An Automatic Method for Assessing the Teaching Impact of Books from Online Academic Syllabi[1]

Kayvan Kousha and Mike Thelwall

Statistical Cybermetrics Research Group, School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK.

E-mail: {k.kousha, m.thelwall}@wlv.ac.uk

Scholars writing books that are widely used to support teaching in higher education may be undervalued due to a lack of evidence of teaching value. Whilst sales data may give credible evidence for textbooks, it may poorly reflect educational uses of other types of books. As an alternative, this article proposes a method to automatically search for mentions of books in online academic course syllabi based on Bing searches for syllabi mentioning a given book, filtering out false matches through an extensive set of rules. The method had an accuracy of over 90% based on manual checks of a sample of 2,600 results from the initial Bing searches. Over a third of about 14,000 monographs checked had one or more academic syllabus mention, with more in the arts and humanities (56%) and social sciences (52%). Low but significant correlations between syllabus mentions and citations across most fields, except the social sciences, suggest that books tend to have different levels of impact for teaching and research. In conclusion, the automatic syllabus search method gives a new way to estimate the educational utility of books in a way that sales data and citation counts cannot.

Introduction

Although there is confusion about the exact nature of teaching scholarship (Kreber, 2002; Boshier, 2009), there is a broad consensus that contributions to teaching are undervalued (Boyer, 1990; Jenkins, 1995; Healey, 2000; Kreber & Cranton, 2000; Chalmers, 2011) and undermined by academic rewards systems (Altbach & Lewis, 1996; Ramsden & Martin, 1996; Skelton, 2004; Young, 2006; see also Jenkins, 1995). This may be due, at least in part, to teaching contributions being less visible than research outputs and less easy to get clear quantitative indicators for. There are many different types of materials that can support teaching (Fincher et al., 2000, p. 888), but among them textbooks are particularly important and monographs can also be useful (Gurung & Martin, 2011; Gurung, Landrum, & Daniel, 2012). Nevertheless, research priorities can discourage academics from writing student focused publications, such as textbooks (Jenkins, 1995) and monographs. Hence there is a need to recognise teaching-related activities and to more frequently reward academics that can demonstrate contributions to teaching as well as research.

Citation analysis has been extensively used to estimate the scholarly impact of academic publications. There do not seem to be any recognised methods to estimate the teaching impact of publications, however, such as textbooks and monographs, which could be used in education. Whilst sales may be a reasonable indicator of educational utility for textbooks, monographs could be read by

---

researchers, practitioners or the general public and hence monograph sales do not always reflect educational uptake. Expert judgments are a logical alternative but would presumably be time-consuming and expensive. For instance, reviewers in the 2008 UK Research Assessment Exercise (RAE) had to judge the research contributions of up to 100 books in some humanities and social sciences subject areas (Kousha, Thelwall, & Rezaie, 2011), and it seems unlikely that they could realistically also evaluate the impact of this many books on teaching. Moreover, opinions of researchers about the teaching value of publications might be subjective, perhaps more subjective than assessments of research (Weller, 2001). Student surveys could also help to assess the value of textbooks but students often do not read their assigned reading (e.g., Sikorski et al., 2002; Sharma, Van Hoof, & Pursel, 2013). For instance, less than 30% of psychology students in one study read course textbooks (Clump, Bauer, & Bradley, 2004). Moreover, student evaluations of teaching and instruction are not necessarily accurate (Spooren, Brockx, & Mortelmans, 2013). A range of alternative metrics has been suggested recently for various aspects of the social usage of articles in response to the research focus of conventional citation metrics (e.g., Priem & Hemminger, 2010; Priem, Piwowar, & Hemminger, 2012; Thelwall, Haustein, Larivière, Sugimoto, 2013), but none of these focus on teaching value. Attempts have also been made to examine the wider impacts of books using library holdings statistics (White, Boell, Yu et al. 2009), book reviews (Gorraiz, Gumpenberger, & Purnell 2014; Kousha & Thelwall, in press), publisher prestige (Giménez-Toledo, Tejada-Artigas, & Mañana-Rodríguez, 2013; Zuccala, Guns, Cornacchia, & Bod, in press 2014) and online reference managers (Hammarfelt, 2014; Kousha & Thelwall, in press) , but these again do not focus on teaching value.

Academic syllabus mentions can be an important indication of teaching scholarship success (Albers, 2003; Thompson, 2007) and can help to evaluate both instructors and course programs (Smith & Razzouk, 1993; Parkes & Harris, 2002). An analysis on 145 syllabi at one American university covering 100 courses across different disciplines, for instance, showed that 86% of syllabi had listed some required reading (Eberly, Newton, & Wiggins, 2001). Instructors that provide students with a list of required readings for courses make expert judgements (e.g., Weiten, 1988; Landrum & Hormel, 2002) and so presence in reading lists can be used to help assess teaching utility (Kousha & Thelwall, 2008). The current study investigates whether counts of mentions in online academic syllabi can be automated in a way that accurately captures the teaching influence of academic monographs and whether the results reflect teaching impacts that would be overlooked by citation counting. If so, the method could help academics or evaluators to be informed about the educational influence of academic books.

Book Impact Assessment

Citation Metrics

Citation Databases: Citations to books can be found using manual checks of reference lists in publications (e.g., Cullars, 1998; Krampen, Becker, Wahner, & Montada, 2007) or cited reference searches, for example in the Web of Science (WoS) (e.g., Bar-Ilan, 2010; Butler & Visser, 2006). Nonetheless, the former method is very time-consuming and therefore impractical for large-scale bibliometric analyses and the latter is largely dependent on citations from academic journals and selected serials. Moreover, journal-based citation metrics have limited value for book impact assessment due to the importance of citations from books to books, especially in the arts, humanities and social sciences (Hicks, 1999, Archambault, Vignola-Gagné, Côté, Larivière, & Gingras, 2006, Nederhof, 2006; Huang & Chang, 2008). The Thomson Reuters Book Citation Index (BKCI) was created in response to an earlier idea (Garfield, 1996), and now indexes the references in about 60,000 books (Book Citation Index, 2014). Similarly, Elsevier's Scopus indexes the references in about 40,000 books (Scopus Content, 2014). Nevertheless, their coverage of scholarly books is far from comprehensive in terms of different languages and publishers. For instance, about 97% of BKCI-indexed books are in

English and about 75% of their publishers are from the USA and England (Torres-Salinas et al., 2014). The above problems in conjunction with other issues, such as the lack of aggregated citation counts for edited volumes, suggest that BKCI should only be used cautiously for evaluative purposes (Leydesdorff & Felt, 2012; Gorraiz, Purnell, & Glänzel, 2013).

Google Books: Despite its lack of a citation index, Google Books (GB) has been used to extract citations from digitised books (Kousha & Thelwall, 2009; Kousha, Thelwall, & Rezaie, 2011) and an automatic method has been developed to identify GB citations with high levels of accuracy and coverage for book impact assessment (Kousha & Thelwall, 2014). Two investigations have found that GB citations to books were much more numerous than citations in conventional citation databases (Scopus and BKCI) in the humanities and in some social sciences but not in the sciences. GB citations to 1,000 books submitted to the 2008 UK RAE in seven book-oriented disciplines were 1.4 times more numerous than Scopus citations and in some humanities fields (Law, History and, Communication, Cultural and Media Studies) the median number of GB citations was three times higher than the median number of Scopus citations to the same set of books (Kousha, Thelwall, & Rezaie, 2011). Another study, of 14,500 BKCI-indexed monographs, found that GB citations to books were 103% to 137% more numerous than BKCI citations in some humanities fields (e.g., history, literature, law, philosophy and arts) but only 8% to 53% as numerous in the sciences (Kousha & Thelwall, in press 2014).

Non-Citation Metrics

Some books can have types of impacts that cannot be traced through citations. For instance, textbooks, introductory science books and novels may have educational or cultural impacts. As a result, attempts have been made to use non-bibliometric indicators to identify the wider impacts of books through library holdings statistics, publisher prestige, book reviews, and online reference managers. These may partly reflect educational value. For instance, library holdings statistics may reveal the information needs of teaching staff or students, academic book reviews may express the educational merits of books, and bookmarks in online reference managers may partly reflect the use of books for teaching and learning activities.

