Received 14 November 2014
Revised 23 March 2015
Accepted 24 March 2015

Altmetrics for the humanities

Comparing Goodreads reader ratings with citations to history books

Alesia A. Zuccala
The Royal School of Library and Information Science, University of Copenhagen, Copenhagen, Denmark and Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam, The Netherlands
Frederik T. Verleysen
Centre for Research & Development Monitoring (ECOOM), Faculty of Political and Social Sciences, University of Antwerp, Antwerp, Belgium
Roberto Cornacchia
Spinque, Utrecht, The Netherlands, and
Tim C.E. Engels
Centre for Research & Development Monitoring (ECOOM), Faculty of Political and Social Sciences, University of Antwerp, and Antwerp Maritime Academy, Antwerp, Belgium

Abstract

Purpose – The purpose of this paper is to assess the value of Goodreads reader ratings for measuring the wider impact of scholarly books published in the field of History.

Design/methodology/approach – Book titles were extracted from the reference lists of articles that appeared in 604 history journals indexed in Scopus (2007-2011). The titles were cleaned and matched with WorldCat.org (for publisher information) as well as Goodreads (for reader ratings) using an API. A set of 8,538 books was first filtered based on Dewey Decimal Classification class 900 "History and Geography", then a subset of 997 books with the highest citations and reader ratings (i.e. top 25 per cent) was analysed separately based on additional characteristics.

Findings – A weak correlation (0.212) was found between citation counts and reader rating counts for the full data set ($n = 8{,}538$). An additional correlation for the subset of 997 books indicated a similar weak correlation (0.190). Further correlations between citations, reader ratings, written reviews, and library holdings indicate that a reader rating on Goodreads was more likely to be given to a book held in an international library, including both public and academic libraries.

Originality/value – Research on altmetrics has focused almost exclusively on scientific journal articles appearing on social media services (e.g. Twitter, Facebook). In this paper we show the potential of Goodreads reader ratings to identify the impact of books beyond academia. As a unique altmetric data source, Goodreads can allow scholarly authors from the social sciences and humanities to measure the wider impact of their books.

Keywords History, Books, Altmetrics, Scholarly communication, Citation analysis, Reader recommendation systems

Paper type Research paper

The authors would like to thank the reviewers of this paper for their many helpful suggestions for improvements. The authors are especially grateful to the Elsevier Bibliometrics Programme (http://ebrp.elsevier.com), and WorldCat.org as well as Goodreads for providing us with datasets as well as API access to the data used in this study. The authors also wish to thank the Center for Digital Humanities at the University of Amsterdam and the Flemish Centre for R&D monitoring (ECOOM) for granting us financial support for this project.

Aslib Journal of Information Management
Vol. 67 No. 3, 2015
pp. 320-336

2050-3806
DOI 10.1108/AJIM-11-2014-0152

1. Introduction

The objective of this study is to assess the value of Goodreads as an altmetric data source. Goodreads is one of "the world's largest sites for readers and book recommendations", and in March of 2015, the Goodreads page indicated a membership of 30 million people, 900 million books, and 34 million reviews (Goodreads, 2015).

Goodreads maintains a registry of both fiction and non-fiction; however, our primary aim is to use this platform to study the visibility of scholarly non-fiction. Specifically, we are interested in books that fit a particular set of criteria: non-fiction contributions to the field of history; published by a university or commercial press; and cited in journal articles that were index by Scopus under the assigned subject category of history (ASJC = 1,202). The present study is guided by the following questions:

(1) To what extent do scholarly History books that are cited in journal articles receive reader ratings on Goodreads?

(2) What are some of the characteristics of History books that are both considerably cited in scholarly journal articles and considerably rated on Goodreads? How do their citations and ratings compare to Goodreads reviews and to library holdings? How are citation scores and ratings distributed over publisher types? To what extent do citation scores and ratings differ across DDC History divisions?

2. The rise of "altmetrics" and public value of historical research

2.1 Research on altmetrics

Altmetrics is considered to be an umbrella term for assessing the presence or acknowledgement of scholarly research on the social web (Priem *et al.*, 2010). The aim of altmetrics is to augment our views on scholarly impact by considering new or "alternative" data sources for measurement, like social bookmarking systems (Haustein and Siebenlist, 2011), online reference managers (Li *et al.*, 2012), Twitter (Haustein *et al.*, 2014; Priem *et al.*, 2012a; Weller *et al.*, 2011), Wikipedia (Nielsen, 2007), or blogs (Shema *et al.*, 2014). Hammarfelt's (2014) research indicates that tweets to books published in the humanities are frequent, but that their coverage on other social media is low. Goodreads is therefore explored in this paper because there have been few studies of altmetric indicators focused on the humanities and because reader ratings on this platform may fill a void associated with altmetrics for scholarly books.

Recent work on altmetric indicators has focused on their relationship to citation counts. According to Thelwall *et al.* (2013) correlation measures for most cases are moderate; hence there really is no systematic evidence that altmetrics are valid proxies for citations. Moreover, citations take a long time to accumulate, while blogs or Tweets are practically immediate and collect rapidly (Thelwall *et al.*, 2013). Additional work by Haustein *et al.* (2014) draws further attention to the correlation problem in altmetrics, specifically for the broader field of biomedicine. When citations to thousands PubMed/WoS research articles were correlated with their Tweets, Spearman's rank measures were found to be quite low. Out of 26 different specialty areas, general and internal medicine presented the highest Spearman value of 0.327, with half of the specialties showing no correlation.

These findings raise questions about the uptake of scholarly documents on social media. Does this uptake help people to become more scientifically literate, more

engaged with science or demonstrate an increased public understanding of a particular field? Nightingale and Scott (2007) suggest that even if highly cited research published in top ranking journals is good for an academic discipline, it may not necessarily be good for society. With this in mind, altmetrics do not necessarily need to become proxies of scientific impact if they can provide us with "new ways to measure (public) engagement with research output" (Bornmann, 2014, p. 2).

2.2 The social impact of history

How do we know if scholarly or scientific research is "good" for society? According to Bornmann (2013, 2015) this is a question that has been circulating amongst policy-makers and academics since the early 1990s. For some time, the answer was situated within a Mode 2 framework for research (Nowotny *et al.*, 2003). With Mode 2, scientists were encouraged to study real-world problems and collaborate in a trans-disciplinary capacity. Research results were expected to generate socially robust knowledge with utilitarian criteria and values (Barré, 2005; Petit, 2004). In addition, scholars have tried to develop different terminology for evaluating the social, cultural, environmental, and economic returns of science; however, Bornmann (2013) indicates that separating the different areas of impact are difficult; moreover scientists "generally dislike impact considerations" (Holbrook and Frodeman, 2011, p. 244) because assessments tend not to do justice to different fields. The "scientific work of an engineer has a different impact than the work of a sociologist or historian" (Bornmann, 2013, p. 219); hence their research results "affect many different aspects of society" (Walker, 2011, para. 9).

