Goodreads Reviews to Assess the Wider Impacts of Books[1]

Kayvan Kousha, Mike Thelwall and Mahshid Abdoli

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, m.abdoli}@wlv.ac.uk

Although peer-review and citation counts are commonly used to help assess the scholarly impact of published research, informal reader feedback might also be exploited to help assess the wider impacts of books, such as their educational or cultural value. The social website Goodreads seems to be a reasonable source for this purpose because it includes a large number of book reviews and ratings by many users inside and outside of academia. To check this, Goodreads book metrics were compared with different book-based impact indicators for 15,928 academic books across broad fields. Goodreads engagements were numerous enough in the Arts (85% of books had at least one), Humanities (80%) and Social Sciences (67%) for use as a source of impact evidence. Low and moderate correlations between Goodreads book metrics and scholarly or non-scholarly indicators suggest that reader feedback in Goodreads reflects the many purposes of books rather than a single type of impact. Although Goodreads book metrics can be manipulated they could be used guardedly by academics, authors, and publishers in evaluations.

Introduction

Many academics write books for research or teaching or to engage with a wider non-academic audience. In the arts and humanities and some social sciences, monographs are often scholars' primary academic outputs (Nederhof, 2006; Huang & Chang, 2008). For instance, in the humanities over a half (55%), in the arts a third (33%), and in the social sciences just under a quarter (22%) of submissions to the 2014 UK Research Excellence Framework (REF) were books compared with only about 0.5% in science, engineering and medicine (REF, 2014). Some scholars also write textbooks to support academic teaching or introductory books to engage with wider audiences. Hence, books could have educational or public interest value (Kousha & Thelwall, 2015). Moreover, academics in the arts and humanities may write books with art, history or literary themes that can aim to culturally enrich a non-academic audience (Small, 2013). A range of indicators is therefore needed in order to provide book authoring scholars with evidence to help them to demonstrate the diverse impacts of their works.

Peer review is often used for the impact assessment of articles, despite the possibility of biases in peer judgments (Weller, 2001). Books tend to be much longer, requiring more time and effort for subject experts to assess. Citation counts can also be a useful quantitative indicator to support peer-review (Moed, 2005), especially when peer-review is not feasible, although it has

---

many limitations (MacRoberts & MacRoberts, 1989), particularly in book-based fields (Hicks, 1999). For instance, the coverage of academic books in the current book citation databases is narrow (Gorraiz, Purnell, & Glänzel, 2013; Torres-Salinas et al., 2013) and there may be indexing problems (Leydesdorff & Felt, 2012). Formal citations to books could be a useful indicator of their scholarly impact. However, academic book reviews may reflect the wider impacts of books, such as educational or cultural influence, rather than, or in addition to their research values. Although scholarly book reviews are rarely cited (Diodato, 1984; Zuccala & van Leeuwen, 2011), they are important academic outputs (Spink, Robins, & Schamber, 1998; Hartley, 2006). In terms of the relationship between the number of book reviews and citations to books, one study showed that the correlations could differ between fields and were considerably higher in Literature (.637), History (.608) and Psychology (.502) than in Biology (.214), Chemistry (.127) and Mathematics (.123) (Gorraiz, Gumpenberger, & Purnell, 2014), suggesting that review and citation counts are most similar in the humanities and social sciences. There have been attempts to use alternative indicators to help assess non-scholarly book impacts, such as counting library holdings ("libcitations") (White, Boell, Yu et al., 2009), library loan statistics (Cabezas-Clavijo et al., 2013), publisher prestige (Giménez-Toledo, Tejada-Artigas & Mañana-Rodríguez, 2013) and web-based book impact (For a review see: Kousha & Thelwall, 2015).

Scholarly book reviews can be an alternative source to aid book impact assessment, especially in book-based fields (Lindholm-Romantschuk, 1998; Spink, Robins, & Schamber, 1998; Zuccala & Van Leeuwen, 2011), although book reviews are rarely regarded as research outputs (East, 2011). In contrast to research articles, which "commonly avoid critical references", scholarly book reviews are a kind of "published peer review" and are "centrally evaluative" (Hyland, 2000, p. 41). Scholarly book reviews are read by many scholars to support their research and teaching (Spink, Robins, & Schamber, 1998; Hartley, 2006). Because of these academic values, scholarly book reviews may be helpful for the impact assessment of books (Shaw, 1991; Nicolaisen, 2002), even though this is not their primary purpose, and there has been an attempt to automatically assess the quality of book reviews for bibliometric evaluations of books (Zuccala, van Someren, & van Bellen, 2014).

Average numbers of reviews per book and book reviews per user in Goodreads are higher than in Amazon.com, but Amazon book reviews are much longer (Dimitrov, Zamal, Piper, & Ruths, 2015). An analysis of 64 English language books that were nominated for book awards (2007-2011) found that although winning books attracted more readers, their Goodreads ratings decreased because they received more attention and criticism from book readers (Kovács & Sharkey, 2014). There is little correlation between book-based metrics (book reviews, and ratings) and social activities (numbers of friends and followers) (Thelwall & Kousha, in press). An analysis of a sample of 8,538 history and geography books found a low but statistically significant positive Spearman correlation (.212) between the number of Goodreads ratings and the number of Scopus citations. A further analysis of 997 books with the most Scopus citations and Goodreads reader ratings also found a weak association between citations and reader ratings (.190),

suggesting that the citation impacts of books, as reported by Scopus, are only loosely reflected in the numbers of Goodreads book reviews (Zuccala, Verleysen, Cornacchia, & Engels, 2015).

The main aim of this study is to assess whether metrics associated with Goodreads book reviews can reflect wider benefits of books such as the research, teaching or cultural. Book reviews are produced in many fields, in which books may have very different types of value. For instance, students commonly use textbooks for their courses in some fields (Gurung & Martin, 2011) and some popular science books may be widely read by non-academics. Hence, both academic textbooks and introductory science books might be highly reviewed in online book review websites but have little citation impact (Kousha & Thelwall, 2016).

Although one study has investigated the relationship between Goodreads reader ratings, Scopus citations and WorldCat library holdings for history books (Zuccala, Verleysen, Cornacchia, & Engels, 2015), the current paper reports an investigation into the scholarly value of Goodreads book metrics using additional book-based metrics including Google Books citations and Wikipedia citations (see methods) in order to give insights into the different types of impacts that book reviews may reflect. It is possible to construct a hybrid metric of the overall engagement of users with a book by totaling all recorded user actions that are relevant to a book. These include reviewing a book, rating it, or listing it as read, to read, currently reading, or added to a collection. The total of all these indicators is called here the *engagement* count. The following questions are addressed.

1. Are Goodreads book metrics (reviews, engagements and average ratings) frequent enough to be a useful source of wider impact evidence for academic books?
2. Do Goodreads book metrics reflect the scholarly or other impacts of academic books?
3. How do the answers to the above questions vary between broad fields and scholarly impact indicators?