Libcitations: Torres-Salinas and Moed (2009) have used library catalogue statistics for books to estimate their impacts. White, Boell, Yu et al. (2009) used the term "libcitation" to refer to the number of libraries holding a book as an indicator of its "cultural benefit" (White, Boell, Yu et al. 2009, p. 1087). If book orders in academic libraries are mainly based on requests by teaching staff, researchers or students, then national or international library holdings catalogues could help to assess the value of books. A study of books from Australian history, philosophy, and political science departments and their libcitation counts in the Libraries Australia union catalog suggested that library holding statistics could reflect the cultural impact of books (White, Boell, Yu et al. 2009). Later studies have also found significant, but low, correlations between library holding counts and citation metrics (e.g., Linmans, 2010; Zuccala & Guns, 2013; Kousha & Thelwall, in press), or low and insignificant correlations (Cabezas-Clavijo et al., 2013). A larger-scale investigation also compared Scopus citations to 59,000 history books and 42,000 literature books with library holding counts from the Association of Research Libraries (ARL), non-ARL libraries and all libraries, finding low Spearman correlations (ranging from 0.288 to 0.244) (Zuccala & Guns, 2013). Another study compared *WorldCat* library holdings with BKCI and GB citations to 2,740 monographs, finding low significant correlations again in the social sciences (r=0.145 for BKCI and 0.234 for GB, n=759), and in the arts and humanities (r=0.141 for BKCI and 0.268 for GB, n=1,262) (Kousha & Thelwall, in press). These studies collectively suggest that the impact of books (including on teaching and culture) as reflected by libcitations is at best loosely related to their research impact and that library holdings may reflect a different type of impact to that of citation counts.

Book Reviews: Book reviews are important in the humanities (e.g., Spink, Robins, & Schamber, 1998; Dilevko et al., 2006; Hartley, 2006). For instance, there are about 15% more book reviews than journal articles in *Thomson Reuter's Arts and Humanities Citation Index* (Zuccala, & van Leeuwen, 2011). An investigation of 200 fiction books found the number of academic reviews in *the Book Review Index* and the number of OCLC library holdings to correlate significantly (r=0.620) (Shaw, 1991). Nicolaisen's (2002) study of 420 sociology books found that monographs with positive reviews tended to attract more *Social SciSearch* citations than do those with negative reviews. Gorraiz, Gumpenberger and Purnell (2014) calculated correlations between Thomson Reuters BKCI citations to books and the number of book reviews, finding much higher correlations in the humanities, such as literature (.637) and history (.608), than in science, such as chemistry (.127) and mathematics (.123). Hence, it seems that humanities book review metrics may partly reveal the scholarly impact of the books. Nevertheless, another study of online book reviews found significant, but low correlations between the numbers of Amazon reviews and BKCI citations to about 2,740 academic monographs, although this association was higher in the social sciences (.223) and the arts and humanities (.189) than in science (0.121) (Kousha & Thelwall, in press). Some book reviews, such as *Choice: Current Reviews for Academic Libraries,* explicitly express the educational merits of books by recommending the value of books for specific segments within higher education (e.g., "Graduate students" or "Upper/lower-division undergraduates") but these have not yet been systematically analysed yet.

Online Reference Managers: Online reference managers, such as *Mendeley*, *CiteULike* and *Zotero*, can be used by researchers, students and others to keep track of their reading material and can also provide evidence of its use. Mendeley reader counts correlate moderately with citation counts for published journals articles in different subject areas (e.g., Li & Thelwall, 2012; Li, Thelwall, & Giustini, 2012; Mohammadi & Thelwall, 2014; Zahedi, Costas, & Wouters, in press, 2014). In the medical sciences, however, citation counts and Mendeley reader counts correlate strongly (about 0.7) and the majority of articles (78%) have one or more readers (Thelwall & Wilson, in press). Mendeley reader counts may partly reflect educational uses of publications because, according to a survey of 860 Mendeley users, publications can be registered in Mendeley for teaching (25%) or other educational activities (13%) as well as for future publications (85%) (Mohammadi, Thelwall, & Kousha, in press). In contrast to academic journal articles, Mendeley users rarely register books and monographs in the site. For instance, only 7% out of 2,739 scientific monographs indexed by BKCI in 2008 had one or more Mendeley reader (Kousha & Thelwall, in press; see also: Hammarfelt, 2014).

Publisher Prestige: Publisher prestige can be used as an indicator of the value of books because better books tend to be published by more prestigious publishers (Donovan & Butler, 2007). An attempt has been made to formalise this by ranking scientific publishers in the social sciences and arts and humanities through a *'Book Publishers Citation Reports',* analogous to the *'Journal Citation Reports'* (Torres-Salinas et al., 2012, see also http://epuc.cchs.csic.es/SPI/). Scientific publisher prestige has mostly been investigated through surveys, however. Garand and Giles (2011), for instance, surveyed about 600 American academics in the political sciences, finding that Cambridge University Press, Princeton University Press, Oxford University Press, University of Chicago Press and Harvard University Press were the highest-profile university presses, in that order (Garand & Giles, 2011). A survey of 3,002 Spanish academics in the social sciences and humanities found that Oxford University Press, Cambridge University Press, Routledge, Springer and John Benjamins were the top-rated publishers, in that order (Giménez-Toledo, Tejada-Artigas & Mañana-Rodríguez, 2013). Using citation data from Scopus-indexed journals (2007-2011), Zuccala, Guns, Cornacchia and Bod, (in press 2014) ranked university presses in history, finding that Oxford University Press, Cambridge University Press, Routledge, Princeton University Press and Palgrave Macmillan were the top five publishers based on total citations and

citations per book in Scopus. Nevertheless, no studies have investigated the prestige of publishers in terms of the educational value of books.

Syllabus Citations: Although many studies of academic course syllabi have assessed pedagogical aspects of courses (e.g., Crittenden, & Wilson, 2006; Pieterse et al., 2009; Mishra, Day, Littles, & Vandewalker, 2011; Homa et al., 2013), only one has used syllabi for impact assessment of the cited publications. It gathered citations from online course syllabi for over 70,000 ISI-indexed articles published in 2003 across 12 science and social science disciplines (Kousha & Thelwall, 2008). The syllabus citations were searched for with automatically generated queries using the first author, article title, journal name, and publication year and then combining them with either "syllabus" or "reading list". Each query was submitted to the three search engines (Google, Yahoo!, Live Search[2]) via their Applications Programming Interfaces (APIs) and their results were combined to give the maximum possible coverage. Syllabus mentions seemed to be a useful indicator in some social sciences fields (e.g., political science and information and library science), but not in others, nor in any of the sciences. The study compared online syllabus mentions and WoS citations only for Information Science & Library Science articles, finding a low but significant Spearman correlation (r=0.231 n=1,987). The method used in this study estimated globally the proportion of non-syllabus mentions (i.e., incorrect results) from the raw search results based on manual checks of a random sample of about 100 URLs from each subject area. This method is not suitable for assessing individual publications, however, because it did not filter out bad matches.

Research questions

The main objective of the current study is to assess whether an automatic method can help the teaching impact assessment of books based on mentions of them in academic syllabi or course reading lists. The focus is monographs from Scopus and BKCI and correlations between syllabus mentions and citations counts are needed to help investigate the degree of similarity between the types of impact that they may reflect.

1. Can an automatic method accurately identify mentions of books in online academic syllabi or course reading list with sufficient coverage to be useful?
2. Do syllabus mentions correlate with citation counts for Scopus books and BKCI monographs and are there disciplinary differences between the magnitudes of the correlations?

Methods

To identify mentions of publications in online academic syllabi or course reading lists we automatically searched for monographs through predefined queries (see below) using the Bing Search API and used heuristics to filter out irrelevant results. The method was evaluated by manual checks to estimate recall and precision.

Data Sets

We investigated about 12,000 Scopus-indexed monographs published during 2005-2010 as the primarily source of research data and about 2,300 BKCI-indexed monographs published in 2008 for secondary analyses. Monographs were chosen from both databases because the citations to individual book chapters, edited books and volume series are not included in the citation counts to whole books in BKCI

---

[2] API search is no longer available for Yahoo! and Google only permits 100 searches per day. Microsoft's Live Search is known now as Bing.

(Leydesdorff & Felt, 2012) and Scopus. We selected the years 2005-2010 to give at least three years for monographs to be mentioned in academic syllabi or to be cited in other publications. We selected only English monographs because these dominate our sources and so we also developed the syllabus search method to capture mentions of books in English language academic syllabi or course reading lists.

To identify monographs we used document type search command "*DOCTYPE (bk)*" in the Scopus advanced search option. We excluded edited books, book series and book chapters to restrict the results to monographs. We also removed records with book editors or different volumes in the Scopus output based on the "*Editors*" and "*Volume*" fields or titles with terms such as "edition" or "volume" (e.g., second edition, vol. II) to limit the data set to individual monographs as far as possible. Eleven subject areas were selected across the arts and humanities, social sciences, sciences and medicine to cover a range of different fields.