To evaluate the impact of academic research, Donovan (2007, 2008) suggests a three-phased approach. The first phase, which is technometric in nature, focuses on the economic returns of research (i.e. technology transfer). The second sociometric phase covers social impact, from either a local or regional standpoint. During the third phase, distinct case studies may be carried out, using quantitative or qualitative data, to uncover complexities that differentiate specific fields. Drawn together, each phase contributes to a more comprehensive picture. With that said, there seems to be no reason why altmetric indicators could not be included in this approach, particularly during the sociometric phase. Overall, the three phases are not just valuable for natural or medical sciences, but for the social sciences and humanities as well. With the humanities, technometric impacts are notoriously difficult to ascertain, thus another more suitable term is needed. We discuss this further in our proposed research framework (i.e. Section 3).

The practice of studying history demands that scholars not only investigate and provide an accurate record of the past but also use this record in an original, thought-provoking analysis. According to Reid and Szreter (2008) history can always be studied for the "sheer pleasure of learning about other times, people and places for their own sake" (para. 1); however, its greater value rests with how it can inform the deliberative process of policymaking. This is partly *Why History Matters*; thus Tosh (2008) understands that different kinds of history lead to critical awareness and other forms that do not. "Historicity" or historical awareness does and can promote an understanding of present-day problems, but in practice, Tosh (2008) is not clear on how this works. A continued academic debate has been necessary to determine what the term "public history" means, or a history that actually administers to the social and political needs of the public.

Why does history matter? Historians and philosophers alike have long since tried to address this question, pointing to several possible explanations. These explanations range from rationalistic or utilitarian to philosophical and psychological. History can be of value to readers as a means to better understand and shape their present; historical awareness through reading can provide the basis of an informed and critical understanding of society, of "educated citizenship" (Tosh, 2008). Certain history books can spark controversies about dealing with painful episodes of the collective past, thereby stimulating collective reflection and public debate (Finkelstein and Birn, 1998). Personal motives also help to explain why individual readers can grow attached to books about history. Historical non-fiction helps many people to reconnect their own present with the society, culture and people of centuries past; thereby satisfying emotional, existential, and aesthetic needs (Ankersmit, 2005; Tollebeek and Verschaffel, 1992).

A "public history" or history that matters requires cooperation between academics and publishers, but there has been some tension in this regard. The inherent tension is rooted in the economic logic of publishing and the expectations that university faculties place on scholars. For instance, publishers may be hesitant to print and distribute specialized academic books if the chances of making a profit on them are slim. Moreover, humanities scholars, especially those working in the USA and Britain, are expected to publish monographs with a strong academic appeal. Research in the form of a book needs to be sufficient enough for the scholar to achieve faculty tenure (USA) or to receive a good score in the RAE/REF (Britain) (Allen and Heath, 2013; Cronin and La Barre, 2004). Some performance-based funding systems may alternatively encourage history scholars to publish articles (Verleysen and Engels, 2012). According to Thompson (2005) publishers have been forced in recent decades to adapt to a contracting market. In addition to printing academic monographs, they have had to diversify their output to include textbooks, trade books, regional interest books, etc. Another way to reconcile the opposing academic-economic logic is to encourage scholars to publish more monographs with both an academic and wider societal appeal. For publishers this implies a better chance of making a profit; for scholars, a better chance of getting their work published in the first place. This solution to the academic-societal balance now needs to be factored into the academic evaluation system.

More than most academic subjects (e.g. highly technical areas of science), historical research maintains an easy readership appeal for citizens. While it is the aim of academic/public libraries (physical or digital) to make books available for reading, historians are also eager to host blogs that will support the public's interest (see Poe, 2014). In the OCLC-WorldCat.org union catalogue, history is now ranked as the largest subject category, with over 2.5 million publications housed in libraries worldwide, and not just academic libraries. This makes history even more numerically substantial than other high-ranking subjects, like literature, philosophy and religion, or business and economics (Worldcat, 2014). From a commercial perspective, books from this field also carry weight: Amazon lists approximately 3.6 million history titles for sale, making it one of the most numerous book subject categories overall, and still slightly outnumbering literary fiction, at approximately three million (Amazon, 2015).

As historical works become salient and ready for public consumption, it is difficult to know just how much they are actually being consumed. It would be useful, for instance, to obtain purchasing statistics from Amazon or circulation statistics from specific libraries (academic or public); however, without such data there are ample reasons to focus on WorldCat.org and Goodreads. Goodreads reader ratings provide

quantitative information about the public uptake of books, whereas Goodreads reviews can tell us what readers actually think of the text. Moreover, Goodreads is linked to friendship networks on Facebook. When people read a particular book on Goodreads, there is at least a strong incentive to share opinions and recommend new books to "friends" from their social network (Trott and Naik, 2012).

3. Research framework

To evaluate the impact of history books, we follow Donovan's (2008) three-phased approach. With this approach, the term technometric is confined to "science, technology, engineering and medicine" and "relates to investment from industry, commercialization, and technology transfer" (Donovan, 2008, p. 2). For a field or subject in the humanities we suggest that it is more appropriate to use the term acadometric.

The acadometric phase of evaluating historical work relates to the root term "academy", which is the society of scholars who work together to create recognized standards. Here, the citation fits well as a practical measure, because citations are generally accepted as "registrations of intellectual property and peer recognition" (Moed, 2005, p. 194). At the very least, when an historian receives a citation from another scholar, his/her in-depth treatment of a subject has been recognized as having met academy standards. The citation might also account for socially construed agreements or disagreements among scholars on matters of historical interpretation. In history, there has been a reported increase in the share of journal articles to about half of the total publications (Den Hertog *et al.*, 2014; Engels *et al.*, 2012), yet within many journal articles, monographs are still the most highly cited item (Jones *et al.*, 1972; Dalton and Charnigo, 2004). This means that it is both reasonable and appropriate to focus on the impact of history books by counting their "citedness" in journal articles.