Methods

The Thomson Reuters Book Citation Index (BKCI) was used for the main analysis because it includes a large number of academic books and citation information. BKCI combines citations from books and citations from journal articles and is therefore slightly different source of *scholarly impact* (Kousha & Thelwall, 2014). In order to investigate different types of scholarly and other impacts, a range of types of data was collected. Google Books citation counts were used for *book-based scholarly impact*. These are likely to be relatively numerous in the arts and humanities compared with traditional citation databases (Kousha & Thelwall, 2009).

Wikipedia was used as a convenient alternative impact source because Wikipedia citations may reflect the informational or non-scholarly impacts of books within society, especially in the arts and humanities and in some areas of the social sciences (Kousha & Thelwall, in press). In addition, the number of library holdings was used as evidence of overall uptake. This presumably reflects the overall *popularity* of a work in terms of the educational or wider cultural benefits of books to library users rather than any particular type of impact (White, Boell, Yu et al., 2009).

Data Sets

A total of 15,928 BKCI-indexed books published during 2008-2010 were identified. Although BKCI starts from 2005, our source had not subscribed to the 2005-2007 data. Over 60,000 academic books were indexed by the Web of Science (The Book Citation Index, 2015) at the time of the study. For BKCI book searches a predefined query was used in the "Publication Name" field[2] in the Thomson Reuters Book Citation Index-Social Sciences & Humanities (BKCI-SSH) and Book Citation Index-Science (BKCI-S). The year 2010 was selected as the latest year to analyse in order to give at least five years for books to be cited and to attract online reviews or ratings in Goodreads. Only English language books were included in order to keep the collection relatively homogeneous. Book chapters, book series and trade publications were excluded for the same reason and because citations to individual book chapters and book series are not always reliable in BKCI (Leydesdorff & Felt, 2012; Kousha & Thelwall, 2014). Books without ISBNs were also omitted because ISBNs were needed to gather Goodreads metrics and library holding statistics (see below).

Due to the relatively small numbers of books in some subject areas and years, related disciplines were combined to generate broader fields for correlation analyses. The OECD Category Scheme for Web of Science subjects was initially used as a practical method to merge specific Web of Science subject areas into six broad fields: Humanities, Social Sciences, Natural Sciences, Agricultural Sciences, Medical and Health Sciences, and Engineering and Technology (Thomson Reuters, 2012). For instance, based on this scheme the WoS category HISTORY & PHILOSOPHY OF SCIENCE was classified under Humanities, although some books in this category might be more related to science (e.g., "Short History of Mathematical Population Dynamics" by Nicolas Bacaër) and it was not possible to manually check these cases. We combined Natural Sciences and Agricultural Sciences to represent "Science" and used the sub-category "Arts" under Humanities to combine books with arts subjects in WoS in the OECD scheme (Architecture, Art, Music, Dance, Theater, Film, Radio & Television). The "TC" field (times cited) field was used in BKCI for counts of citations to books.

Goodreads reviews and ratings

Every book can have a dedicated page in Goodreads. This page describes the book and includes metadata, such as the author name, publisher, and publication year. In addition, the page records the number of Goodreads users that have reviewed the book, the average rating of the book given by these members (on a scale of 1 to 5), and the most recent member reviews of the book. The Goodreads Applications Programming Interface (API) (http://www.goodreads.com/api) was used to automatically extract the number of text reviews, ratings and engagements for each

---

2 . (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*)

Scopus and BKCI book via its ISBN. For this, a program was added to the free *Webometric Analyst* software (http://lexiurl.wlv.ac.uk, the "Get Goodreads Reviews Stats From ISBNs" button in the Books tab).

Some books in the BKCI (10%) dataset had more than one ISBN (up to three). This is possible because different editions (e.g., second edition) and formats (e.g., e-book, paperback and hardcover) can have separate ISBNs. The Goodreads Search API gives aggregated statistics for text reviews ("WorkTextReviews"), the total number of ratings, text reviews and engagements ("WorkReviews") and average user ratings out of five ("AverageRating") for books with multiple editions. Out of 376 BKCI books with three ISBNs in the WoS dataset, all were found in Goodreads using the third WoS ISBN, and most were also found using the first (359) or second (289) ISBN. Thus, the last ISBN was used to search for books in Goodreads, followed by either the first or second ISBN to extract any Goodreads metrics from all available ISBNs. Ultimately, 90% (14,111 out of 15,928) of BKCI-indexed books were found in Goodreads.

Google Books citations

Google Books API searches were used in Webometric Analyst (Google Book Search in the Books tab) to count citations from books to the sample of BKCI books. Although Google Books does not include a citation index, Webometric Analyst automatically constructs Google Books queries that are designed to identify citations and then removes false matches from the results (for more details see: Kousha & Thelwall, 2014). These queries combined the last name(s) of up to three authors, a phrase search for the first ten words from the book title, and the publication year, as in the following example.

Kemp-Welch "Poland under Communism: A Cold War History" 2008

Books with one or two word titles (e.g., "Birds" or "Breast Cancer") were excluded because short titles can generate many false matches. For books with three, four or five words in their titles, publisher names were also added to the queries to increase the precision of the searches. Publisher names in both BKCI were standardised before adding them to the queries, such as "University Press of Kentucky" instead of "UNIV PRESS KENTUCKY" in BKCI.

Larsson "Dissociative Recombination of Molecular Ions" "Cambridge University Press" 2008

Wikipedia citation searches

The Bing API in Webometric Analyst was used to locate Wikipedia articles citing books in the data set (for method details, see: Kousha & Thelwall, in press). The strategy described above for Google Books citation searches was also used to generate queries for BKCI books but the *site:wikipedia.org/wiki/* command was added at the end of each query to limit the search results to Wikipedia articles, as shown in the example below.

Chemero "Radical Embodied Cognitive Science" 2009 "MIT Press" site:wikipedia.org/wiki

WorldCat library holdings (Libcitation)

The OCLC WorldCat library holding catalogue (http://oclc.org/worldcat/catalog.en.html) includes more than 2.3 billion books and other items from academic and non-academic libraries (e.g., national or public libraries) around the world (see, worldcat.org/webservices/registry/xsl/search-advanced) and is the main free international source of library collection information. The WorldCat Search API was used with the ISBNs of Scopus and BKCI books, again through Webometric Analyst, in order to record the library holdings for each book. Almost all BKCI-indexed books (99.3%) across the selected fields and years were found through WorldCat API searches, perhaps because WorldCat is one of the main information providers for Goodreads (https://www.goodreads.com/topic/show/766331-announcement-goodreads-to-import-worldcat-library-of-congress-data-to). Note that WorldCat reports combined statistics for all editions of a book, even if they have different ISBNs and so the WorldCat data is not strictly comparable with the other sources, which are restricted to a single publication year.