For BKCI monographs, we used an alphabetical query[3] in the "*Publication Name*" field in the *Book Citation Index-Science (BKCI-S)* and *Book Citation Index-Social Sciences & Humanities (BKCI-SSH)*. We limited the results to 'Books' and the year to 2008 (i.e., the middle of the Scopus range). As for the Scopus data set, we removed edited books and volume series from the BKCI output (the *BE* and *VL* fields respectively). Related disciplines were combined based on their BKCI categories to represent three broad fields: science and medicine; social sciences; and arts and humanities. For instance, we merged *History, Religion, Government & Law, Literature, Philosophy, Anthropology, Music, Art, Theater, Film, Radio & Television* to form the arts and humanities set.

Some monographs in both the Scopus and BKCI data sets had multiple distinct subjects. For instance, the book "*Role of China in Global Dirty Industry Migration*" by Haitian Lu had been assigned three BKCI subjects (*Business & Economics; Environmental Sciences & Ecology; Government & Law*). We excluded books that fit into more than one of the three broad fields to avoid monographs with very diverse topics affecting the disciplinary analysis. We also excluded monographs with titles consisting of less than three words to avoid retrieving many false matches in the syllabus searches (see below).

Automatic Syllabus Searching and Filtering
We automatically generated and ran queries to identify mentions of the Scopus and BKCI monographs in online syllabi. The free *Webometric Analyst* software (http://lexiurl.wlv.ac.uk) was used to conduct automatic searches based on the Bing API and to report mentions of books in academic syllabi or course reading lists. We used Bing because it is the only major search engine that supports API searches on a sufficient scale. We tested different queries by combing different bibliographic information about the monographs with syllabus-relevant descriptors to retrieve mentions of books in online syllabi with high accuracy and coverage. We then limited the results to academic websites and used additional heuristics to include correct matches and to exclude false results. Details of the stages are given below.

Stage 1: Syllabus Query for API Searching
To locate mentions of monographs in syllabi or course reading lists through the Bing API, we searched for the first author last name and the first (up to) seven terms of the monograph title as a phrase search and combined them with either "syllabus" or "course description". For each record, Webometric Analyst ran the two queries at the same time and combined the results (after removing duplicates) using the OR operator (the vertical bar "|" in Webometric Analyst).

---

[3] . (A* OR B* OR C* OR D* OR E* OR F* OR G* OR H* OR I* OR J* OR K* OR L* OR M* OR N* OR O* OR P* OR Q* OR R* OR S* OR T* OR U* OR V* OR W* OR X* OR Y* OR Z* OR 0* OR 1* OR 2* OR 3* OR 4* OR 5* OR 6* OR 7* OR 8* OR 9*)

Tetlock "Expert political judgment: How good is it" "course description"**|**Tetlock "Expert political judgment: How good is it" "syllabus"

We removed Scopus monographs with single word titles (103, 0.8%) or two word titles (423, 3.3%) from the data set (e.g., "*Modernism*", *"Interpreting Music"*) and also added the publisher's name to monographs with three words in their titles to reduce the number of false matches from the searches (see below). We also converted the publisher names mentioned in the Scopus or BKCI outputs to variants that give the maximum recall (e.g., "Springer" and "Elsevier" instead of "Springer New York" and "Elsevier Ltd.").

*Gutmann "Identity in democracy" "Princeton University" "course description"***|***Gutmann "Identity in democracy" "Princeton University" syllabus*

We sometimes did not include publisher names and never included the last names of the second and subsequent authors because our initial manual checking showed that some instructors give abbreviated information about books (e.g., the first author and title) and so adding extra information to the queries would have considerably reduced the number of correct matches.

We selected the terms 'syllabus' and 'course description' to identify relevant online syllabi based on manual checks of search results during the study and general recommendations about the necessary structural components of course syllabi in higher education (O'Brien, Millis, & Cohen, 2008; Afros & Schryer, 2009). A PhD study of syllabus components in undergraduate courses, for instance, found that the majority (about 98%, n=180) of the course syllabi reviewed across science, social sciences and science departments in Taiwanese universities included a "*course description*" subsection (Tung, 2010). Although our initial manual checks found many other terms that could be used to identify online academic course syllabi (e.g., *Course Syllabus, Outline Syllabus, Syllabus Schedule, Syllabus Plan, Syllabus Information)*, it was impractical to cover all these terms in the automatic searches because each search term after the first doubles the number of Bing API searches without necessarily retrieving many new matches. Nevertheless, we also used the more general term '*syllabus*' because it covers a range of relevant descriptors (e.g., *Syllabus Content, Syllabus Title*) and hence is relatively powerful.

Stage 2: Limiting the Results to Academic Sites
In order to assess the academic teaching influence of monographs, we restricted the Bing API search results to a list of over 24,600 university websites from the *Webometrics Ranking of World Universities* (http://www.webometrics.info). This appears to be the most current and comprehensive list of websites of universities and academic institutions. In addition, we also included any search results from second-level academic domains for the U.S. (.edu/), Australia (.edu.au/), Malaysia (.edu.my/), Taiwan (.edu.tw/), Poland (.edu.pl/), and Singapore (.edu.sg/). We also included any search results from ".ac." academic domains for 22 countries (e.g., .ac.uk, .ac.nz, .ac.jp, ac.ir, ac.za). For some countries we included both types of academic domains (.edu.cn/ and .ac.cn/ for China).

We excluded all results from *Academia* (academia.edu) because of many false matches from the online CVs and publication lists of academics. We also removed results from several digital libraries with academic domains, such as *CiteSeerX* (*citeseerx.ist.psu.edu*), and online library catalogues (e.g., *digilib.ui.ac.id* or *lib.hcu.ac.th*) due to many false results from book lists and catalogue pages.

Stage 3: Identifying Correct Syllabus Matches
Because the results included some false matches after Stage 2, additional rules were added to identify correct academic syllabus matches. Based on manual checks, a set of rules was defined for the titles,

descriptions and URLs of the syllabus search results from academic websites. The results of Stage 2 were kept only if they matched Stage 3a, Stage 3b or Stage 3c.

Stage 3a: Search results title matches: The title field in the Bing API results was used to restrict results to selected terms that had a high probability of indicating that the webpage was a syllabus. For these terms, wildcard matching (*) was used to cover multiple term endings. The terms checked for included *syllabus, syllabi, module, course description*, course objective*, course outline, course detail*, course syllabus, course schedule, course information, course title, course overview, course structure, course content, teaching module, reading list, selected reading*, required reading*, recommended reading** and *additional reading**. There were also many portmanteau terms, such as c*oursesyllabus*, *coursetitle, teachingmodule*, and different ways to combine related terms, such as '*Content of Course*', '*Description of Course*' or '*Schedule of Class*' and these were also added to the rules. We also found many week-related syllabus results, such as '*History of philosophy course: week 1' or 'WEEK SEVEN: Introduction to biology*", and hence also included the combination of '*week*' and numbers (e.g., 1 or one, up to 14).

Stage 3b: Search results description matches: Many syllabi had just course names as their titles, such as '*Modern Imperialism*', '*Clinical Binocular Vision*' or '*Introduction to Quantum Mechanics*' but no syllabus-related terms*.* Nevertheless, it was sometimes possible to identify syllabi based on the text in their description fields, such as '*Course Description:', 'Outline Syllabus' and 'Tentative Syllabus'*. Hence, we included additional rules, which were mostly case sensitive, to capture syllabus-related text within course syllabi. Some examples of the case sensitive descriptors selected are *Course description, Course syllabus, Module description, Syllabus summary, Syllabus overview, Outline of lecture, Class syllabus, Syllabus and reading*, Detailed syllabus, Syllabus and assignment*, Syllabus and synopsis, Syllabus and schedule, Syllabus plan, Recommended reading*, Required textbook*, Syllabus and course information, Outline of lecture** and *Course objectives*. We also included different case formats, such as '*Course Description*', '*Course description*', '*COURSE DESCRIPTION',* and covered both *'and'*, and *'&'* (e.g., '*Syllabus and readings*' and '*Syllabus & readings*').

Stage 3c: Search results URL matches: We added extra rules, similar to those above, to locate syllabus mentions in URLs because in some valid cases the search results titles and descriptions did not include any syllabus-related words. Examples of URLs with relevant syllabus mentions are shown below. In some cases we used wildcards (e.g., syll*).

> *http://www.ants.edu/**syllabi**/sp09_theo772s.pdf*
> *http://www.macalester.edu/~harris/**syllabus**/crf.html*
> *http://hompi.sogang.ac.kr/gsis/2011/contents/download/**syll**_06s_03.doc*
> *http://cstl-cla.semo.edu/bwgentry/ui336/**course_description**.htm*
> *http://www.crs.cuhk.edu.hk/download/**CourseOutline**/1213T2/CURE3143.pdf*
> *http://www.office.usp.ac.jp/~klinger.w/class/jcmu/**CourseReadings**.html*
> *https://secure.garrett.edu/ExWeb/**CourseMaterials**/F2010_21-683.pdf*

We also found additional descriptors based on different combinations of academic semesters and years in the URLs (e.g., *fall2000, spring-2009, winter_2011, summer.2013* and *2009winter*) and generated a list of descriptors for these.

