online communication and interactivity. *The Internet and Higher Education*, 5(1), 1-19. https://doi.org/10.1016/S1096-7516(01)00077-X

Dimitrov, S.; Zamal, F.; Piper, A.; Ruths, D. (2015). Goodreads vs Amazon: The Effect of Decoupling Book Reviewing and Book Selling. In: *Ninth International AAAI Conference on Web and Social Media*, pp. 602-605. England: Oxford University. Available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10557/10452

Elsevier (2018). Scopus Content Coverage Guide. Available at: https://www.elsevier.com/solutions/scopus/content

Engels, T.C.E.; Ossenblok, T.L.B.; Spruyt, E.H.J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000-2009. *Scientometrics*, 93 (2), 373-390. https://doi.org/10.1007/s11192-012-0680-2

Garand, J.C.; Giles, M.W. (2011). Ranking Scholarly Publishers in Political Science: An Alternative Approach. *PS: Political Science and Politics*, 44 (2), 375-383. https://doi.org/10.1017/S1049096511000229

Giménez-Toledo, E. (ed.). (2017). *La edición académica española. Indicadores y características*. Madrid: Federación de Gremios de Editores de España. Available at: http://ilia.cchs.csic.es/SPI/spi-fgee/docs/EAEV2.pdf

Giménez-Toledo, E.; Tejada-Artigas, C.; Mañana-Rodríguez, J. (2013). Evaluation of scientific books' publishers in social sciences and humanities: Results of a survey**.** *Research Evaluation*, 22 (1), 64-77. https://doi.org/10.1093/reseval/rvs036

Giménez-Toledo, E.; Mañana-Rodríguez, J.; Engels, T.C.E.; Ingwersen, P.; Siverten, G.; Verleysen, F.T.; Zuccala, A. (2016). Taking scholarly books into account: current developments in five European countries. *Scientometrics*, 107 (2), 685-699. https://doi.org/10.1007/s11192-016-1886-5

Gorraiz, J.; Purnell, P.J.; Glänzel, W. (2013). Opportunities for and limitations of the book citation index. *Journal of the American Society for Information Science and Technology*, 64 (7), 1388–1398. https://doi.org/10.1002/asi.22875

Gorraiz, J.; Gumpenberger, C.; Purnell, P.J. (2014). The power of book reviews: a simple and transparent enhancement approach for book citation indexes. *Scientometrics*, 98 (2), 841-852. https://doi.org/10.1007/s11192-013-1176-4

Gutiérrez García, C.; Vega-Álvarez, M.P.; Gómez-Alonso, M.T.; Pérez-Pueyo, A. (2016). Análisis de las bibliografías de las asignaturas de pedagogía y didáctica de los grados en ciencias de la actividad física y el deporte en España. *Retos: nuevas tendencias en Educación Física, Deporte y Recreación*, (30), 138-143. Available at: https://recyt.fecyt.es/index.php/retos/article/view/45045

Halevi, G.; Nicolas, B.; Bar-Ilan, J. (2016). The Complexity of Measuring the Impact of Books. *Publishing Research Quarterly*, 32 (3), 187-200. https://doi.org/10.1007/s12109-016-9464-5

Hartley, J. (2006). Reading and writing book reviews across the disciplines. *Journal of the American Society for Information Science and Technology*, 57 (9), 1194-1207. https://doi.org/10.1002/asi.20399

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometric*s, 44 (2), 193-215. https://doi.org/10.1007/BF02457380

Huang, M.-h.; Chang, Y.-w. (2008). Characteristics of research output in social sciences and humanities: From a research evaluation perspective. *Journal of the American Society for Information Science and Technology*, 59 (11), 1819-1828. https://doi.org/10.1002/asi.20885

Kousha, K.; Thelwall, M. (2008). Assessing the impact of disciplinary research on teaching: An automatic analysis of online syllabuses. *Journal of the American Society for Information Science and Technology*, 59 (13), 2060-2069. https://doi.org/10.1002/asi.20920

Kousha, K.; Thelwall, M. (2015). An automatic method for extracting citations from Google Books. *Journal of the Association for Information Science and Technology*, 66 (2), 309-320. https://doi.org/10.1002/asi.23170

Kousha, K.; Thelwall, M. (2016a). An automatic method for assessing the teaching impact of books from online academic syllabi. *Journal of the Association for Information Science and Technology*, 67 (12), 2993-3007. https://doi.org/10.1002/asi.23542