Analyses

Spearman correlations were calculated between Goodreads book metrics, citation indicators and library holding statistics in order to find evidence of Goodreads book metrics reflecting a type of scholarly or other impact. Correlations were calculated independently for each field and year to avoid mixing data in a way that would reduce the correlation coefficients (Fairclough & Thelwall, 2015). The percentage of books with a non-zero score on an indicator was used to judge the prevalence of a particular data type. This is not enough to decide whether an indicator is useful and so average values per book were also calculated for each data source. The geometric mean was used for this because it is more stable than the arithmetic mean for highly skewed data sets, like those analysed here. It is also a better central tendency indicator than the median for most of the data because the preponderance of low values makes the median too coarse grained for effective comparisons (Thelwall, 2016).

An additional analysis was conducted for a total of 5,343 books matched in both BKCI and Scopus published during 2008-2010 in order to allow comparisons between BKCI and Scopus citation counts.

A follow-up content analysis was conducted on a random sample of 50 BKCI books with at least one Goodreads book review from each of the Social Sciences, Arts and Humanities, Medical Sciences, Science, and Engineering (n=250 altogether). One review per book was randomly selected for books with more than one review. Three independent coders and a predefined classification scheme (see below) were used to identify (a) the contents and (b) typical types of impacts in Goodreads book reviews. Cohen's kappa (Cohen, 1960) was calculated for each subject area to assess the inter-rater agreement rates between coders.

  a) The contents of Goodreads book reviews:
     • A social comment that is not about the content of the book (e.g., "My daughter gave me this book for Christmas!" or "My teacher wrote this book.")

- Just an opinion about the book (e.g., "I thought it would be more entertaining." or "Absolutely fascinating read and a great idea.")
- Links to other reviews (e.g., "Read about this book in this article: http://www.nytimes.com/2010/06/09/dining/09curious.html?hpw")
- A brief description of the book, giving some information about books beyond just an opinion.
- A full book review, giving a substantial amount of information about the contents of the book, such as an overview, a list of chapters, a story overview, or a comment on the writing style.
- Other - texts that do not clearly match the other classes, such as "really?"
  b) Types of impacts in book reviews:
- Evidence of research or scientific impact (e.g., "This book is relevant for my research/PhD" or "I will use this for my future papers/publications").
- Evidence of teaching impact ("I'll use this book for my teaching" or "It is the text book for the course I am studying").
- Evidence of professional or technical impact (e.g., "Strongly recommend for anyone who works in the field of Computer Engineering" or "Recommend for anyone working in the U.K. health system").
- Other types of impacts, such as cultural, artistic or societal (e.g., "Probably the best general history of Thailand in the English language" or "Recommend for everyone interested in First World War era music").

Inter-coder consistency was assessed using Cohen's kappa for each category and subject (Table A in the online appendix, https://dx.doi.org/10.6084/m9.figshare.3122173.v3). The coders were three experienced content analysts, one with an information science degree and two with information science PhDs. Despite the use of a coding scheme, pre-coding testing, three rounds of pilot testing and coding scheme refinements before the main task, the kappa values were low overall due to the complex and subjective nature of the concept of impact, the relative scarcity of positive examples of impact (Viera & Garrett, 2005). The values were compared to standard heuristic guidelines, which suggest that at least 0.21 indicates fair agreement, 0.41+ indicates moderate agreement and 0.61+ indicates substantial agreement and 0.81+ indicates almost perfect agreement (Landis & Koch, 1977). Considering the cross-subject average, the only category for which there was at least moderate agreement between all three pairs of coders was 'Teaching Impact'. The 'Type' category achieved close to moderate agreement overall and the 'Professional Impact' category had fair agreement overall, but the 'Research Impact' and 'Other Impact' categories had no real agreement.

The results were reported by averaging the scores of the coders, which should make them more reliable overall than suggested by the kappa values because they represent the average opinions of three trained expert coders with impact evaluation subject knowledge. Nevertheless, the Research Impact and Other Impact results should be treated as speculative and the Type and Professional Impact results should be treated with some caution.

Results

BKCI-indexed books

For the BKCI-indexed dataset the average (geometric mean) number of Goodreads engagements for Arts and Humanities BKCI books is particularly high (Table 1). In contrast, in Medical Sciences, Science and Engineering the average numbers of BKCI citations are two to three times higher than the average number of Goodreads engagements. The average number of Goodreads engagements per book seems to be high enough for a viable indicator (Table 1), especially in the Arts (ranging from 3.5 to 4.3) and Humanities (2.9 to 4.7). Moreover, over 80% of books in Arts and Humanities published in 2008-2010 had at least one Goodreads engagement. Nevertheless, the proportion of books with at least one Goodreads review seems to be too low for routine impact assessment purposes in all fields.

Table 1. Goodreads book metrics, citation and library holding indicators for books indexed by BKCI during 2008, 2009 and 2010 across six fields (n=15,928). Statistics in the table: Number of books; percentage of books with at least one citation, library holding or review (geometric mean).

Broad field (merged Web of Science subjects)	Year	BKCI citations	GB citations	Wikipedia citations	Library holdings	Goodreads reviews	Goodreads engage.*
Social Sciences (e.g., Sociology; Education; Psychology; Political Science; Economics and Business; Management)	2008	1,665; 79.5% (5.7)	1,641; 86.7% (6.0)	1,641; 39.1% (0.6)	1,665; 99.5% (362.6)	1,404; 15.3% (0.2)	1,404; 72.8% (2.2)
	2009	2,338; 76.9% (4.6)	2,293; 81.4% (4.4)	2,293; 33.1% (0.5)	2,338; 99.4% (364.3)	2,218; 14.1% (0.2)	2,218; 62.7% (2.0)
	2010	2,310; 74.7% (3.4)	2,280; 78.5% (3.5)	2,280; 29.8% (0.4)	2,310; 98.9% (360.4)	1,857; 13.7% (0.1)	1,857; 65.8% (1.6)
Humanities (e.g., History; Philosophy; Ethics; Religion; History & Philosophy of Science; Cultural Studies; Linguistics; Literature)	2008	1,012; 85.2% (6.3)	985; 87.9% (6.3)	985; 54.4% (1.2)	1,012; 99.5% (434.9)	873; 27.3% (0.3)	873; 84.7% (4.7)
	2009	1,213; 86.4% (5.1)	1,178; 84% (4.9)	1,178; 48.5% (0.9)	1,213; 99.3% (436.9)	1,164; 23.8% (0.3)	1,164; 75.3% (3.8)
	2010	1,148; 81.4% (4.1)	1,118; 81% (3.8)	1,118; 48.4% (0.9)	1,148; 98.6% (471.1)	870; 25.6% (0.3)	870; 79.4% (2.9)
Arts (e.g., Art; Music; Theatre; Dance; Film; Radio & Television; Architecture)	2008	186; 83.3% (4.1)	179; 88.3% (5.6)	179; 58.1% (1.5)	186; 99.5% (553.2)	159; 25.2% (0.3)	159; 85.5% (4.1)
	2009	211; 83.4% (3.8)	193; 82.4% (3.8)	193; 52.8% (1.1)	211; 100% (567.2)	195; 21.5% (0.2)	195, 79.5% (3.5)