> *http://spa.sdsu.edu/forms/CJ_302_**Fall_2011**_Kaplan.pdf*

http://www.d.umn.edu/~lknopp/geog3334-1/Spring%2006.htm
*http://home.uchicago.edu/~paul/StanilandPLSC36100**Winter12***.pdf
*http://www.uh.edu/~trdegreg/**summer2011***econ3355.htm
*http://memo.cgu.edu.tw/yu-yen/**2011fall**-wl.mht.htm

Stage 4: Excluding False Syllabus Matches

In the next step we defined a set of rules to exclude false matches based upon the titles, descriptions and URLs of the results from stage 3. This was again based on manual checks identifying common cases of false matches.

Stage 4a: Removing false matches with title checks: The incorrect matches after Stage 3 were mainly from academic CVs, profile pages, digital libraries, book databases and catalogues. For example, there were false matches from online CVs because some books were mentioned in the CVs of their authors (or reviewers) along with syllabus-related terms, such as educational background narratives (e.g., "Developed Course Syllabus", "Designed Course Syllabus", "Lectures and Course Structure" or "Course Module Taught"). In order to cope with this type of false match we automatically removed results with CV-related terms s in their title (e.g., *CV, C.V., resume, vitae, my profile, my publications*).

To remove false matches from digital libraries, book lists or databases we ignored any results with a range of related terms in their title, including *New Book, Electronic Resource, Book Review, Newsletter, Read Unlimited Book*, Reserve Book*, All Books, New Item*, New Bookmark*, New Title, Book Catalog*, Search Result*, News and Browse by* [Author, Title, Subject, Year]*.

Stage 4b: Removing false matches with URL checks: Additional rules were devised to remove matches that appeared to be CVs based upon the terms within the URLs of the filtered search results (e.g., URLs ending with *.cv.pdf, .vitae.pdf, .c.v.pdf, .cv.htm*, .cv.doc*, Curriculum_Vitae, Curriculum.Vitae, Curriculum-Vitae* or *CurriculumVitae*). There were also false matches from Excel files (.xls and .xlsx) from academic libraries listing a huge number of books and other publications for library acquisition services (e.g., http://www.lib.hcu.ac.th/BookFile/GRAD/MEG.xls) and so all results with Excel file extensions were removed. False matches within digital libraries (e.g., http://citeseerx.ist.psu.edu/showciting?cid=6447860) were removed by automatically removing all results from the digital library website (e.g., citeseerx.ist.psu.edu).

The filtered results from Stage 4 were matched with Scopus or BKCI citation counts (Table 1).

Table 1. Results at different stages of the automatic syllabus mention detection process.

Syllabus matches	Step 1: Raw Bing API search results	Step 2: Restriction to academic websites	Step 3: Restriction to syllabus descriptors	Step 4: CV and library false matches removed
No. (% with at least one syllabus) median (mean) maximum per book	133,468 (95%) 8 (11.1)376	38,334 (70%) 2 (3.2) 220	22,766 (41%) 0 (1.9)194	19,184 (38%) 0 (1.6) 158

Results

The Accuracy and Coverage of the Syllabus Search Method

Random Scopus Monographs Search Results Checks

To estimate the accuracy (precision) and coverage (recall) of the automatic syllabus search method, the first author conducted a manual check of a sample of 2,000 results (1,000 matching after stage 4b and 1,000 filtered out of the original Bing API results by stages 3 or 4) and a reference librarian independently did the same task for another sample of 600 results (300 matching and 300 filtered out). In both cases the checks were for whether the monographs were mentioned in the context of academic course reading lists or not. We did not cross-check the results with more people because identifying the correct or incorrect matches in course syllabi or other online documents (e.g., library catalogues, book lists or academic CVs) seemed to be very straightforward and not subjective.

The overall estimated recall and precision were 92.2% and 93.2% for the first author's manual checks and 91.3% and 95.6% for the independent checker (Table 2), suggesting that the automatic method gives high accuracy and coverage (over 90%). Note, however, that this check excludes all results not found by the original Bing API searches, such as password protected syllabi, syllabi in pages not found or not reported by Bing (for search engine coverage issues, see: Bar-Ilan & Peritz, 2004), and syllabi containing books mentioned in unusual ways.

Table 2. Estimates of the accuracy and coverage (within the Bing API results) of the automatic syllabus mention searches.

Manual checking experiments	Sampled records	Estimated precision	Estimated recall	Relevant results	False matches	Missing relevant syllabus results
Author tests	1,000	**93.2%**	**92.2%**	932	68 (6.8%)	78 (7.8%)
Independent tests	300	**95.6%**	**91.3%**	287	13 (4.3)	26 (8.6%)

Most of the false matches in the final sample were from online CVs because the rules were insufficient to exclude all online CVs. For instance, the book "*Fragmented Intimacy: Addiction in a Social World*" by Peter J. Adams was in his profile page (https://www.fmhs.auckland.ac.nz/en/soph/peter-adams/gambling-research.html), which contains biographic information about course development. Moreover, some online CVs were stored in databases and had no CV-relevant terms in their URLs or titles. A few false matches also derived from mentions of books in book reviews, essays, and book lists in academic websites.

As shown in Table 2, about 8% of the relevant syllabus mentions in the Bing API results were missed. A common reason was absence of syllabus-related terms in the titles, descriptions and URLs of the search results (e.g., <u>Title</u>: '*ANTH-UA 104 (South Asia) 1 - New York University*'; <u>Description</u>: '*2007 Enemy Lines: Childhood, Warfare, and Play in Batticaloa. ...*'; <u>URL</u>: '*http://anthropology.as.nyu.edu/docs/IO/17923/ANTHUA104Teja.pdf*'. In other cases there were unusual syllabus-related terms that were not included in the algorithm (e.g., *http://www.msrchm.edu/common/doc/**detailedCurriculum**.doc*). In addition, some non-English language course syllabi were also overlooked (e.g., *Corso di studio, descripción del curso,* and *Kurslitteratur*).

Top Search Results Checks
Extra manual checks were performed for the four most mentioned books from Scopus and BKCI because extreme cases seem particularly likely to have false matches and these false matches may be influential if the results are used for research evaluations.

-   Sugrue, T.J. (2010). The origins of the urban crisis: Race and inequality in postwar Detroit. Princeton University Press. *(Scopus, Arts and Humanities highly syllabus mentioned).*
-   Angrist, J.D., Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. *(Scopus, Social Science highly mentioned).*
-   De Berg, M. et al. (2008). Computational geometry: Algorithms and applications. Springer. *(Scopus, Science highly syllabus mentioned).*
-   Bartels, LM. (2008). Unequal Democracy: The Political Economy of the New Gilded Age. Princeton University Press. *(BKCI highly syllabus mentioned)*

The manual checks by the first author found high recall (91%-93%) and precision (93%-96%) for all books.

Total syllabus mentions vs. total citation counts
The number of Scopus citations (186,764) to books is substantially (973%) higher than the number of syllabus mentions (19,184) of them (Table 3). About 55% of the monographs had at least one Scopus citation and 38% had at least one academic syllabus mention. Unsurprisingly, the distribution of the Scopus citations is highly skewed because of many highly cited monographs. In the arts and humanities, however, the Scopus citation medians and the academic syllabus mention medians are the same (1) and 56% of books had at least one syllabus mention, probably reflecting the importance of monographs for teaching in book-based subjects (e.g., law, history, literature, philosophy). In contrast, the chemistry Scopus citation median (18) is much higher than the syllabus mention median (0) and 90% of books had at least one citation (vs. 27% with at least one syllabus mention), probably reflecting that Scopus-indexed chemistry monographs were mostly used for research rather than teaching. Presumably Scopus does not index chemistry textbooks, for example. In the social sciences about 52% of the monographs had at least one syllabus mention and 67% had at least one Scopus citation, perhaps reflecting the use of the same monographs for both research and academic teaching in this area.

The relationship between syllabus mentions and Scopus citations
Table 3 shows that there are significant, but low, correlations between Scopus citations and academic syllabus mentions for all monographs in the study (r=0.271, p = 0.01, n=19,184). The correlations are higher in the social sciences (.522), computer science (.325) and arts and humanities (.300) and much lower in chemistry (.110), physics and astronomy (.124), and engineering and material sciences (.200). The mostly low but significant correlations suggest that the academic teaching influence and research impact of monographs are only loosely related in most areas of scholarship. This is an important result but is not surprising because most disciplines probably rely upon educational textbooks for teaching, although monographs may be more used in teaching in the social sciences, arts and humanities. The stronger correlation in the social sciences suggests that more cited monographs tend to be recommended more by instructors. The weaker correlation in the arts and humanities, in contrast, could be due to the relative unimportance of citations from journal articles in these disciplines, however, rather than the lack of a relationship between teaching and research impact.