Kousha, K.; Thelwall, M. (2016b). Can Amazon.com reviews help to assess the wider impacts of books? *Journal of the Association for Information Science and Technology*, 67 (3), 566-581. https://doi.org/10.1002/asi.23404

Kousha, K.; Thelwall, M.; Abdoli, M. (2017). Goodreads reviews to assess the wider impacts of books. *Journal of the Association for Information Science and Technology*, 68 (8), 2004-2016*.* https://doi.org/10.1002/asi.23805

Larivière, V.; Macaluso, B. (2011). Improving the coverage of social science and humanities researchers' output: The case of the Érudit journal platform*. Journal of the American Society for Information Science and Technology*, 62 (12), 2437-2442. https://doi.org/10.1002/asi.21632

Leydesdorff, L.; Felt, U. (2012). Edited Volumes, Monographs and Book Chapters in the Book Citation Index (BKCI) and Science Citation Index (SCI, Sosci, A&HCI). *Journal of Scientometric Research*, 1 (1), 28-34. https://doi.org/10.5530/jscires.2012.1.7

López-Navarro, I.; Moreno, A.I.; Quintanilla, M.Á.; Rey-Rocha, J. (2015). Why do I publish research articles in English instead of my own language? Differences in Spanish researchers' motivations across scientific domains. *Scientometrics*, 103 (3), 939-976. https://doi.org/10.1007/s11192-015-1570-1

Marta-Lazo, C.; Grandío Pérez, M.M.; Gabelas Barroso, J.A. (2014). La educación mediática en las titulaciones de Edu-

cación y Comunicación de las universidades españolas. Análisis de los recursos recomendados en las guías docentes. *Revista de Comunicación Vivat Academia*, XVII (126), 63-78. https://doi.org/10.15178/va.2014.126.63-78

Mas-Bleda, A.; Thelwall, M. (2016). Can alternative indicators overcome language biases in citation counts? A comparison of Spanish and UK research. *Scientometrics*, 109 (3), 2007-2030. https://doi.org/10.1007/s11192-016-2118-8

MECD (2017). *Panorámica de la edición española de libros 2016*. Ministerio de Educación, Cultura y Deporte: Secretaría General Técnica. Available at: https://sede.educacion.gob.es/publiventa/panoramica-de-la-edicion-de-libros-2016/edicion/21061C

Moed, H.F.; Luwel, M.; Nederhof, A.J. (2002). Towards Research Performance in the Humanities. *Library Trends*, 50 (3), 498-520. https://www.ideals.illinois.edu/handle/2142/8415

Nederhof, A.J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66 (1), 81-100. https://doi.org/10.1007/s11192-006-0007-2

Osca-Lluch, J.; Veyrat, A.; Morales, J. (2013). El consumo de información en Humanidades. *Arbor,* 189 (760), a026. https://doi.org/10.3989/arbor.2013.760n2012

Prieto-Paíno, C. (2013). *Tipología, características y género de las bibliografías recomendadas. El caso de Humanidades, ciencias sociales y especialidades científico-técnicas en la Universidad de Salamanca* [Dissertation]. Salamanca: Universidad de Salamanca. Available at: https://gredos.usal.es/jspui/handle/10366/123403?mode=full

Romanos de Tiratel, S. (2000). Conducta informativa de los investigadores argentinos en Humanidades y Ciencias Sociales. *Revista Española de Documentación Científica*, 23 (3), 267-285. https://doi.org/10.3989/redc.2000.v23.i3.324

San-Martín Gutiérrez, S.; Jiménez Torres, N.; Jerónimo Sánchez-Beato, E. (2016). La evaluación del alumnado universitario en el Espacio Europeo de Educación Superior. *Aula Abierta*, 44 (1), 7-14. https://doi.org/10.1016/j.aula.2015.03.003

Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometric*s, 107 (2), 357-368. https://doi.org/10.1007/s11192-016-1845-1

Sivertsen, G.; Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. *Scientometrics*, 91 (2), 567-575. https://doi.org/10.1007/s11192-011-0615-3

Slattery, J.M.; Carlson, J.F. (2005). Preparing an effective syllabus: Current best practices. *College Teaching*, 53 (4), 159-164. https://doi.org/10.3200/CTCH.53.4.159-164