The sociometric, or second phase of evaluating impact can be constructed in two ways: with "libcitations" and with "altmetrics". The "libcitation", which was first introduced by White *et al.* (2009), denotes the perceived "cultural" benefit of a particular book, and the "altmetric" measure for that same book identifies its actual cultural uptake via social media. In this paper we confine ourselves to this sociometric approach. The case study approach, which Donovan (2008) indicates as the third evaluative phase, would require the analysis of Goodreads reviews, which is beyond the scope of this paper.

4. Research methods

4.1 Bibliographic data collection

The sampling process for our study began with a set of 42,353 filtered and merged variations of book titles (see Table III) from Scopus history journal article reference lists. All titles were cited at least once during the publication years 2007-2011 and identified as being registered on Goodreads during the 2006-2013 period. The Scopus data were granted to our research team in 2012 via the Elsevier Bibliometrics Research Programme (http://ebrp.elsevier.com/) and transferred to a Microsoft SQL database. All of the Scopus journals provided were assigned a classification code of 1202 (i.e. for history); but many were also multi-disciplinary with additional classification codes for subjects like political science, economics, science studies, religion, literary theory, and literary criticism.

The 42,353 book titles were matched with bibliographic records from the WorldCat.org union library catalog using an API developer key. The purpose of this procedure was to complete the bibliographic entries for each book by locating its publisher name, publisher location, OCLC accession number, and ISBN, including its

associated union catalog holding count. All holding counts (or "libcitations") were distinguished specifically between those from the academic libraries of the American Association of Research Libraries (ARL_Counts) and those from other international libraries (NONARL_Counts). The administration office at OCLC-WorldCat.org later provided us with a list of Dewey Decimal Classification (DDC) codes and descriptors corresponding to each book title's OCLC accession number.

After integrating the Scopus book references with data from WorldCat.org, we then conducted the API retrieval on Goodreads (www.goodreads.com). This search involved matching known ISBNs for the cited book titles with ISBNs and titles registered in Goodreads. The API downloads were conducted over the course of one week in November 2013, and the following information from Goodreads was collected: rating counts for the books, average star ratings (up to five stars), and review counts.

To focus our research on the field of history, we selected the books linked to the Main Dewey Decimal Classification of History and Geography (DDC = 900). Figure 1 outlines the main steps associated with the data sampling procedures, leading to an initial set of $n = 8{,}538$ book titles, and a smaller data set of interest, comprised of $n = 997$ titles. Table I presents the descriptive statistics related to the $n = 8{,}538$ data set. In Table II,

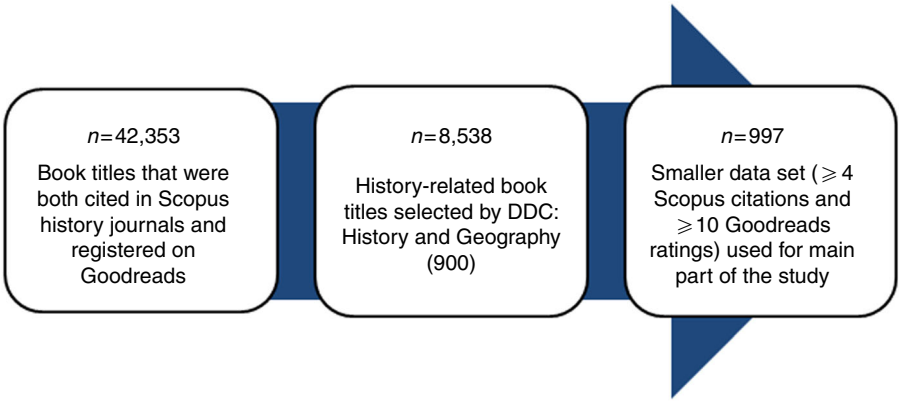


Figure 1.
Stages of the data collection process

	Scopus citations	Goodreads reader ratings	Goodreads reviews	Goodreads average star ratings
Mean	4.34	238.11	10.41	2.95
Median	2	2	0	3.60
SD	9.25	10,903.04	153.57	1.64
Skewness	19.29	89.40	55.81	−0.98
SE of skewness	0.03	0.03	0.03	0.03
Minimum	1	0	0	0.00
Maximum	433	996,125	11,541	5.00
Percentile 25	1	1	0	2.50
Percentile 50	2	2	0	3.60
Percentile 75	4	10	2	4.00

Note: $n = 8{,}538$

Table I.
Descriptive statistics for the Scopus journal citations (2007-2011) and Goodreads reader ratings and reviews (2006-2013)

Book Titles Cited in Scopus history journals (2007-2011)	Rated by Readers on Goodreads (2006-2013) Reader ratings below the 75th percentile (< 10)	Reader ratings at or above the 75th percentile ($> = 10$)
Citations below the 75th (< 4)	$n = 4{,}641$ Low-impact books	$n = 1{,}242$ Social impact books
Citations at or above the 75th percentile ($> = 4$)	$n = 1{,}658$ Academic impact books	$n = 997$ Books under examination

Data structure
SOURCE_ID: 618933
ISBN: 292775962
SOURCE_TITLE: *Anglos and Mexicans in the Making of Texas, 1836-1986*
CITE_COUNT: 20
RATINGS_COUNT: 32
AVGSTAR_RATING: 3.70
REVIEWS_COUNT: 2
PUBLISHER_NAME: University of Texas Press
ARL_COUNT: 105
NON_ARL COUNT: 698
DDC: 976.4
Main class: History and geography

Table II.
Division of the set of 8,538 history books into four partitions based on the 75th percentile values for Scopus journal citation counts and Goodreads ratings

the citation counts and reader rating counts associated with the 75th percentile in the frequency distribution were used to partition the $n = 8{,}538$ dataset into groups, leading to a focused analysis on a group of 997 books – i.e., the books belonging to the top 25 per cent both in terms of citation counts ($\geqslant 4$) and in terms of reader ratings ($\geqslant 10$).

4.2 Data set caveats

When working with large data sets, particularly newly developed sets that have been amalgamated from distinct sources, certain challenges are expected in terms of data standardization and cleaning. Our first challenge was to identify and work with as many cited titles as possible that had been recorded in Scopus as "non-sourced" reference material. The term "non-sourced" simply means that the cited titles are not indexed in Scopus and, hence, did not possess an internal Scopus Identification Number (Scopus ID). First, we automatically separated the non-sourced references into two categories: one category that followed a journal referencing format and another category in which the references could potentially be books (e.g. recorded with only one title and no volume or page numbers). References that "appeared" to be for non-sourced journal articles (or other types of grey literature) were removed from the original data set, while those that appeared to be a book were retained.