Table 3. Spearman correlations between academic syllabus mentions and citations to Scopus books (2005-2010) by field.

Discipline	No. of books	Syllabus mentions No. (% with syllabus) median (mean) max	Scopus citations No. (% with cites) median (mean) max	Correlation: Syllabus mentions vs. Scopus citations
Arts and humanities	1,892	4,076 (56%) 1 (2.1) 97	16,383 (59%) 1 (8.6) 1,142	.300**
Social sciences	1,215	4,005 (52%) 1 (3.2) 154	19,566 (67%) 3 (16.1) 966	.522**
Economics and Finance	1,467	2,270 (33%) 0 (1.5) 158	12,613 (48%) 0 (8.5) 855	.263**
Psychology	299	423 (49%) 0 (1.4) 20	2,308 (57%) 1 (7.7) 161	.217**
Environmental and earth sciences	1,352	1,422 (34%) 0 (1) 35	15,293 (59%) 2 (11.3) 594	.245**
Physics and astronomy	398	426 (23%) 0 (1) 37	2,848 (28%) 0 (7.1) 135	.124*
Chemistry	603	587 (27%) 0 (0.9) 88	40,434 (90%) 18 (67) 4,968	.110*
Engineering and material sciences	1,852	2,583 (38%) 0 (1.3) 79	43,876 (62%) 4 (23.6) 773	.200**
Computer science	810	1,485 (37%) 0 (1.8) 101	8,517 (38%) 0 (10.5) 328	.325**
Mathematics	348	930 (42.5%) 0 (2.6) 89	17,458 (44%) 0 (50.1) 5,400	.236**
Medical Sciences	1,734	977 (24%) 0 (0.5) 19	7,468 (43%) 0 (4.3) 929	.274**
Total	11,970	19,184 (38%) 0 (1.6) 158	186,764 (55%) 1 (15.6) 5,400	.271**

* Significant at the p = 0.05 level. ** Significant at the p = 0.01 level.

To check whether the results differ by year, we calculated separate annual correlations between syllabus mentions and Scopus citations. There were differences between years, although they were not huge (ranging from 0.221 to 0.317) and fluctuated from year to year rather than systematically increasing or decreasing. We have not found an obvious cause for the fluctuations and speculate that they may be due to annual changes in the coverage of books by Scopus.

In order to check the extent to which the Scopus selection of monographs influences the results, we conducted a secondary analysis of 2,300 BKCI-indexed monographs published in 2008. There are low but significant correlations between academic syllabus mentions and BKCI citations in the three broad fields analysed (Table 4). The correlation is higher in the arts and humanities and the social sciences than in science and medicine. Thus, it seems that the Scopus and BKCI selections of monographs have had an impact on the correlation results.

Table 4. Secondary analysis of BKCI indexed monographs in 2008 by field.

Discipline	No. of books	Syllab. mentions No. (% with syllabus) median (mean) max	Scopus citations No. (% with cites) median (mean) max	Correlations: Syllabus mentions and BKCI citations
Arts and humanities (e.g.,	1,078	1,412 (37%)	11,414 (84%)	.385**

History, Religion, Law, Literature and Philosophy)		0 (1.3) 47	5 (10), 196	
Social Sciences (e.g., Social Sciences, Psychology, Sociology and Education)	685	929 (35%) 0 (1.4) 78	7,596 (84%) 5 (11) 272	.365**
Science and medicine (e.g., Engineering, Life Sciences, Computer Science, General & Internal Medicine)	530	717 (32%) 0 (1.3) 33	6,627 (51.6%) 1 (12.5) 339	.107*
Total	2,293	3,058 (35%) 0 (1.3) 78	25,637 (77%) 4 (11.1) 339	.303**

* Significant at the p = 0.05 level. ** Significant at the p = 0.01 level.

Discussion

The automatic syllabus citation extraction method described above was shown by the manual checks to have accuracy and coverage over 90% (at least within the original Bing API results), both overall and for the most mentioned monographs. Hence, the method seems likely to be effective enough at identifying public online mentions of monographs in academic syllabi to be useful in practice. Its main practical limitation is its overall coverage because its results are presumably a very small fraction of the number of syllabus mentions (both public and private, online and offline) of the monographs investigated. Nevertheless, the figures may be reasonable for comparative purposes to help assess the teaching influence of monographs, especially in teaching-oriented and book-based fields and when evaluators, promotion committees and academic departments need to be informed about teaching influence of books rather than just their research impact. For instance, about 16.5% of the submissions to the 2008 UK RAE were books, but a third of all submissions in the social sciences and humanities were books compared with only 1% in the sciences (Kousha, Thelwall, & Rezaie, 2011). Perhaps some of the submitted books had teaching impact rather than, or in addition to, research impact.

Arts and humanities monographs were more commonly mentioned in academic course syllabi than were science monographs, confirming that there are disciplinary differences in the extent to which monographs are used in education. In further support of this, in the arts and humanities 18% of monographs had at least one syllabus mention but no Scopus citations, in which is larger than in the social sciences (12%) and science and medicine (11%). Most importantly, this shows that some monographs have a measurable educational impact even though they have no citation impact (see also tables 6, 7 and 8 in the Appendix). However, citations to monographs may be underrepresented in Scopus due to its low coverage of books and monographs.

Table 5. The proportion of monographs with at least one syllabus mention and no Scopus citations.

Arts and humanities (n=1,892)		Social science (n=2,981)		Science and medicine (n=7,097)		All monographs (n=11,970)	
At least one syllabus mention but no citations	At least one citation but no syllabus mentions	At least one syllabus mention but no citation	At least one citation but no syllabus mentions	At least one syllabus mention but no citations	At least one citation but no syllabus mentions	At least one syllabus mention but no citations	At least one citation but no syllabus mentions
18%(341)	21.5%(407)	11.8%(352)	26%(775)	10.9%(773)	32.3%(2290)	12.2%(1466)	29%(3472)

About half of the syllabus mentions of Scopus monographs in this study were from U.S.-affiliated institutions, followed by the UK (7.5%), Canada (5.8%) and India (5.7%) (Table 9 in the Appendix). Surprisingly, there were only 132 (0.7%) syllabus mentions from Australian universities (.edu.au/) compared with 401 (2.1%) from Polish Universities, suggesting that policies about publishing academic syllabi online could influence the results. Moreover, the method used here targets English language academic syllabi and could presumably be extended to capture non-English syllabi.

The relatively low correlations between syllabus mentions and citations counts for academic monographs, especially in the sciences, suggests that there are distinct ways in which teaching and research are supported by books and there seems to be little relationship between teaching and research (for a review see, Hattie, & Marsh, 1996), at least for monographs. For instance, some scientific books may have little value for undergraduate teaching because they are primarily written for experts and in hierarchical areas of science few undergraduates may be capable of understanding them. However, the medium correlation in the social sciences suggests that some monographs may be useful for both teaching and research. For instance, the 20 social sciences monographs that were most mentioned in academic syllabi also received many citations in Scopus (see Table 7 in the Appendix). This supports the argument that "stronger relationships between research and teaching exist in so-called 'soft' disciplines, such as humanities and social sciences, than in 'hard' ones, such as natural sciences" (Feldman, 1987 as quoted by Griffiths, 2004, p. 711).

Limitations
The automatic syllabus search method described in this study has a number of limitations.

Results beyond university websites: Many academic syllabi are only accessible in the open web and they are not searchable by commercial search engines because they are stored in password protected databases or websites. Moreover, there could also be relevant results from scientific societies, organisations and research institutions outside university websites with academic teaching programs. For instance, there were some relevant excluded syllabus mentions from *The American Psychological Association (APA)* (http://www.apa.org/divisions/div43/Syllabi/6600-01SooHooFamTherapySyll1.doc) and *The American Political Science Association (APSA)* (http://www.apsanet.org/Files/Syllabi/currahundergrad.doc) both using non-academic domains. Moreover, results from social networks and blogs were excluded because they contain many irrelevant pages. However, some syllabus mentions from non-academic websites such as *wordpress.com* (e.g., https://genderandpolitics.wordpress.com/syllabus-2/) or *slideshare.net* (e.g., http://www.slideshare.net/MarkWhit/statebuilding-syllabus) are relevant, perhaps authored by creative professors that do not want to be restricted by their academic web spaces. Hence, the results of this study are only a subset of the online university extracted course syllabi indexed by Bing.