Sorli Rojo, A.; Mochón Bezares, G.; Martín-Carretero, C. (2011). Reseñas en revistas científicas españolas de Ciencias Sociales y Humanidades: análisis de la pro-

ducción entre 2007 y 2009. *Revista Española de Documentación Científica*, 34 (4), 526-544. https://doi.org/10.3989/redc.2011.4.841

Spink, A.; Robins, D.; Schamber, L. (1998). Use of Scholarly Book Reviews: Implications for Electronic Publishing and Scholarly Communication. *Journal of the American Society for Information Science and Technology*, 49 (4), 364-374. https://doi.org/10.1002/(SICI)1097-4571(19980401)49:4<364::AID-ASI6>3.0.CO;2-3

Thelwall, M. (2017). Three practical field normalised alternative indicator formulae for research evaluation. *Journal of Informetrics*, 11 (1), 128-151. https://doi.org/10.1016/j.joi.2016.12.002

Thelwall, M.; Sud, P. (2012). Webometric research with the Bing Search API 2.0. *Journal of Informetrics*, 6 (1), 44-52. https://doi.org/10.1016/j.joi.2011.10.002

Thompson, J.W. (2002). The Death of the Scholarly Monograph in the Humanities? Citation Patterns in Literary Scholarship. *Libri*, 52 (3), 121-136. https://doi.org/10.1515/LIBR.2002.121

Torres-Salinas, D.; Moed, H.F. (2009). Library Catalog Analysis as a tool in studies of social sciences and humanities: An exploratory study on published book titles in Economics. *Journal of Informetrics*, 3 (1), 9-26. https://doi.org/10.1016/j.joi.2008.10.002

Torres-Salinas, D.; Robinson-García, N.; Jiménez-Contreras, E.; Delgado López-Cózar, E. (2012). Towards a "Book Publishers Citation Reports". First approach using the "Book Citation Index". *Revista Española de Documentación Científica*, 35(4), 615–624. https://doi.org/10.3989/redc.2012.4.1010

Torres-Salinas, D.; Robinson-García, N.; Cabezas-Clavijo, Á.; Jiménez-Contreras, E. (2014a). Analyzing the citation characteristics of books: edited books, book series and publisher types in the book citation index. *Scientometrics*, 98 (3), 2113-2127. https://doi.org/10.1007/s11192-013-1168-4

Torres-Salinas, D.; Robinson-García, N.; Campanario, J.M.; Delgado López-Cózar, E. (2014b). Coverage, field specialisation and the impact of scientific publishers indexed in the Book Citation Index. *Online Information Review*, 38 (1), 24-42. https://doi.org/10.1108/OIR-10-2012-0169

UNE (2016). *Las editoriales universitarias en cifras 2016*. Available at: http://www.une.es/media/Ou1/Image/webdiciembre2016/InformeComercioInteriorUNE-2015conlogoMECD.pdf

Vila-Suero, D.; Villazón-Terrazas, B.; Gómez-Pérez, A. (2012). datos.bne.es: A library linked dataset. *Semantic Web Journal,* 4 (3), 307-313. Available at: http://oa.upm.es/14467/1/4.datosBNEdataset.pdf

White, H.D.; Boell, S.K.; Yu, H.; Davis, M.; Wilson, C.S.; 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*, 60 (6), 1083-1096. https://doi.org/10.1002/asi.21045

Williams, L.M.; Cody, S.A.; Parnell, J. (2004). Prospecting for New Collaborations: Mining Syllabi for Library Service Opportunities. *The Journal of Academic Librarianship*, 30 (4), 270-275. https://doi.org/10.1016/j.acalib.2004.04.009

Zuccala, A.A; Leeuwen, T. van (2011). Book reviews in humanities research evaluations. *Journal of the American Society for Information Science and Technology*, 62 (10), 1979-1991. https://doi.org/10.1002/asi.21588

Zuccala, A.A.; Guns, R.; Cornacchia, R.; Bod, R. (2015a). Can we rank scholarly book publishers? A bibliometric experiment with the field of history. *Journal of the Association for Information Science and Technology,* 66 (7), 1333-1347. https://doi.org/10.1002/asi.23267

Zuccala, A.A.; Verleysen, F.T.; Cornacchia, R.; Engels, T.C. (2015b). Altmetrics for the humanities: Comparing Goodreads reader ratings with citations to history books. *Aslib Journal of Information Management*, 67 (3), 320-336. https://doi.org/10.1108/AJIM-11-2014-0152