With the API interface to WorldCat.org, we could then further verify whether or not the potential book titles were in fact books. If a cited title matched a title recorded in the union library catalogue, we confirmed that the referenced document was a book. Nevertheless, the API matching and download procedure (i.e. for publisher names, publishing locations, ISBNs, and OCLC accession number) was not perfect and some of the titles could have been incorrectly matched, for instance, in cases where similar fiction book titles were matched with those that were non-fiction. Scholars often lack

precision in their referencing practices; thus with the data standardizing procedure we ultimately found that it was necessary to merge title variations (see Table III).

With the final API procedure using Goodreads, many of the ISBNs that we obtained from our WorldCat.org download did not lead to a match in Goodreads. If the ISBN did not match, a secondary title search was used, and if the title was also not found, we assumed that the book was simply not registered for a public review. Goodreads is, however, a reliable platform for retrieving information about books because it recognizes that identical titles can have different ISBNs. A different ISBN is usually assigned to the same book if a new edition has been printed at a different time and if it appears in different formats (e.g. e-book, hardcover, paperback). An ISBN can also be recorded as a ten-digit number or a 13-digit number. For example, in Table II above, the book titled *European Feminisms, 1700-1950: A Political History* has both an ISBN-10 (084734194) and an ISBN-13 (9780804734202), where the latter is an algorithmic transformation of the first. The benefit of using Goodreads is that it recognizes one book title as being a source for a rating and review regardless of the number of editions that were printed. This means that the ratings that we collected for our analysis were for a particular "work" and not for a particular edition.

Finally, it is important to clarify that Goodreads tends to have a strong English-language bias, owing to the fact that it was developed in the USA. Moreover we were restricted to working with books in the most basic sense of the term. Although WorldCat.org was useful for obtaining publisher information for every book title, it did not support a level of distinction between "fiction" and "non-fiction", nor was there a markup tag in the catalogue that would enable us to distinguish precisely if we were working with a monograph or an edited book.

5. Data analysis and results

5.1 Scholarly books rated on goodreads

Certain books that have been recognized (i.e. cited) as academic histories may be of interest to readers outside the scholarly communication system. To measure the association between academic and social impact (i.e. in terms of rater visibility on Goodreads), we calculated the Spearman's ρ correlation between the citation counts and the reader ratings for the full dataset of $n = 8{,}538$. The value obtained was a significant but weak positive correlation of 0.212 ($p < 0.01$). Overall, this means that when the citations received by *History and Geography* books (DDC = 900) in the journal literature increase, Goodreads readership ratings also have a slight tendency to

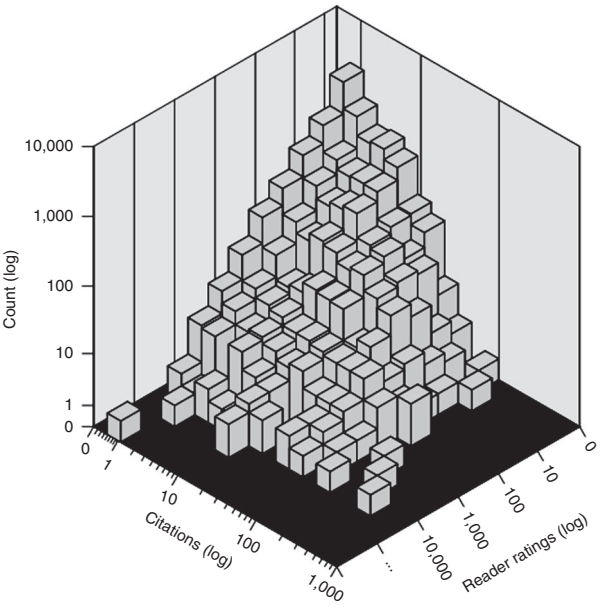
DATABASE_ID	ISBN	Title (cited variations)	Citecount
1631477	0804734194	European Feminisms	4
1631472	0804734194	European Feminisms 1700-1950	1
2136038	0804734194	European Feminisms 1700-1950. A Political History	2
5318151	0804734194	European Feminisms 1700-1950: A Political History	3
5177203	0804734194	European Feminisms, 1700-1950. A Political History	1
2150431	0804734194	European Feminisms, 1700-1950: A Political History	8
8340201	0804734194	European Feminisms, 1700-1950; A Political History	1
5519599	0804734194	European Feminisms,1700-1950: A Political History	1
MERGED_ID	*ISBN*	*Title (merged variations)*	*Citecount*
8340201	0804734194	European Feminisms, 1700-1950; A Political History	21

Table III. Example of merged variations for titles referenced in Scopus and summed citation counts

increase. With the full data set divided into groups (see Table II) a similar, but slightly weaker correlation of 0.190 ($p < 0.01$) was found specifically for the 997 books under examination.

The 3D histogram below (Figure 2), gives an idea of why these correlations are low, as it shows that most books obtain both low citations and low reader ratings. Note that the count axis is on a log scale as well (as are citations and reader ratings). The skewness of the data (see also Table I) is even stronger for the reader ratings than for the citations. No particular relation between citations and reader ratings appears from the histogram.

Figure 2 also illustrates that the groups shown in Table II should not be considered principally distinct. Rather, books may have impact within one or both of the academic and social communication systems studied, where citations are significant to the first and reader ratings are significant to the second. As shown in Figure 3, the group of books under examination fits at the intersection. On one hand, texts from this



Note: n=8,538

Figure 2.
3D histogram of the distribution of books as a function of their number of citations and reader ratings (log scales)

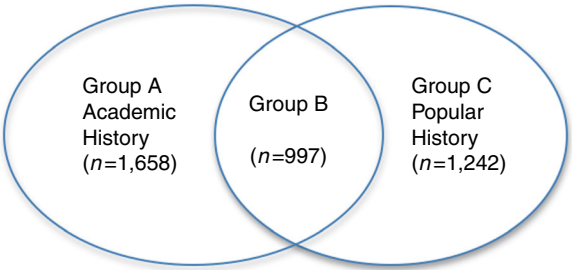


Figure 3.
Venn diagram of the books under study (n = 997) at the intersection of books with an academic impact and books with a wider impact

particular group could be characterized as "popular(ized) academic histories" – i.e., texts that were originally meant for an academic audience but have migrated into public consciousness. On the other hand, they could be popular histories that happened to have been well cited by academics. The remainder of our analyses will focus specifically on this dataset comprised of $n = 997$ observations.