Results from non-English Academic Syllabi: The automatic syllabus mention search method is likely to miss many non-English academic syllabus mentions because the keywords used in the search query and filtering search terms were largely in English (although *syllabus* and *curriculum vitae* are Latin and used internationally), missing mentions of monographs in academic course syllabus in Spanish (e.g., *'programa de estudios'*), Italian (e.g., *'Programma del corso'*) or German (e.g., '*Kursplan'* or *'Kursprogramm'*), French (e.g., '*Descriptif du cours'*) and other languages.

Other issues: Lecturers may recommend the same teaching resources in course syllabi from different years with slightly different URLs (e.g., .../*syllabus/**Fall2012**/* and .../*syllabus/**Fall2014**/*). We did not exclude such cases, although it would presumably have been possible to develop a new set of heuristics

to do so. Nevertheless, updated syllabi may contain the modified reading lists and so multiple occurrences reflect the positive indicator of continued use over time. Moreover, for practical reasons, we only used the two descriptors 'syllabus' and 'course description' in the queries. However, there are other relevant terms that we did not use in the queries (e.g., *Course Outline* or *Course Module*) to avoid having to submit too many Bing API searches. In addition, we used heuristics to filter the syllabus mentions. Although the manual checks estimated that this method was accurate enough to estimate the teaching influence of books (over 90%), there may be individual books for which it is much less accurate. Finally, we excluded monographs (4%) with very short titles (single or two words) from our main data set to prevent retrieving too many incorrect matches. This is a practical limitation because any research assessment using syllabus mentions would have to evaluate all monographs, irrespective of title length. For such books, searches may need to be manually constructed on a case by case basis, perhaps including additional bibliographic information, such as the last names of the three first authors and the publisher name.

Conclusions

The ability to assess the teaching impact of publications should be useful for evaluators, funders and promotion committees when educational impact is valued rather than, or in addition to, research impact. For example, promotion committees might consider an uncited monograph to be a failure unless they had the time to read it or could find out that it had been mentioned in many academic syllabi. The method described here can help to estimate the educational impact of monographs inside academia based on their mentions in online academic syllabi. It may be more useful than using book sales alone because these could reflect an unknown combination of educational, research or practitioner uses.

In answer to the first research question, mentions of academic monographs in online academic course syllabi can be automatically retrieved with high precision and with high coverage relative to the initial set of Bing API results.

In answer to the second research question, there were low but significant correlations between the number of academic syllabus mentions and citation counts for monographs. The correlations ranged from 0.522 in the social sciences to 0.110 in chemistry for Scopus and from 0.385 in the arts and humanities to 0.107 in science and medicine for the BKCI data set. The overall results, combined with previous evidence of stronger relationships between research and teaching in the humanities and social sciences than in the hard and natural sciences (Griffiths, 2004), suggest that the academic syllabus mention searches could be particularly helpful to assess the teaching influence of books and monographs in the social sciences, arts and humanities.

In the future, or as a future project, if more universities make their course syllabi available online or systematically register them in a public database then it would be much easier to harvest academic course reading lists automatically with a high degree of accuracy. This would also pave the way to create an educational index of the teaching influence or impact of textbooks or monographs in different subjects and at different educational levels, which would provide additional insights into the relationship between research and teaching.

Future research might extend the method to other languages and include the websites of a range of scientific organizations, societies and research centres outside universities. Future investigations could also compare the educational impacts of textbooks with other books. Although it is sometimes difficult to distinguish between teaching books mainly written for students and other types of books on the basis

of their titles, some publishers have attempted to do so, including Cambridge University Press (http://www.cambridge.org/gb/academic/textbooks/) and Oxford University Press (http://ukcatalogue.oup.com/category/academic.do, see the 'Undergraduate Textbook' subcategory under each subject).

References

Afros, E., & Schryer, C. F. (2009). The genre of syllabus in higher education. Journal of English for Academic Purposes, 8(3), 224-233.

Albers, C. (2003). Using the syllabus to document the scholarship of teaching. Teaching Sociology, 31(1), 60-72.

Altbach, P. & Lewis, L. (1997). The academic profession in international perspective. In PG Altbach (ed.), The international academic profession, The Carnegie Foundation for the Advancement of Teaching, New Jersey, pp. 3–48.

Archambault, É., 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.

Bar-Ilan, J., & Peritz, B. C. (2004). Evolution, continuity, and disappearance of documents on a specific topic on the Web: A longitudinal study of 'informetrics'. Journal of the American Society for Information Science and Technology, 55(11), 980 - 990.

Bar-Ilan, J. (2010). Citations to the "Introduction to informetrics" indexed by WOS, Scopus and Google Scholar. Scientometrics, 82(3), 495-506.

Book Citation Index (2014). http://wokinfo.com/products_tools/multidisciplinary/bookcitationindex/

Boshier, R. (2009). Why is the scholarship of teaching and learning such a hard sell? Higher Education Research and Development, 28(1), 1-15.

Boyer, E. L. (1990). Scholarship revisited. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching. http://184.168.109.199:8080/jspui/bitstream/123456789/2134/1/ED326149.pdf

Butler, L., & Visser, M. (2006). Extending citation analysis to non-source items. Scientometrics, 66(2), 327–343.

Cabezas-Clavijo, A., Robinson-García, N., Torres-Salinas, D., Jiménez-Contreras, E., Mikulka, T., Gumpenberger, C., Wemisch, A., & Gorraiz, J. (2013). Most borrowed is most cited? Library loan statistics as a proxy for monograph selection in citation indexes. In: Proceedings of 14th International Conference of the International Society for Scientometrics and Informetrics (pp. 1237-1252). Vienna, Austria: AIT Austrian Institute of Technology GmbH Vienna.

Chalmers, D. (2011). Progress and challenges to the recognition and reward of the scholarship of teaching in higher education. Higher Education Research and Development, 30(1), 25-38.

Clump, M. A., Bauer, H., & Bradley, C. (2004). The extent to which psychology students read textbooks: A multiple class analysis of reading across the psychology curriculum. Journal of Instructional Psychology, 31(3), 227-232.

Crittenden, V. L., & Wilson, E. J. (2006). Content, pedagogy, and learning outcomes in the international marketing course. Journal of Teaching in International Business, 17(1-2), 81-101.

Cullars, J. (1998). Citation characteristics of English-language monographs in philosophy. Library & Information Science Research, 20(1), 41–68.

Dilevko, J., McMillan, B., Allison-Cassin, S., Aspinall, J., & Mauro, C. (2006). Investigating the value of scholarly book reviews for the work of academic reference librarians. Journal of Academic Librarianship, 32(5), 452-466.

Donovan, C., & Butler, L. (2007). Testing novel quantitative indicators of research "quality," esteem and "user engagement:" An economics pilot study. Research Evaluation, 16(4), 231-242.

Eberly, M. B., Newton, S. E., & Wiggins, R. A. (2001). The syllabus as a tool for student-centered learning. The Journal of General Education, 50(1), 56–74.

Feldman, K. (1987). Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: a review and exploration. Research in Higher Education, 26(3), 227-298.

Fincher, R. -. E., Simpson, D. E., Mennin, S. P., Rosenfeld, G. C., Rothman, A., McGrew, M. C., Penelope A. Hansen, P.A, Mazmanian, P. E., & Turnbull, J. M. (2000). Scholarship in teaching: An imperative for the 21st century. Academic Medicine, 75(9), 887-894.

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.

Garfield, E. (1996). Citation indexes for retrieval and research evaluation. Consensus Conference on the Theory and Practice of Research Assessment, Capri.

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.

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.

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.

Griffiths, R. (2004). Knowledge production and the research-teaching nexus: the case of the built environment disciplines. Studies in Higher Education, 29(6), 709-726.

Gurung, R. A. A., Landrum, R. E., & Daniel, D. B. (2012). Textbook use and learning: A North American perspective. Psychology Learning and Teaching, 11(1), 87-98.

Gurung, R. A. R., & Martin, R. C. (2011). Predicting textbook reading: The textbook assessment and usage scale. Teaching of Psychology, 38(1), 22-28.

Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. Scientometrics. 101(2), 1419-1430.

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.

Hattie, J., & Marsh, H. W. (1996). The relationship between research and teaching: A meta-analysis. Review of Educational Research, 66(4), 507-542. http://wokinfo.com/products_tools/multidisciplinary/bookcitationindex/

Healey, M. (2000). Developing the scholarship of teaching in higher education: A discipline- based approach. Higher Education Research & Development, 19(2), 169-189.

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. Scientometrics, 44(2), 193–215.

Homa, N., Hackathorn, J., Brown, C. M., Garczynski, A., Solomon, E. D., Tennial, R., Sanborn, U. A. &, Gurung, R. A. R. (2013). An analysis of learning objectives and content coverage in introductory psychology syllabi. Teaching of Psychology, 40(3), 169-174.