Subject area	Year						
	2010	224; 79.9% (3.1)	220; 80.9% (3.5)	220; 60.5% (1.5)	224; 99.1% (632.1)	185; 30.3% (0.4)	185, 89.7% (4.3)
Medical Sciences (e.g., General & Internal Medicine; Surgery; Nursing; Immunology; Pharmacology & Pharmacy)	2008	258; 72.9% (4.2)	244; 79.5% (3.4)	244; 34.4% (0.5)	258; 98.8% (351.0)	235; 9.4% (0.1)	235; 67.7% (1.6)
	2009	485; 54.2% (2.0)	468; 57.3% (1.4)	468; 20.5% (0.3)	485; 99.6% (368.1)	459; 5.7% (00.0)	459; 46.6% (0.9)
	2010	846; 54.5% (1.5)	813; 47.2% (0.9)	813; 18.2% (0.2)	846; 99.1% (335.5)	807; 4.6% (0.0)	807; 42.6% (0.7)
Science (e.g., Chemistry; Physics; Computer Science; Mathematics; Environmental Sciences & Ecology; Genetics)	2008	460; 66.7% (5.9)	437; 70.3% (2.4)	437; 35.5% (0.6)	460; 99.8% (440.6)	406; 11.3% (0.1)	406; 70.7% (2.0)
	2009	816; 57.6% (3.1)	785; 59.1% (1.5)	785; 28.4% (0.4)	816; 99.6% (384.7)	763; 9.4% (0.1)	763; 58.3% (1.5)
	2010	1,172; 63% (3.0)	1,114; 53.2% (1.1)	1,114; 23.4% (0.3)	1,172; 99.7% (337.9)	1,074; 5.7% (0.1)	1,074; 48.8% (0.9)
Engineering (e.g., Mechanics; Materials Science; Electrical & Electronic Engineering; Construction & Building Technology)	2008	291; 70.1% (6.3)	276; 69.2% (2.2)	276; 26.8% (0.3)	291; 98.6% (359.4)	256; 8.2% (0.1)	256; 66% (1.5)
	2009	561; 60.1% (3.2)	536; 57.5% (1.3)	536; 22.2% (0.2)	561; 99.8% (419.3)	538; 5.2% (0.0)	538; 51.1% (1.0)
	2010	732; 71% (3.7)	700; 52.9% (1.1)	700; 18.4% (0.2)	732; 99.9% (328.8)	648; 3.5% (0.0)	648; 46.6% (0.7)
All subjects areas		15,928; 73.4% (4.0)	15,460; 73.8% (3.1)	15,460; 34.6% (0.6)	15,928; 99.3% (386.1)	14,111; 14.1% (0.2)	14,111; 64.6% (1.9)

+Books with less than three words in their titles were excluded for Google Books and Wikipedia citation searches. *This includes number of ratings; text reviews and read listings (e.g., "to read" or "read").

Patterns of similarity between Goodreads and citation metrics

There are generally significant positive correlations between the Goodreads book metrics and the other indicators in all of the studied disciplines and years, although there are some disciplinary differences (Tables 2-7). Hence, in general, academic books with more formal citations or library holdings tend to obtain more online reviews, engagements and ratings in Goodreads. In Social Sciences, Medical Science and Engineering correlations between BKCI citations and Goodreads metrics are higher than between other indicators, such as Google Books and Wikipedia citations. In the Humanities, however, where books and monographs are important scholarly outputs, Goodreads book metrics have stronger associations with Google

Books citations. In the Arts, the highest correlations are often between Goodreads book metrics and Wikipedia citations, perhaps because cultural, artistic, or entertainment impacts are reflected in both sources. In Science, however, there is no pattern in the associations between Goodreads book metrics and other citation indicators, perhaps because scientific books are less reviewed in Goodreads (Table 1) or have other types of impacts. There are commonly significant moderate positive correlations between WorldCat library holdings and Goodreads engagements in Social Sciences, the Arts and the Humanities, indicating that in these fields both indicators may reflect the general popularity of academic books rather than any specific type of impact.

In summary, the results are consistent with Goodreads book metrics reflecting multiple types of intellectual impacts, including scholarly, and perhaps educational and cultural impacts.

Table 2. Spearman correlations between Goodreads book metrics and citation and other indicators in Social Sciences for BKCI-indexed books in 2008, 2009 and 2010 respectively.

Indicators	BKCI citations	GB citations	Wiki. Citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
BKCI citations	1	.513** .511** .470**	.178** .172** .196**	.219** .133** .125**	**.269** **.260** .216**	**.335** **.339** **.322**	**.261** **.272** .194**
GB citations		1	.299** .282** .301**	.243** .126** .101**	.182** .204** **.250**	.309** .269** .317**	.223** .207** **.212**
Wiki. citations			1	.135** .143** .143**	.172** .206** .218**	.220** .255** .297**	.151** .196** .197**
Library holdings				1	.346** .214** .199**	.371** .312** .229**	.263** .232** .132**
Goodreads reviews					1	.555** .527** .540**	.432** .422** .397**
Goodreads engage.						1	.710** .747** .713**
Goodreads ave. rating							1

** Significant at the p = 0.01 level. The highest correlations between citation and Goodreads metrics for each year are highlighted in bold.

Table 3. Spearman correlations between Goodreads book metrics and citation and other indicators in Humanities for BKCI-indexed books in 2008, 2009 and 2010, respectively.

Indicators	BKCI citations	GB citations	Wiki. Citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating

Indicators	BKCI citations	GB citations	Wiki. citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
BKCI citations	1	.534** .462** .450**	.167** .141** .188**	.190** .184** .235**	.232** .207** .243**	.342** .279** .301**	**.288**** .168** .170**
GB citations		1	.379** .278** .315**	.176** .137** .120**	**.310**** **.264**** **.251****	**.349**** **.313**** **.342****	.241** **.212**** **.185****
Wiki. citations			1	.170** .220** .254**	.284** .245** .239**	.296** .312** .303**	.188** .185** .182**
Library holdings				1	.312** .239** .187**	.389** .364** .225**	.215** .243** .173**
Goodreads reviews					1	.675** .628** .636**	.352** .437** .360**
Goodreads engage						1	.637** .685** .670**
Goodreads ave. rating							1

** Significant at the p = 0.01 level. The highest correlations between citation and Goodreads book metrics for each year are highlighted in bold.