5.2 Characteristics of books with high citations and high reader ratings

Further analyses using the Group B data set ($n = 997$) indicate that reader ratings accumulated on Goodreads correlate highly with written reviews (Spearman's $\rho = 0.871$) (see Table IV). This is to be expected. As ratings tend to increase so do the number of reviews, but of particular interest is the fact that ratings also correlate significantly with non-ARL library holding counts (Spearman's $\rho = 0.467$). Essentially, if a book is rated and/or reviewed often on Goodreads, it is also likely to have been accessible at libraries that were not a member of the ARL. In other words, it is present in many different types of libraries, both public and academic, worldwide.

To ascertain the degree to which the books from Group B might be characterized as academic histories or popular histories, we make use of a top-50 book publisher list for the field of history, produced by Zuccala *et al.* (2015). All book titles published by a top ranking university press (e.g. Oxford University Press) were assigned to Category 1. Those published by lower ranking university presses (i.e. not included in the top 50) were assigned to Category 2 and all additional titles published by commercial publishers were assigned to Category 3.

Note from Figure 4 that books printed by commercial publishers and cited by History journals indexed by Scopus ($\geqslant 4$ citations) were more likely than books printed by university presses to receive ratings from readers on Goodreads. A pairwise comparison of the mean reader rating counts (31 for lower ranked university presses, 131 for top ranked university presses, and 1,241 for commercial presses) shows a statistically significant difference at the 0.05 level between the latter two. Due to highly skewed distributions we also see numerous outliers for both the top ranking university press and the commercial publisher category. In terms of citations no difference is apparent, also not from the means (ten for lower ranked university presses, 16 for top ranked university presses, and 15 for commercial presses).

Table IV. Spearman's ρ correlations for Group B ($n = 997$), based on citations, Goodreads reader ratings, written reviews, library holding counts for the Association for Research Libraries (ARL) in the USA, and counts of international library holdings not affiliated with the ARL in the USA (non-ARL)

	Reader ratings	Reader reviews	ARL holdings	Non-ARL holdings
Citations				
ρ	*0.190*	*0.143*	*0.208*	*0.072*
Significance	< 0.000000	< 0.000000	< 0.000000	< 0.000000
Reader ratings				
ρ	*1.000*	*0.871*	*0.101*	*0.467*
Significance		< 0.000000	0.000669	< 0.000000
Reader reviews				
ρ		1.000	*0.072*	*0.448*
Significance			0.011453	< 0.000000
ARL holdings				
ρ			1.000	*0.566*
Significance				< 0.000000

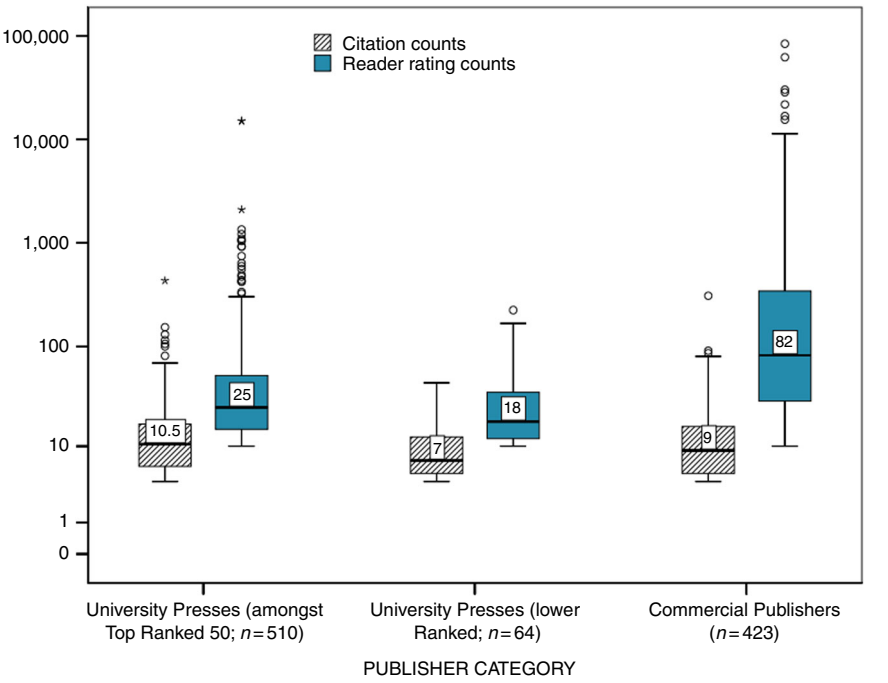


Figure 4.
Boxplots of citation counts and Goodreads reader rating counts for book titles in Group B ($n = 997$) based on publisher category

When we investigated the outliers found in the commercial category (Figure 4), two titles in particular stood apart from the rest. The first is the book *Persepolis: The Story of a Childhood*, written by Marjane Satrapi and published by Pantheon in 2000. This title had received over 80,000 Goodreads ratings at the time of our study. The second most frequently rated title, with over 60,000 Goodreads ratings, was the book written by Barak Obama titled: *Dreams from My Father: A Story of Race and Inheritance* (Three Rivers Press, New York, NY). Both books share a common theme as historical memoirs (i.e. narrative histories) and have also been recognized within the scholarly communication system (i.e. 11 and 12 citations, respectively, 2007-2011).

Amongst the outlier groups associated with the university press category (Figure 3), we have *The Histories*, also known as *The History*, or *The History of Herodotus*. This is a foundational work within the field of history, which has been published repeatedly in different editions. With respect to the University of Chicago Press edition, we found a top rating count of 15,120. The second most highly rated academic title on Goodreads, with a count of 2,105 was the work of Walker D. Howe, titled: *What Hath God Wrought: The Transformation of America, 1815-1848*, published by Oxford University Press.