Huang, M., & Chang, Y. (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.

Jenkins, A. (1995). The research assessment exercise, funding and teaching quality. Quality Assurance in Education, 3(2), 4-12.

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.

Kousha, K., & Thelwall, M. (2009). Google Book Search: Citation analysis for social science and the humanities. Journal of the American Society for Information Science and Technology, 60(8), 1537-1549.

Kousha, K., & Thelwall, M. (2014). An automatic method for extracting citations from Google Books. Journal of the Association for Information Science and Technology. doi: 10.1002/asi.23170. http://www.scit.wlv.ac.uk/~cm1993/papers/AutomaticGoogleBooksCitationsPreprint.pdf

Kousha, K., & Thelwall, M. (in press). Can Amazon.com reviews help to assess the wider impacts of books? Journal of the Association for Information Science and Technology. http://www.koosha.tripod.com/AmazonReviewstoAssessBooks-Preprint.pdf

Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus. Journal of the American Society for Information Science and Technology, 62(11), 2147-2164.

Krampen, G., Becker, R., Wahner, U., & Montada, L. (2007). On the validity of citation counting in science evaluation: Content analyses of references and citations in psychological publications. Scientometrics, 71(2), 191–202.

Kreber, C. (2002). Controversy and consensus on the scholarship of teaching. Studies in Higher Education, 27(2), 151-167.

Kreber, C., & Cranton, P. A. (2000). Exploring the scholarship of teaching. Journal of Higher Education, 71(4), 476-495.

Landrum, R. E., & Hormel, L. (2002). Textbook selection: Balance between the pedagogy, the publisher, and the student. Teaching of Psychology, 29(3), 245-248.

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.

Li, X., & Thelwall, M. (2012). F1000, Mendeley and traditional bibliometric indicators. In: Proceedings of the 17th International Conference on Science and Technology Indicators. Montréal, Canada. pp. 451–551.

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. Scientometrics, 91(2), 461–471.

Linmans, A. J. M. (2010). Why with bibliometrics the humanities does not need to be the weakest link. Indicators for research evaluation based on citations, library bindings and productivity measures. Scientometrics, 83(2), 337–354.

Mishra, J., Day, K., Littles, D., & Vandewalker, E. (2011). A content analysis of introductory courses in music education at NASM-accredited colleges and universities. Bulletin of the Council for Research in Music Education, (190), 7-19.

Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. Journal of the Association for Information Science and Technology, 65(8), 1627-1638.

Mohammadi, E., Thelwall, M. & Kousha, K. (in press). Can Mendeley bookmarks reflect readership? A survey of user motivations. Journal of the Association for Information Science and Technology. http://www.scit.wlv.ac.uk/~cm1993/papers/CanMendeleyBookmarksReflectReadershipSurvey_preprint.pdf

Nederhof, A. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. Scientometrics, 66(1), 81–100.

O'Brien, J. G., Millis, B. J., & Cohen, M. W. (2008). The course syllabus: A learning-centered approach. 2nd edition San Francisco: Calif.

Parkes, J., & Harris, B. (2002). The purposes of a syllabus. College Teaching, 50(2), 55-61.

Pieterse, A. L., Evans, S. A., Risner-Butner, A., Collins, N. M., & Mason, L. B. (2009). Multicultural competence and social justice training in counseling psychology and counselor education: A review and analysis of a sample of multicultural course syllabi. Counseling Psychologist, 37(1), 93-115.

Priem, J., & Hemminger, B. M. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social web. First Monday, 15(7). http://firstmonday.org/ojs/index.php/fm/article/view/2874/2570

Priem, J., Piwowar, H., & Hemminger, B. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. Retrieved from http://arXiv.org/html/1203.4745v1

Ramsden, P., & Martin, E. (1996). Recognition of good university teaching: Policies from an Australian study. Studies in Higher Education, 21(3), 299-315.

Scopus Content: Book Expansion project update (2014). http://blog.scopus.com/posts/scopus-content-book-expansion-project-update

Sharma, A., Van Hoof, B., & Pursel, B. (2013). An assessment of reading compliance decisions among undergraduate students. Journal of the Scholarship of Teaching and Learning, 13(4), 103 – 123.

Shaw, D. (1991). An analysis of the relationship between book reviews and fiction holdings in OCLC. Library and Information Science Research, 13(2), 147-154.

Sikorski, J. F., Rich, K., Saville, B. K., Buskist, W., Drogan, O., & Davis, S. F. (2002). Student use of introductory texts: Comparative survey findings from two universities. Teaching of Psychology, 29(4), 312-313.

Skelton, A. (2004) Understanding 'teaching excellence' in higher education: a critical evaluation of the National Teaching Fellowships Scheme. Studies in Higher Education, 29(4), 451–466.

Smith, M. F. & Razzouk, N. Y. (1993). Improving classroom communication: The case of the course syllabus. Journal of Education for Business, 68(4), 215-221

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, 49(4), 364-374.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. Review of Educational Research, 83(4), 598-642.

Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. PLOS ONE, 8(5), e64841.

Thelwall, M. & Wilson, P. (in press). Mendeley readership Altmetrics for medical articles: An analysis of 45 fields, Journal of the Association for Information Science and Technology. http://www.scit.wlv.ac.uk/~cm1993/papers/MendeleyInScienceAltmetricsPreprint.pdf

Thompson, B. (2007). The Syllabus as a communication document: Constructing and presenting the syllabus. Communication Education, 56(1), 54-71.

Torres-Salinas, D., & 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, 3(1), 9-26.

Torres-Salinas, D., Robinson-García, N., Campanario, J. M., & López-Cózar, E. D. (2014). Coverage, field specialisation and the impact of scientific publishers indexed in the book citation index. Online Information Review, 38(1), 24-42.

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), 615-620.

Tung, Y.T. (2010). A case study of undergraduate course syllabi in Taiwan. PhD. Dissertation, University of North Texas. http://digital.library.unt.edu/ark:/67531/metadc28487/m2/1/high_res_d/dissertation.pdf

Weiten, W. (1988). Objective features of introductory psychology textbooks as related to professors' impressions. Teaching of Psychology, 15(1), 10-16.

Weller, A. C. (2001). Editorial peer review: Its strengths and weaknesses. Medford, N.J: Information Today.

White, H.D., Boell, S.K., Yu, H., Davis, M., Wilson, C.S., & Cole, F.T. (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.

Young, P. (2006) Out of balance: lecturers' perception of differential status and rewards in relation to teaching and research. Teaching in Higher Education, 11(2), 191–202.

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. Scientometrics, DOI 10.1007/s11192-014-1264-0

Zuccala, A., & Guns, R. (2013). Comparing book citations in humanities journals to library holdings: Scholarly use versus "perceived cultural benefit" (RIP). In: Proceedings of 14th International Conference of the International Society for Scientometrics and Informetrics (pp.353–360). Vienna, Austria: AIT Austrian Institute of Technology GmbH Vienna.

Zuccala, A., & van Leeuwen, T. (2011). Book reviews in humanities research evaluations. Journal of the American Society for Information Science and Technology, 62(10), 1979-1991.

Zuccala, A., Guns, R., Cornacchia, R., & Bod, R. (in press, 2014). Can we rank scholarly book publishers? A bibliometric experiment with the field of history. Journal of the Association for Information Science and Technology. http://www.illc.uva.nl/evaluating-humanities/RankingPublishers(Preprint_2014).pdf

Zuccala, A., Guns, R., Cornacchia, R., & Bod, R. (in press, 2014). Can we rank scholarly book publishers? A bibliometric experiment with the field of history. Journal of the Association for Information Science and Technology. Retrieved from http://www.illc.uva.nl/evaluatinghumanities/RankPublishers(Mar5_2014Preprint).pdf

Appendix

Table 6. The 20 Arts and Humanities books that were mentioned most often in academic syllabi or course reading lists.