Table 4. Spearman correlations between Goodreads book metrics with citation and other indicators in the Arts for BKCI-indexed books in 2008, 2009 and 2010, respectively.

Indicators	BKCI citations	GB citations	Wiki. citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
BKCI citations	1	.540** .500** .432**	.208** .142* .160*	.102 .072 .124	.298** .105 .120	.286** .098 .228**	.233** .043 .069
GB citations		1	.358** .373** .309**	.120 .060 .099	**.352**** .160* .242**	.358** **.310**** .275**	.268** .168* .162*
Wiki. citations			1	.330** .345** .387**	.297** **.256**** **.264****	**.382**** .297** **.310****	**.289**** **.174*** **.169****
Library holdings				1	.274** .239** .253**	.396** .345** .253**	.180* .254** 0.037
Goodreads reviews					1	.636** .621** .681**	.408** .304** .361**
Goodreads engage.						1	.653** .663** .560**

Indicators							Goodreads ave. rating
Goodreads ave. rating							1

** Significant at the p = 0.01 level. The highest correlations between citation and Goodreads book metrics for each year are highlighted in bold.

Table 5. Spearman correlations between Goodreads book metrics and citation and other indicators in Medical Science for BKCI-indexed books in 2008, 2009 and 2010 respectively.

Indicators	BKCI citations	GB citations	Wiki. citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
BKCI citations	1	.371** .462** .412**	.097 .181** .171**	.032 -.046 -.044	**.290** .197** **.189**	**.331** **.351** **.293**	**.275** **.301** **.372**
GB citations		1	.353** .267** .240**	.043 -.032 -.040	.188** **.220** .133**	.286** .296** .231**	.204** .286** .220**
Wiki. citations			1	.171** 0.042 -0.008	.152* .112* 0.027	.136* .220** .255**	.153* .171** .134**
Library holdings				1	.221** .122** .047	.077 .07 .015	.124 .049 .059
Goodreads reviews					1	.465** .384** .364**	.341** .387** .381**
Goodreads engage.						1	.659** .737** .693**
Goodreads ave. rating							1

* Significant at the p = 0.05 level. ** Significant at the p = 0.01 level. The highest correlations between citation and Goodreads book metrics for each year are highlighted in bold.

Table 6. Spearman correlations between Goodreads metrics and citation and other indicators in Science for BKCI-indexed books in 2008, 2009 and 2010 respectively.

Indicators	BKCI citations	GB citations	Wiki. citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
BKCI citations	1	.249** .347** .307**	.081 .225** .216**	.140** .011 -.012	0.064 .130** .153**	.127* .274** **.277**	**.152** .205** **.202**
GB citations		1	.420** .438** .275**	.112* .105** .008	**.141** .168** **.159**	**.203** .292** .212**	.121* **.242** .187**
Wiki. citations			1	.053 .051 .071*	.053 **.192** .149**	.125* **.300** .252**	0.023 .241** .163**

Indicators	BKCI citations	GB citations	Wiki. citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
Library holdings				1	.222** .203** .153**	.251** .176** .148**	.203** .179** .142**
Goodreads reviews					1	.471** .492** .390**	.337** .435** .409**
Goodreads engage.						1	.733** .748** .698**
Goodreads ave. rating							1

* Significant at the p = 0.05 level. ** Significant at the p = 0.01 level. The highest correlations between citation and Goodreads book metrics for each year are highlighted in bold.

Table 7. Spearman correlations between Goodreads book metrics and citations and other indicators in Engineering for BKCI-indexed books in 2008, 2009 and 2010 respectively.

Indicators	BKCI citations	GB citations	Wiki. citations	Library holdings	Goodreads reviews	Goodreads engage.	Goodreads ave. rating
BKCI citations	1	.279** .299** .291**	.148* .133** .170**	.081 .007 .083*	**.143*** **.159**** **.148****	.160* **.249**** **.295****	**.157*** .143** **.201****
GB citations		1	.304** .325** .260**	.114 .103* .081*	.107 .137** .166**	**.178**** .221** .164**	.156* .164** .167**
Wiki. citations			1	.048 .022 .067	.096 .151** .122**	.087 .210** .157**	.108 **.212**** .129**
Library holdings				1	.123* .114** .091*	.247** .109* .061	.113 .087* -.024
Goodreads reviews					1	.402** .341** .316**	.351** .346** .326**
Goodreads engage.						1	.684** .701** .645**
Goodreads ave. rating							1

* Significant at the p = 0.05 level. ** Significant at the p = 0.01 level. The highest correlations between citation and Goodreads book metrics for each year are highlighted in bold.

Books in both BKCI and Scopus

As a further analysis, Scopus books published in 2008-2010 were extracted and the ISBNs from the BKCI data set were matched against Scopus so that the two databases could be compared. A

total of 5,343 books were matched in this way to assess whether their citation counts differed significantly between databases and whether the databases had different relationships with Goodreads book metrics. The geometric mean number of citations per book and the percentage of books with at least one citation are much higher for BKCI-data than for Scopus data across all fields and years (Table B in the appendix section, https://dx.doi.org/10.6084/m9.figshare.3122173.v3). Scopus indexes an average (geometric mean) of less than 1 citation per book in all fields and years, except for Science and Medicine in 2008, whereas the WoS averages range from 2.6 to 9.4 (see also the Discussion). There are, on average, more citations from Google Books than from Scopus or Wikipedia, so BKCI and Google Books citations give the most citations for books and are probably the most appropriate for scholarly-related comparisons with Goodreads metrics. Although prevalence alone does not guarantee meaning for an indicator, low values undermine attempts to gauge the value of an indicator through correlation tests or, in practice, to distinguish between the impacts of different outputs.

Although there are generally significant positive correlations between BKCI citations and Goodreads book metrics across most fields and years, the correlations between Scopus citations and Goodreads book metrics are much weaker across most fields and year (Tables C-E in the appendix). Surprisingly, given the similar natures of BKCI and Scopus, the correlations between BCKI and Scopus citations are even lower than the correlations between BCKI and Goodreads metrics in Social Sciences and the Arts and Humanities, although not in Science and Medicine. Moreover, the correlations between BKCI and Google Books citations are moderate (from 0.364 to 0.497), whereas the correlations are much weaker between Scopus and Google Books citations (from 0.104 to 0.225). This may be partly due to the low numbers of Scopus citations to books (see the Discussion) and suggests that Scopus citations to books only partially reflect the scholarly impact of the books, despite the substantial number of books and journal articles indexed in Scopus (http://www.elsevier.com/solutions/scopus/content).