At the cross-section of both the academic and social communication system certain specialties of historical research may have a greater transfer potential from academia to readers amongst the broader public. To determine whether this was the case for any of the ten DDC history and geography divisions, we compared their mean citations counts and mean reader ratings using two-sided *t*-tests. Table V presents the number of books for each of the ten DDC history and geography divisions, the mean, median, percentile 95 and maximum number of citations, and the mean, median, percentile 95, and maximum number of reader ratings. It shows that only one mean citation value, that of

Table V.
Mean, median, percentile 95, and maximum number of citations and reader ratings per DDC history and geography division

	DDC history and geography divisions									
	900 – History	910 – Geography and travel	920 – Biography and genealogy	930 – History of the ancient world (to ca. 499)	940 – History of Europe	950 – History of Asia	960 – History of Africa	970 – History of North America	980 – History of South America	990 – History of other areas
Number of books	104	44	6	59	302	108	28	312	23	11
Citations										
Mean	26	12	11	22	14	13	10	13	10	12
Median	15	8	11	10	10	7	7	9	7	11
Percentile 95	80	34	19	62	38	32	19	36	22	21
Maximum	154	40	19	433	131	309	30	80	23	21
Reader ratings										
Mean	320	324	295	872	369	1,196	1,121	703	42	35
Median	43	49	49	21	34	30	36	42	35	28
Percentile 95	1,680	927	1,295	11,326	1,399	3,570	11,387	1,877	127	72
Maximum	4,296	4,737	1,295	15,120	28,597	84,303	15,572	62,463	142	72

the main division of "History" (900) is significantly higher ($p < 0.05$) than that of several other divisions (i.e. the divisions 910, 940, 950, 960, 970, and 980). Among the mean reader ratings, there are no significant differences. Hence at least in terms of mean scholarly and social impact few differences appear within DDC History and Geography.

From Table V it appears, however, that the mean citations and reader ratings are systematically higher than the corresponding median values, especially in the case of the reader ratings. From the percentile 95 and maximum values it is clear that this is mainly because of the high outlier values found in most of the divisions. For example, within the DDC section of geography and travel (910), one of the titles that attracted a high level of readership interest was Thor Heyerdahl's first book, titled *The RA Expeditions*. This text is a historical account of the Norwegian ethnographer and adventurers voyage from north Africa to South America in a papyrus boat crafted like those illustrated in Egyptian wall paintings. At the time of our study, Heyerdahl's book had received 439 reader ratings on Goodreads, and had obtained a citation count of six in the journal literature (2007-2011).

6. Discussion

6.1 Main findings

History books cited in scholarly journal articles have an impact in academia, and may appeal to the broader public. Citation counts extracted from articles published in history journals indexed by Scopus point to the general impact of books in academia; while reader ratings and reviews on Goodreads serve as an indicator of impact beyond academia. On the basis of these two measures, we have identified a subset of history books that have both an appeal to academic historians as well as readers from the broader public. For both the whole set of 8,538 history and geography books (DDC = 900), as well as the subsection of 997 books with citation counts and reader ratings above the 75th percentile values of 4 and 10, respectively, we observed a low correlation between citations and reader ratings. In other words, the journal citation-public readership association is not particular strong. This result is in line with previous research on altmetrics (e.g. Thelwall *et al.*, 2013; Haustein *et al.*, 2014).

Further correlations between citations, reader ratings, written reviews, and library holdings indicate that a reader rating was more likely to be given to a book on Goodreads if the book was held in an international library, including both public and academic libraries. Coupling the data to a classification of publishers (Zuccala *et al.*, 2014) also shows that commercial presses have published 42 per cent of the books that have achieved both academic and social impact. This means that either academic historians are often relying on commercial publishers to print and distribute work that they think appeals to a broader readership, or that certain popular texts are accepted as relevant sources within the scholarly communication system. Lastly, we show that in terms of citations and reader ratings there are few differences between history books that belong to different DDC divisions.

6.2 Limitations

It is important to indicate a few limitations to using Goodreads as an altmetric data source. The first limitation relates to the fact that it is primarily a platform used by English-speaking readers. At present this corresponds well with the English-language bias we see in academic journals indexed by Scopus. This means that the correlations made between citations to books in journals and reader ratings to the same books on Goodreads may not indicative of a broad section of cultures and languages.

The second drawback to Goodreads is that it was not possible to determine if the public's recognition of many of the books came prior to or after other forms of public appreciation, for example, news announcements of special awards. If the social interest amongst readers began with Goodreads then we might conclude that the impetus for public interest was the nature of the book itself. If not, other forms of social media (e.g. news articles) may have motivated readers to appreciate and critique certain history books.

A third drawback relates to the nature of this study, which involved isolating book titles and standardizing book and publisher metadata. Since it was not feasible to clean the data perfectly in terms of all titles, publication dates, and authors, we did not focus too precisely on constructing a controlled time frame for the analyses. Note that in previous studies, like that of Thelwall *et al.* (2013) a "sign" test was introduced to be sure that journal articles had been exposed to the same citation delay and usage uptake biases on other forms of social media (e.g. Twitter). When comparing journal citations to reader ratings we may have also introduced a test that would normalize for different uptake periods. Scholarly readers and Goodreads readers both require time to read a history book, but a Goodreads star rating is submitted on the basis of a quick decision, while the citation must be processed through peer review and lengthy publishing procedures before it is used as piece of "metric" evidence.

One final limitation relates to our method of using the DDC for classifying books in the broad field of history. Despite the fact that the DDC is one of the most widely used method of organizing books in the English-speaking world, the use of other classification systems, such as the Library of Congress Classification (LCC) or the Universal Decimal Classification (UDC), could possibly yield additional information. Compared to other subjects that have been classified, however, these three most commonly used classification systems have been assessed as being adequate, complete, and systematic for the field of history (Zins and Santos, 2011).

6.3 Future research

Readers registered on Goodreads are much more likely to give a star rating to a history book than spend time writing a review (see Table I). To understand how readers have criticized, appreciated, or recommended History books on Goodreads, it would be useful to carry out opinion-based analyses of full-text reviews. Trott and Naik (2012) previously examined reviews on Goodreads and found that participants "employ a wide range of terms in discussing the appeal of books" and often suggest "possible read-alikes" (p. 321). A critical facet of reader interactions is the trust factor: persons "who seem to know and trust the reviewer or other commentators are more likely to be swayed into stating that they will read the book" (p. 321). Exploiting the integration of Goodreads with Facebook might also help to confirm the concept of "lay" reader, which is difficult to establish given that both academics and non-academic (lay) persons alike make use of social networking platforms.

Zuccala *et al.* (2014) previously examined reviews that were published for academic historians in the journal *American Historical Review*. Here, scholarly reviewing was observed as a specific type of process in which scholars were invited to formulate critical assessments of the writing style and academic credibility of colleagues who had published a new book. With Goodreads reviews, a similar linguistic technique of analyses might be used in a new investigation concerning the social impact of historical literature. It is not part of the scope of this paper to employ the same methods as Zuccala *et al.* (2014); however, a future approach might be designed to compare

linguistic differences and similarities between academic reviews and public reviews. A comparative analysis of reviews from the two realms might also provide deeper insight into levels of public interest.