Scopus Book	Scopus subject	Academic syllabi	Scopus citations
Sugrue T.J. The origins of the urban crisis: Race and inequality in postwar Detroit. Princeton University Press, 2010.	Arts and Human.	97	426
Chan W.-T. A source book in Chinese philosophy. Princeton University Press, 2008.	Arts and Human.	55	0
Schreibman S., SUnsworth J. A Companion to Digital Humanities. John Wiley and Sons, 2007.	Arts and Human.	48	2
Paret P., Craig G.A., Gilbert F. Makers of modern strategy from Machiavelli to the nuclear age. Princeton University Press, 2010.	Arts and Human.	46	0
Cooper F. Colonialism in question: Theory, knowledge, history. University of California Press, 2005.	Arts and Human.	45	279
Payne C.M. I've got the light of freedom: The organizing tradition and the Mississippi freedom struggle. University of California Press, 2007.	Arts and Human.	45	22
Fredrickson G.M. Racism: A short history. Princeton University Press, 2009.	Arts and Human.	37	175
Hockings P. Principles of visual anthropology. De Gruyter Mouton, 2009.	Arts and Human.	32	1

Scopus Book	Scopus subject	Academic syllabi	Scopus citations
Trachtenberg M. The craft of international history: A guide to method. Princeton University Press, 2009.	Arts and Human.	29	29
Macdonald S. A Companion to Museum Studies. John Wiley and Sons, 2007.	Arts and Human.	28	2
Duranti A. A Companion to Linguistic Anthropology. Blackwell Publishing Ltd, 2007.	Arts and Human.	27	8
Kachru B.B., Kachru Y., Nelson C.L. The Handbook of World Englishes. John Wiley and Sons, 2008.	Arts and Human.	26	6
Bowen J.R. Why the French don't like headscarves: Islam, the state, and public space. Princeton University Press, 2010.	Arts and Human.	26	129
Bhatia T.K., Ritchie W.C. The Handbook of Bilingualism. John Wiley and Sons, 2008.	Arts and Human.	21	2
Naremore J. More than night: Film noir in its contexts. University of California Press, 2008.	Arts and Human.	20	43
Farber D. Taken hostage: The Iran hostage crisis and America's first encounter with radical Islam. Princeton University Press, 2009.	Arts and Human.	19	6
Van Dyke R.M., Alcock S.E. Archaeologies of Memory. Blackwell Publishers Ltd, 2008.	Arts and Human.	18	8
Erskine A. A Companion to the Hellenistic World. Blackwell Publishing Ltd, 2007.	Arts and Human.	18	1
Kennedy G.A. A new history of classical rhetoric. Princeton University Press, 2009.	Arts and Human.	18	0
Pickering M. Research methods for cultural studies. Edinburgh University Press, 2008.	Arts and Human.	17	1

Table 7. The 20 Social Sciences, and Economics books that were mentioned most often in academic syllabi or course reading lists.

Scopus Book	Scopus subject	Academic syllabi	Scopus citations
Angrist J.D., Pischke J.-S. Mostly harmless econometrics: An empiricist's companion. Princeton University Press, 2008.	Soc. Sci. Economics	158	640
Acemoglu D. Introduction to modern economic growth. Princeton University Press, 2008.	Soc. Sci. Economics	154	172
Keohane R.O. After hegemony: Cooperation and discord in the world political. Princeton University Press, 2005.	Soc. Sci. Economics	154	966
Gali J. Monetary policy, inflation, and the business cycle: An introduction to the new Keynesian framework. Princeton University Press, 2009.	Soc. Sci. Economics	111	109
Tirole J. The theory of corporate finance. Princeton University Press, 2010.	Soc. Sci. Economics	102	229
Slaughter A.-M. A new world order. Princeton University Press, 2009.	Soc. Sci. Economics	85	935
Reinhart C.M., Rogoff K.S. This time is different: Eight centuries of financial folly. Princeton University Press, 2009.	Soc. Sci. Economics	79	960
Smelser N.J., Swedberg R. The handbook of economic sociology. Princeton University Press, 2010.	Soc. Sci. Economics	78	179
Ikenberry G.J. After victory: Institutions, strategic restraint, and the rebuilding of order after major wars. Princeton University Press, 2009.	Soc. Sci. Economics	77	570
Rodrik D. One economics, many recipes: Globalization, institutions, and economic growth. Princeton University Press, 2008.	Soc. Sci. Economics	61	213
Portes A., Rumbaut R.G. Immigrant America: A portrait. University of California Press,	Soc. Sci.	57	26

Scopus Book	Scopus subject	Academic syllabi	Scopus citations
2006.	Economics		
Clark G. A farewell to alms: A brief economic history of the world. Princeton University Press, 2008.	Soc. Sci. Economics	56	228
Snow D.A., Soule S.A., Kriesi H. The Blackwell Companion to Social Movements. Blackwell, 2007.	Soc. Sci. Economics	51	27
Shapiro I. The state of democratic theory. Princeton University Press, 2009.	Soc. Sci. Economics	45	105
McNeil J.A., Frey R., Embrechts P. Quantitative risk management: Concepts, techniques, and tools. Princeton University Press, 2010.	Soc. Sci. Economics	43	460
White T., Folkens P. The Human Bone Manual. Elsevier, 2005.	Soc. Sci. Economics	40	101
Biddle S. Military power: Explaining victory and defeat in modern battle. Princeton University Press, 2010.	Soc. Sci. Economics	35	91
Rotberg R.I. When states fail: Causes and consequences. Princeton University Press, 2010.	Soc. Sci. Economics	33	72
Warren M.R. Dry bones rattling: Community building to revitalize American democracy. Princeton University Press, 2010.	Soc. Sci. Economics	33	157
Gutmann M.C. The meanings of macho: Being a man in Mexico City. University of California Press, 2006.	Soc. Sci. Economics	32	20

Table 8. The 20 Sciences books that were mentioned most often in academic syllabi or course reading lists.

Scopus Book	Scopus subject	Academic syllabi	Scopus citations
De Berg M., Cheong O., Van Kreveld M., Overmars M. Computational geometry: Algorithms and applications. Springer, 2008.	Computer Science	101	1
Jackson M.O. Social and economic networks. Princeton University Press, 2010.	Mathematics	89	465
Crabtree R.H. The Organometallic Chemistry of the Transition Metals. John Wiley, 2005.	Chemistry	88	1757
Ashenden P. The designer's Guide to VHDL. Elsevier, 2008.	Computer Science	88	28
Enderle J., Blanchard S., Bronzino J. Introduction to Biomedical Engineering. Elsevier, 2005.	Engineering	79	85
Mallat S. A. Wavelet Tour of Signal Processing. Elsevier, 2009.	Engineering	66	545
Bazaraa M.S., Sherali H.D., Shetty C.M. Nonlinear Programming: Theory and Algorithms. John Wiley, 2005.	Mathematics	54	3065
Harris D.M., Harris S.L. Digital Design and Computer Architecture. Elsevier, 2007.	Computer Science	53	31
Rao S.S. The finite Element Method in Engineering. Elsevier, 2005.	Engineering	52	91
Donahoo Michael J., Calvert Kenneth L. TCP/IP Sockets in C: Practical Guide for Programmers. Elsevier, 2009.	Computer Science	50	3
Karl H., Willig A. Protocols and Architectures for Wireless Sensor Networks. John Wiley, 2006.	Engineering	50	742
Pinedo M.L. Scheduling: Theory, algorithms, and systems. Springer, 2008.	Mathematics	45	0
Pahl G., Beitz W., Feldhusen J., Grote K.-H. Engineering design: A systematic approach. Springer, 2007.	Engineering	43	0
Albert J., Gentleman R., Parmigiani G., Hornik K. Bayesian computation with R. Springer, 2009.	Mathematics	41	0

Hirsch C. Numerical Computation of Internal and External Flows: The Fundamentals of Computational Fluid Dynamics. Elsevier Ltd, 2007.	Chemistry	39	115
Cassandras C.G., Lafortune S. Introduction to discrete event systems. Springer, 2008.	Engineering	38	0
Stein E.M., Shakarchi R. Real analysis: Measure theory, integration, and Hilbert spaces. Princeton University Press, 2009.	Mathematics	37	19
Hobbie R.K., Roth B.J. Intermediate physics for medicine and biology. Springer, 2007.	Physics	37	0
Hauer F.R., Lamberti G.A. Methods in Stream Ecology. Elsevier, 2007.	Environment Science	35	37
Witten I., Bainbridge D., Nichols D. How to Build a Digital Library. Elsevier, 2010.	Computer Science	33	27

Table 9. The 15 countries containing the most automatic syllabus mentions found in this study.

Country	Domain (academic)	No.	%
US	.edu/	9,560	49.8
UK	.uk/ (.ac.uk/)	1,448	7.5
Canada	.ca/	1,115	5.8
India	.in/ (edu.in/ or ac.in/ or other)	1,089	5.7
Poland	.pl/ (edu.pl/)	401	2.1
Hong Kong	.hk (edu.hk)	384	2.0
Italy	.it/	381	2.0
Sweden	.se/	348	1.8
Turkey	.tr (.edu.tr/)	313	1.6
Japan	.jp/ (.ac.jp/)	289	1.5
Germany	.de/	248	1.3
Israel	.il/ (ac.il/)	237	1.2
Taiwan	.tw/ (.edu.tw/)	236	1.2
China	.cn/ (.edu.cn/, .ac.cn/ or other)	205	1.1
Switzerland	.ch/	142	0.7