Content analysis of Goodreads reviews

Most reviews only expressed an opinion (positive or negative) about the book, ranging from 39% in Social Sciences to 60% in Engineering (Table 8). Nevertheless, about a third of Goodreads reviews went further by containing a brief book description. This was most common in Medical Sciences (40%) and rarest in Engineering (27%). Full book reviews were least rare in the Social Sciences (15%), followed by Medical Sciences (7%), Science (4%), Arts and Humanities (3%) and Engineering (1%). This suggests that there are disciplinary differences in the extent to which reviewers write comprehensive book reviews in Goodreads (but see the Limitations). The low figure for Arts and Humanities is particularly surprising, given the importance of book reviews in this field and presumably the prevalence of review-writing skills. For instance, about a quarter of the BKCI books had at least one Goodreads review in the Arts and Humanities, which is more than in the other subject areas (Table 1), and so there may be many more full book reviews in the Arts and Humanities than the other categories but they were not always randomly selected

because there were even more informal reviews. This may be because Arts and Humanities books reach a much wider public that may be tempted to post a short informal review.

Table 8. Goodreads book review types by field (n=250).

Fields	Social comment not about the book	Only an opinion about the book	Links to other reviews	A brief book description	A full book review	Other
Social Sciences	4.7%	39.3%	2.7%	32%	14.7%	6.7%
Arts and Humanities	2.7%	45.3%	2.0%	36.7%	2.7%	10.7%
Sciences	3.4%	40.4%	4.8%	37.7%	4.1%	9.6%
Medical Sciences	2.7%	45.3%	0.0%	39.9%	7.4%	4.7%
Engineering	1.3%	60.0%	0.7%	26.7%	1.3%	10.0%

In Science more Goodreads reviews had information about the professional or technical benefits of academic books than in the Arts and Humanities and Social Sciences (Figure 1). In contrast, in the Arts and Humanities and Social Sciences over a third (each 35%) of book reviews had information about cultural, history or art values of the books reviewed. Few Goodreads reviews discussed the research or scientific impacts of the books reviewed (2%-7%), and slightly more had information about educational or teaching benefits, although this was rare in the Arts and Humanities and Medical Sciences. It seems possible that educational value would be higher in the Arts and Humanities than suggested by this figure, since these are discursive subject areas and so the value of published books in education may be more universal and hence implicit in reviews.
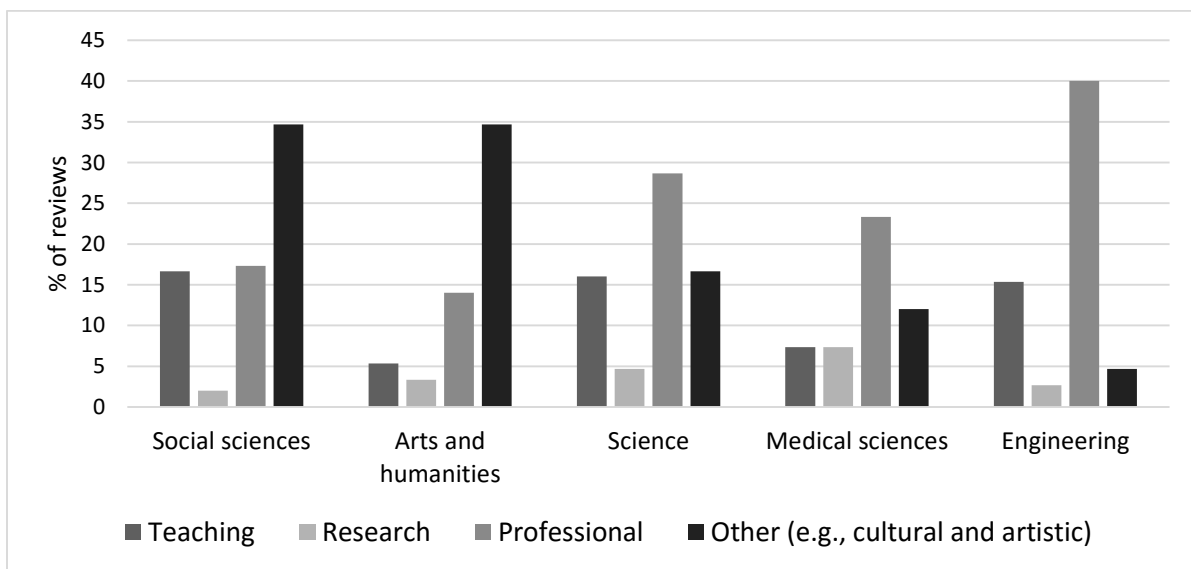

Figure 1. The percentage of Goodreads reviews with different types of impact by field.

Discussion

About 27% of the books (4,240 out of 15,928) had no BKCI citations and most of these uncited books (2,351 out of 4,240 or 55%) had at least one Goodreads engagement. The books with the most Goodreads engagements but no BKCI citations are from the Arts, Humanities, and Social Sciences (17 out of 20), with subject areas including History, Philosophy, Literature, Business, and Film, Radio, & Television (Table F in the appendix, https://dx.doi.org/10.6084/m9.figshare.3122173.v3). Although this is at least partly due to low average citation counts in these areas, it also consistent with Goodreads book metrics reflecting the cultural or educational benefits of academic books in the arts and humanities and some social science subject areas. For example, the book "*Portfolios of the Poor: How the World's Poor Live on 2 Dollars a Day*" published by Princeton University Press, had received 2,690 Goodreads social engagements but had no BKCI citations. It had also been acquired by over 1,000 libraries, suggesting that it might have primarily cultural or educational value rather than research impact as reflected by BKCI citations. In support of Goodreads tending to reflect public interest, the moderate correlations between Goodreads engagements and library holdings in the Social Sciences, the Arts and Humanities are higher than in Medicine, Sciences and Engineering, where scholarly outputs seem less likely to attract a public audience. In contrast, all except one of the 20 books with the most BKCI citations but no Goodreads engagements are in Science, Medical Sciences and Engineering (Table G in the appendix).

Goodreads engagements and BKCI citations most strongly correlate in Social Sciences, Medical Sciences and Engineering, whereas in the Humanities the correlations are higher between Goodreads engagements and Google Books citations. One reason for this difference could be that books in article-based fields attract many BKCI citations from WoS-indexed articles, whilst in book-based fields, Google Books citation searches can find relatively higher numbers of citations from other books (Kousha & Thelwall, 2014). Nevertheless, significant correlations do not demonstrate cause-and-effect relationships. Hence, the findings suggest, but do not prove, that Goodreads engagements reflect a degree of article-type scholarly impact in Social Sciences, Medical Sciences and Engineering, whereas they reflect book-type scholarly/educational/cultural impact in the Humanities.