7. Conclusion

Goodreads needs to be examined further with respect to other academic subject areas; however, the field of history has been a convincing starting point for investigating how scholarly books are received and evaluated beyond academia. For academics, Goodreads possesses its greatest potential as an altmetric research tool if a book that has been received well within the academic community as scholarly, accurate, and thought provoking piece of work can also attract substantial public interest. There are many forms of social impact, but as this study shows, Goodreads could make a strong contribution to a complementary approach to metric evaluations, particularly with the use of citations and reader ratings together. As such, the Goodreads book rating and reviewing system definitely belongs to the realm of altmetrics, where research is now continuing with an ever-expanding inventory of social media platforms.

References

Allen, N. and Heath, O. (2013), "Reputations and research quality in British political science: the importance of journal and publisher rankings in the 2008 RAE", *The British Journal of Politics and International Relations*, Vol. 15 No. 1, pp. 147-162.

Amazon (2015), "History", available at: www.amazon.com/History-Books/b?node = 9 (accessed 12 March 2015).

Ankersmit, F. (2005), *Sublime Historical Experience*, Standford University Press, Stanford, CA.

Barré, R. (2005), "S&T indicators for policy making in a changing science-society relationship", in Moed, H., Glänzel, W. and Schmoch, U. (Eds), *Handbook of Quantitative Science and Technology Research*, Springer, Dordrecht, pp. 115-131.

Bornmann, L. (2013), "What is societal impact of research and how can it be assessed? A literature survey", *Journal of the American Society of Information Science and Technology*, Vol. 64 No. 2, pp. 217-233.

Bornmann, L. (2014), "Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics", *Journal of Informetrics*, Vol. 8 No. 4, pp. 895-903.

Bornmann, L. (2015), "Usefulness of altmetrics for measuring the broader impact of research: A case study using data from PLOS and F1000Prime", *Aslib Journal of Information Management*, Vol. 67 No. 3, pp. XX-XX.

Cronin, B. and La Barre, K. (2004), "Mickey mouse and milton: book publishing in the humanities", *Learned Publishing*, Vol. 17 No. 2, pp. 85-98.

Dalton, M.S. and Charnigo, L. (2004), "Historians and their information sources", *College & Research Libraries*, Vol. 65 No. 5, pp. 400-425.

Den Hertog, P., Jager., C.-J., Vankan, A., Te Velde, R., Veldkamp, J., Aksnes, D., Sivertsen, G., Van Leeuwen, T. and Van Wijk, E. (2014), "Science, technology & innovation indicators. Thematic paper 2: scholarly publication patterns in the social sciences and humanities and their relationship with research assessment", available at: http://dialogic.nl/documents/other/sti2_themepaper2.pdf (accessed 9 March 2015).

Donovan, C. (2007), "The qualitative future of research evaluation", *Science and Public Policy*, Vol. 34 No. 8, pp. 565-574.

Donovan, C. (2008), "The Australian research quality framework: a live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research", in Coryn, C.L.S. and Scriven, M. (Eds), *Reforming the Evaluation of Research. New Directions for Evaluation*, , Vol. 118, Jossey-Bass, San Francisco, CA, pp. 47-60.

Engels, T.C.E., Ossenblok, T.L.B. and Spruyt, E.H.J. (2012), "Changing publication patterns in the social sciences and humanities, 2000-2009", *Scientometrics*, Vol. 93 No. 2, pp. 373-390.

Finkelstein, N.G. and Birn, R.B. (1998), *A Nation on Trial: The Goldhagen Thesis and Historical Truth*, Henry Holt & Company Inc., New York, NY.

Goodreads (2015), "About Goodreads", available at: www.goodreads.com/about/us (accessed 12 March 2015).

Hammarfelt, B. (2014), "Using altmetrics for assessing research impact in the humanities", *Scientometrics*, Vol. 101 No. 2, pp. 1419-1430.

Haustein, S. and Siebenlist, T. (2011), "Applying social bookmarking data to evaluate journal usage", *Journal of Informetrics*, Vol. 5 No. 3, pp. 446-457.

Haustein, S., Peters, I., Sugimoto, C.R., Thelwall, M. and Larivière, V. (2014), "Tweeting biomedicine: an analysis of tweets and citations in the biomedical literature", *Journal of the Association for Information Science and Technology*, Vol. 65 No. 4, pp. 656-669.

Holbrook, J.B. and Frodeman, R. (2011), "Peer review and the exante assessment of societal impacts", *Research Evaluation*, Vol. 20 No. 3, pp. 239-246.

Jones, C., Chapman, M. and Carr Woods, P. (1972), "The characteristics of the literature used by historians", *Journal of Librarianship*, Vol. 4 No. 3, pp. 137-156.

Li, X., Thelwall, M. and Giustini, D. (2012), "Validating online reference managers for scholarly impact measurement", *Scientometrics*, Vol. 91 No. 2, pp. 461-471.

Moed, H.F. (2005), *Citation Analysis in Research Evaluation*, Springer, Dordrecht.

Nielsen, F. (2007), "Scientific citations in wikipedia", *First Monday*, Vol. 12 No. 8, doi:10.5210/fm.v12i8.1997, available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/indexphp/fm/article/view/1997/1872 (accessed 8 January 2014).

Nightingale, P. and Scott, A. (2007), "Peer review and the relevance gap: ten suggestions for policy-makers", *Science and Public Policy*, Vol. 34 No. 8, pp. 543-553.

Nowotny, H., Scott, P. and Gibbons, M. (2003), "Introduction: 'Mode 2' revisited: the new production of knowledge", *Minerva*, Vol. 41 No. 3, pp. 179-194.

Petit, J.C. (2004), "Why do we need fundamental research?", *European Review*, Vol. 12 No. 2, pp. 191-207.

Poe, M. (2014), "New books in history. Discussions with historians about their new books", available at: http://newbooksinhistory.com/ (accessed 17 January 2014).

Priem, J., Costello, K. and Dzuba, T. (2012a), "Prevalence and use of Twitter among scholars", *Figshare*, available at: http://dx.doi.org/10.6084/m9.figshare.104629 (accessed 10 January 2014).

Priem, J., Taraborelli, D., Groth, P. and Neylon, C. (2010), "Alt-metrics: a manifesto", available at: http://altmetrics.org/manifesto/ (accessed 7 January 2014).

Reid, A. and Szreter, S. (2008), "Making history. The changing face of the profession in Britain. History & policy", available at: www.history.ac.uk/makinghistory /resources/articles/history_and_policy.html (accessed 2 December 2013).