As reported in Tables C-E in the appendix, the correlations between Scopus citations and Goodreads metrics are mostly weaker than between BKCI or Google Books citations and Goodreads metrics for books indexed in both BKCI and Scopus. It is not clear whether the difference is due to the use of broader categories or more selective coverage in the former case. The main reason for this seems to be the low numbers of citations to books from Scopus publications. For instance, 57% (3,059 out of 5,343) of the books in both BKCI and Scopus had at least one BKCI citation but no Scopus citations, whereas only 3% (153 out of 5,343) had one or more Scopus citations but no BKCI citations, despite the larger coverage of Scopus from books and articles (from 1996 onwards) than in WoS databases (Ball & Tunger, 2006). For instance, although the book "*Fiber Optical Parametric Pmplifiers, Oscillators and Related Devices*"

published in 2008 had 188 BKCI citations, in the record for this book in Scopus had a Times Cited score of 0. A cited reference search for the above book in Scopus retrieved 311 matches. These matches were documents indexed by Scopus that referenced the book, although the book had not been matched with its Scopus record so that the citations were not assigned to the book. Thus, technical problems with Scopus citation matching for books seem to have resulted in many citations having been omitted and the results are therefore unreliable.

Limitations

The Thomson Reuters BKCI was used as the primarily source of the citation data here. However, a practical limitation of using BKCI in this study is that it predominantly covers books in English from a small number of publishers (Gorraiz, Purnell, & Glänzel, 2013; Torres-Salinas et al. 2014) and claims that "priority is given to books and book series that have relatively greater citation impact" (Testa, 2012). More importantly, citations to different book editions are sometimes missing from the BKCI citation counts (Leydesdorff & Felt, 2012; Gorraiz, Purnell, & Glänzel, 2013; Chi, Thijs, & Glanzel, 2015). Hence, results from BCKI citations might be underestimates for books with different editions in our data set. In contrast, the Google Books citations, Wikipedia citations and WorldCat library holdings methods capture citations to different editions and hence are comparable with Goodreads engagement counts. Hence, results from the correlation analysis between BCKI citations and other metrics should be guardedly interpreted in the absence of a database to provide aggregated citations for multiple editions of each book.

Only English language BKCI-indexed books were used and most Goodreads users are presumably from the USA and other mainly English-speaking countries. For instance, according to Alexa.com, 44% of Goodreads visitors are from the United States (http://www.alexa.com/siteinfo/goodreads.com), giving a US-bias to Goodreads metrics. It therefore seems likely that the value of Goodreads would be much lower for languages other than English and also for books that address national audiences outside of the USA (e.g., Nigerian law). Moreover, over a three-quarters of Goodreads' users are female (Thelwall, in press), giving a gender bias. A future study might use books outside these citation databases to investigate whether the value of Goodreads would be lower for a wider spectrum of academic books, which seems possible. Future research might also extend the content analysis of Goodreads book reviews, using a larger sample size, to sub disciplines (e.g., history or literature) and compare the contents of Goodreads reviews with academic reviews (e.g., published reviews).

While there is no specific evidence of systematic manipulation by Goodreads reviewers (see Wijnhoven & Bloemen, 2014), online reviews and ratings of books can be manipulated (BBC, 2015; Hu, Bose, Koh, & Liu, 2012) and hence, as with most social web metrics, should not directly be used for research evaluation (Birkholz & Wang, 2011; Thelwall & Kousha, 2015; Wouters & Costas, 2012). For instance, authors have the ability to include positive reviews about their books through registering pseudonymous Goodreads accounts. Nevertheless, Goodreads metrics could

help authors and publishers to assess feedback from book readers inside and outside of academia because manipulation is a lesser problem for self-evaluation (Wouters & Costas, 2012).

Conclusions

This study examines whether Goodreads book metrics can be used to assess the research, teaching, cultural or other impacts of books. In answer to the first research question, the overall results show that there are enough Goodreads engagements for impact assessment purposes in the Arts, the Humanities and to some degree in Social Sciences. In the Arts (85%), Humanities (80%), and Social Sciences (67%) at least two thirds of the BKCI-indexed books had one or more Goodreads engagement, and a majority had engagements in Science (59%), Engineering (55%) and Medical Sciences (52%), showing that Goodreads book metrics are more widespread in book-based fields. Moreover, the Goodreads engagements geometric mean in the Arts (4) is higher than for BKCI citations (3.7), whereas in Science the Goodreads average (1.5) is much lower than for BKCI citations (4), confirming the particular value of Goodreads in some fields. There are very few Goodreads reviews in all fields, although a quarter of Arts and Humanities and 15% of Social Sciences BKCI-indexed books had at least one text review compared with 9% in Science, 6.5% in Medical Sciences and 5.6% in Engineering. Thus, Goodreads seems to give useful impact evidence in the Arts, the Humanities and perhaps also the Social Sciences.

In answer to the second research question, there are generally significant positive correlations between the Goodreads book metrics and other citation indicators and library holding counts in most of the fields and years (see Tables 2-7), although there are also disciplinary differences. The Spearman correlations are stronger between Goodreads engagements and BKCI citations in Social Sciences, Medical Sciences and Engineering. In Humanities the correlations are higher between Goodreads engagements and Google Books citations, probably due to importance of book citations in the humanities (Kousha & Thelwall, 2014). In the Arts Goodreads engagements correlate more strongly with Wikipedia citations, perhaps because many books in this field may have cultural or educational (informational) uptake as reflected in Wikipedia citations (Kousha & Thelwall, in press). The result also confirms that Wikipedia citations and library holdings have the strongest associations in the Arts (.330, .345 and .387), whereas in Science, Medical Sciences and Engineering there are generally very low correlations between Wikipedia citations and library holdings, suggesting that both indicators reflect the cultural or educational benefits of books to some extent. Future studies focusing on the types of science, engineering and social science books that get reviewed would be interesting to see whether they have wider societal, educational or cultural benefits or whether the reviews are written by other scholars.

In answer to the third research question, there are substantial disciplinary differences between the broad fields analysed as well as between more comparable subject areas. For instance, Goodreads book metrics are more numerous in the Arts and Humanities than in other fields. However, the correlations between Goodreads book metrics are stronger with BKCI citations in Social Sciences, Medical Sciences and Engineering, and with Google Books citations in the

Humanities, and with Wikipedia citations in the Arts, suggesting that Goodreads book metrics may reflect the multiple diverse intellectual impacts of books, such as scholarly, educational, cultural, and informational contributions. Surprisingly, Scopus citations had lower correlations with Goodreads metrics (Table C-E in the appendix) than did other scholarly indicators (BKCI and GB citations) throughout this study, confirming the previously found weak relationship (.190 and .212) between Scopus citations and Goodreads ratings (Zuccala, Verleysen, Cornacchia, & Engels, 2015). One reason for this weak association could be that Scopus does not fully integrate existing citations to books in its citation reports (see discussion).

In conclusion, Goodreads metrics are recommended for self-evaluations by publishers and authors in the arts, humanities and perhaps also the social sciences. However, because of the potential for manipulation, extreme caution should be exercised if they are used in formal evaluations – such as requesting an honesty declaration from the author and publisher. When used, they may partly reflect different types of impacts, including educational, cultural and informational, and the type of impact is likely to vary substantially by field and book type.

References

Ball, R., & Tunger, D. (2006). Science indicators revisited – Science Citation Index versus SCOPUS: A bibliometric comparison of both citation databases. Information Services & Use, 26(4), 293-301.

BBC (2015). Amazon targets 1,114 'fake reviewers' in Seattle lawsuit. http://www.bbc.co.uk/news/technology-34565631

Birkholz, J., & Wang, S. (2011). Who are we talking about?: The validity of online metrics for commenting on science. Paper presented in: altmetrics11: Tracking scholarly impact on the social Web. An ACM Web Science Conference 2011 Workshop, Koblenz (Germany), 14-15. http://altmetrics.org/workshop2011/birkholz-v0

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, Vienna, Austria, Vol. 2, pp. 1237-1252. Retrieved October 02, 2013, from http://arxiv.org/ftp/arxiv/papers/1305/1305.1488.pdf.

Chi, P., Thijs, B., & Glänzel, W. (2015). The challenges to embody a new data source: The Book Citation Index. ISSI Newsletter, 11(1), 24-29.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37-46.

Dimitrov, S., Zamal, F., Piper, A., & Ruths, D. (2015). Goodreads vs Amazon: The effect of decoupling book reviewing and book selling. In Proceedings of the Ninth International AAAI Conference on Web and Social Media, May 26-29, 2015, (pp. 602-605).Oxford University. http://piperlab.mcgill.ca/pdfs/Goodreads_ICWSM_2015.pdf

Diodato, V. (1984). Impact and scholarliness in arts and humanities book reviews: A citation analysis. Proceedings of the 47th Annual Meeting of the American Society for Information Science, 21, 217–220.

East, J.W. (2011). The scholarly book review in the humanities: An academic Cinderella? Journal of Scholarly Publishing, 43(1), 52–67.

Fairclough, R., & Thelwall, M. (2015). The influence of time and discipline on the magnitude of correlations between citation counts and quality scores. Journal of Informetrics, 9(3), 529–541.

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.

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

Hartley, J. (2006), Reading and writing book reviews across the disciplines. Journal of the American Society for Information Science, 57(9) 1194–1207.

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

Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. Decision Support Systems, 52(3), 674-684.

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.

Hyland K. (2000). Praise and criticism: interactions in book reviews, in K. Hyland (ed.), Disciplinary Discourses: Social Interactions in Academic Writing, Longman, Harlow, pp. 41–62.

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, 66(2), 309–320.

Kousha, K., & Thelwall, M. (2015). Web indicators for research evaluation. Part 3: books and non-standard outputs, El profesional de la información, 24(6), 724-736.

Kousha, K. & Thelwall, M. (2016). 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.

Kousha, K., & Thelwall, M. (in press). Are Wikipedia citations important evidence of the impact of scholarly articles and books? Journal of the Association for Information Science and Technology.

Kovács, B., & Sharkey, A. J. (2014). The paradox of publicity: How awards can negatively affect the evaluation of quality. Administrative Science Quarterly, 59(1), 1-33.

Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159-174.

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.

Lindholm-Romantschuk, Y. (1998). Scholarly book reviewing in the social sciences and humanities: The flow of ideas within and amongst disciplines. Westport, CT: Greenwood Press.

MacRoberts, M.H., & MacRoberts, B.R. (1989). Problems of citation analysis: A critical review. Journal of the American Society for Information Science, 40(5), 342-349.

Moed, H. F. (2005). Citation analysis in research evaluation. New York, New York: Springer.

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

Nicolaisen, J. (2002). The scholarliness of published peer reviews: A bibliometric study of book reviews in selected social science fields. Research Evaluation, 11(3), 129-140.

REF (2014). Results and submissions in Research Excellence Framework. http://results.ref.ac.uk/

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.

Small, H. (2013). The value of the humanities. Oxford, UK: Oxford University Press.

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.

Testa, J. (2012). The book selection process for the Book Citation Index in Web of Science. http://wokinfo.com/media/pdf/BKCI-SelectionEssay_web.pdf

Thomson Reuters. (2012). OECD Category Scheme. http://ipscience-help.thomsonreuters.com/incitesLive/globalComparisonsGroup/globalComparisons/subjAreaSchemesGroup/oecd.html

The Book Citation Index (2015). Thomson Reuters. http://wokinfo.com/products_tools/multidisciplinary/bookcitationindex/

Thelwall, M. (2016). The precision of the arithmetic mean, geometric mean and percentiles for citation data: An experimental simulation modelling approach. Journal of Informetrics, 10(1), 110-123.

Thelwall, M. (in press). Book genre and author gender: romance>paranormal-romance to autobiography>memoir. Journal of the Association for Information Science and Technology. http://wlv.openrepository.com/wlv/bitstream/2436/612738/1/GoodreadsBookGenreAuthorGenderPreprint.pdf

Thelwall, M., & Kousha, K. (2015). Web indicators for research evaluation, part 1: Citations and links to academic articles from the web. El Profesional de la Información, 24(5), 587-606. doi:10.3145/epi.2015.sep.08

Thelwall, M., & Kousha, K. (in press). Goodreads: A social network site for book readers. Journal of the Association for Information Science and Technology. http://www.scit.wlv.ac.uk/~cm1993/papers/GoodReadsASocialNetworkSiteForBookReaders_preprint.pdf

Torres-Salinas, D., Rodríguez-Sánchez, R., Robinson-García, N., Fdez-Valdivia, J., & García, J. A. (2013). Mapping citation patterns of book chapters in the Book Citation Index. Journal of Informetrics, 7(2), 412-424.

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.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. Family Medicine, 37(5), 360-363.

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.

Wijnhoven, F., & Bloemen, O. (2014). External validity of sentiment mining reports: Can current methods identify demographic biases, event biases, and manipulation of reviews? Decision Support Systems, 59(1), 262-273.

Wouters, P., & Costas, R. (2012). Users, narcissism and control: tracking the impact of scholarly publications in the 21st century. In: E. Archambault, Y. Gingras, V. Larivière (Eds) Proceedings of the 17th International Conference on Science and Technology Indicators. Montreal: Science-Metrix and OST (Vol. 2, pp. 487-497).

Zuccala, A., Verleysen, F. T., Cornacchia, R., & Engels, T. C. (2015). Altmetrics for the humanities: Comparing Goodreads reader ratings with citations to history books. Aslib Journal of Information Management, 67(3), 320-336.

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, 2011.

Zuccala, A., van Someren, M. & van Bellen, M. (2014). A machine-learning approach to coding book reviews as quality indicators: Toward a theory of megacitation. Journal of the Association for Information Science and Technology, 65(11), 2248–2260.