Shema, H., Bar Illan, J. and Thelwall, M. (2014), "Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics?", *Journal of the American Society for Information Science and Technology*, Vol. 65 No. 5, pp. 1018-1027.

Thelwall, M., Haustein, S., Larivière, V., Sugimoto, C.R. (2013), "Do altmetrics work? Twitter and ten other social web services", *PLoS ONE*, Vol. 8 No. 5, p. e64841.

Thompson, J.B. (2005), *Books in the Digital Age. The Transformation of Academic and Higher Education Publishing in Britain and the United States*, Polity Press, Cambridge.

Tollebeek, J. and Verschaffel, T. (1992), *De Vreugden Van Houssaye. Apologie Van de Historische Interesse*, Wereldbibliotheek, Amsterdam, NL.

Tosh, J. (2008), *Why History Matters*, Palgrave Macmillan, London.

Trott, B. and Naik, Y. (2012), "Finding good reads on goodreads. Readers take RA into their own hands", *Reference & User Services Quarterly*, Vol. 51 No. 4, pp. 319-323.

Verleysen, F.T. and Engels, T.C.E. (2012), "Historical publications at flemish universities, 2000-2009", *Journal of Belgian History – Revue Belge d'Histoire Contemporaine – Belgisch Tijdschrift voor Nieuwste Geschiedenis*, Vol. 42 No. 4, pp. 110-143.

Walker, R. (2011), "Conceptualizing the social impact of science (Conceptualizando el impacto social de la ciencia)", available at: http://sisobproject.wordpress.com/2011/07/20/conceptualizing-the-social-impact-of-science-2/ (accessed 14 January 2014).

Weller, K., Dröge, E. and Puschmann, C. (2011), "Citation analysis in Twitter: approaches for defining and measuring information flows within tweets during scientific conferences", in Rowe, M., Stankovic, M., Dadzie, A.-S. and Hardey, M. (Eds), *Making Sense of Microposts (#MSM2011), Workshop at the Extended Semantic Web Conference(ESWC 2011), Heraklion*, pp. 1-12, available at: http://files.ynada.com/papers/msm2011.pdf (accessed 7 January 2014).

White, H., Boell, S.K., Yu, H., Davis, M., Wilson, C.S. and Cole, F.T.H. 2 (2009), "Libcitations: a measure for comparative assessment of book publications in the humanities and social sciences", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 6, pp. 1083-1096.

WorldCat (2014), "Find items in libraries near you. 2 billion items available here through a library", available at: www.WorldCat.org/ (accessed 17 January 2014).

Zins, C. and Santos, P.L.V.A.C. (2011), "Mapping the knowledge covered by library classification systems", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 5, pp. 877-901.

Zuccala, A., van Someren, M. and van Bellen, M. (2014), "A machine-learning approach to coding book reviews as quality indicators: towards a theory of mega-citation", *Journal of the American Society for Information Science and Technology*, Vol. 65 No. 11, pp. 2248-2260.

Zuccala, A., Guns, R., Cornacchia, R. and Bod, R. (2015), "Can we rank scholarly book publishers? A bibliometric experiment with the field of history", *Journal of the American Society for Information Science and Technology*, doi:10.1002/asi.23267 (in press).

Further reading

Priem, J., Piwowar, H.A. and Hemminger, B.M. (2012b), "Altmetrics in the wild: using social media to explore scholarly impact", available at: http://arxiv.org/abs/ 1203.4745 (accessed 8 January 2014).

Corresponding author

Dr Alesia A. Zuccala can be contacted at: spl465@iva.ku.dk

This article has been cited by:

1. DesrochersNadine Nadine Desrochers Nadine Desrochers is an Assistant Professor at the Université de Montréal's LIS school (EBSI). She holds an MLIS from the Western University and a PhD in French Literature from the University of Ottawa. Her research examines the paratext and perceptions surrounding cultural products, readers' advisory, and the information-seeking habits of creators. Her work has been published in Information Research, the Canadian Journal of Information and Library Science, and Library & Information Science Research. She is the Co-Editor of the collective book Examining Paratextual Theory and its Applications in Digital Culture (with Daniel Apollon; IGI Global, 2014). LaplanteAudrey Audrey Laplante Audrey Laplante is an Associate Professor at the Université de Montréal's LIS school (EBSI). She is a Member of the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT). She received a PhD (2008) and a Master's (2001) in Information Science from the McGill University and Université de Montréal, respectively. Her research interests focus primarily on the information-seeking behaviour of music researchers and amateurs, and on systems for music information retrieval and discovery. Her research has been published in a variety of outlets, including Library & Information Science Research, the Proceedings of the International Society for Music Information Retrieval, and the collective book New Directions in Children's and Adolescents' Information Behavior Research (Emerald Group Publishing, 2014). MartinKim Kim Martin Kim Martin is a PhD Candidate in Library and Information Science at the Western University. Her research interests include the use of social media by humanities scholars and the role of serendipity in the research process of the historian. Working in both academia and the wider public sphere to promote humanities education, she works hard to maintain an interdisciplinary approach throughout her doctoral work by participating in the growing world of the digital humanities. Quan-HaaseAnabel Anabel Quan-Haase Anabel Quan-Haase is an Associate Professor of Information and Media Studies and Sociology, Western University. She is the past President of the Canadian Association of Information Science and a Council Member of CITAMS, a section of the American Sociological Association. She is the Author of Technology and Society: Social Networks, Power and Inequality (Oxford University Press, 2013; 2nd ed., 2016). Her research focuses on social media, social networks, social capital, and innovation. SpiteriLouise Louise Spiteri Louise Spiteri is an Associate Professor at the School of Information Management, Dalhousie University. Dr Spiteri holds an MLIS from the Western University, and PhD in Information Studies from the University of Toronto. Dr Spiteri's research examines the contribution of social tagging and user-generated content to information retrieval systems and library catalogues. Dr Spiteri's research has been published in a variety of journals, including Journal of Documentation, Cataloging & Classification Quarterly, and Journal of Information Science. École de bibliothéconomie et des sciences de l'information, Université de Montréal, Montréal, Canada Faculty of Information and Media Studies, University of Western Ontario, London, Canada School of Information Management, Dalhousie University, Halifax, Canada . 2016. Illusions of a "Bond": tagging cultural products across online platforms. *Journal of Documentation* **72**:6, 1027-1051. [Abstract] [Full Text] [PDF]