Why Do Papers Have Many Mendeley Readers But Few Scopus-Indexed Citations and Vice Versa?[1]

Mike Thelwall
Statistical Cybermetrics Research Group, School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK.
Tel. +44 1902 321470. Fax +44 1902 321478, Email: m.thelwall@wlv.ac.uk

Abstract
Counts of citations to academic articles are widely used as indicators of their scholarly impact. In addition, alternative indicators derived from social websites have been proposed to cover some of the shortcomings of citation counts. The most promising such indicator is counts of readers of an article in the social reference sharing site Mendeley. Although Mendeley reader counts tend to correlate strongly and positively with citation counts within scientific fields, an understanding of causes of citation-reader anomalies is needed before Mendeley reader counts can be used with confidence as indicators. In response, this article proposes a list reasons for anomalies based upon an analysis of articles that are highly cited but have few Mendeley readers, or vice versa. The results show that there are both technical and legitimate reasons for differences, with the latter including communities that use research but do not cite it in Scopus-indexed publications or do not use Mendeley. The results also suggest that the lower of the two values (citation counts, reader counts) tends to underestimate of the impact of an article and so taking the maximum is a reasonable strategy for a combined impact indicator.
Keywords: Mendeley; altmetrics; citation analysis

Introduction

Although citation counts are widely used in formal and formal research evaluations as indicators of scholarly impact, they have many limitations. In particular, citations take time to accrue whilst follow-up research is conducted and published, and citations from the academic literature may not reflect an article's non-academic impact. In response to the latter point, a range of alternative indicators have been proposed to supplement citation counts. These include patent citations as indicators of commercial value (Meyer, 2000; Trajtenberg, 1990), web citations or tweets as evidence of wider or public interest (Eysenbach, 2011) and syllabus mentions as evidence of educational impact (Kousha & Thelwall, 2008). An important limitation of most indicators derived from the web is the ease with which they can be manipulated, making them unsafe for most formal evaluations (Wouters & Costas, 2012). Although many early alternative indicators are difficult to calculate in practice for large sets of articles, those generated from social websites, such as Twitter, can often be calculated automatically on a large scale using Applications Programming Interfaces (APIs) that allow automated retrieval (Priem, Taraborelli, Groth, & Neylon, 2011). Many of these new indicators correlate positively and significantly with academic citations to articles, but typically with low correlation coefficients (Costas, Zahedi, & Wouters, in press; Thelwall Haustein Larivière Sugimoto 2013). Other analytical strategies are also needed to more fully evaluate these indicators, however (Sud & Thelwall, 2014).

[1] Journal of Librarianship & Information Science, to appear

One particularly promising new indicator is Mendeley reader counts (Gunn, 2013; Haustein & Siebenlist, 2011). Although counts of readers of articles in the social reference sharing site Mendeley (Henning & Reichelt, 2008) seem to predominantly reflect scholarly impact, and hence largely duplicate rather than supplement citation counts, they appear earlier than citations (Maflahi & Thelwall, in press) because they are not affected by publication delays and so have value as early impact indicators. This characteristic is particularly important for decisions relating to emerging research areas or recent research.

Many studies have shown that counts of Mendeley readers correlate with citation counts for individual journal articles within a field, whichever field is analysed (Bar-Ilan, 2012; Haustein, Larivière, Thelwall, Amyot, & Peters, 2014; Li & Thelwall, 2012; Li, Thelwall, & Giustini, 2012; Thelwall, & Wilson, in press; Zahedi, Costas, & Wouters, 2014). Mendeley users tend to be younger than average and include a small proportion of master's and bachelor's degree students (less than 20% - perhaps much less due to ambiguity in Mendeley's user category names) that presumably do not publish, as well as professional users such as medical doctors and librarians (Mohammadi, Thelwall, Haustein, & Larivière, in press). Reflecting this, eminent senior researchers seem to rarely use Mendeley (Mas-Bleda, Thelwall, Kousha, & Aguillo, 2014). Those that do use Mendeley may register articles in it and are likely to read them in the future or to have read them before registering them (Mohammadi, Thelwall, & Kousha, in press). It is therefore reasonable to consider Mendeley as a source of predominantly academic readership information with a bias towards younger academics and educational uses. For convenience, the term *reader* is used in the remainder of this article to refer to people that register an article in Mendeley even though not all registered articles are read by the person that registered them.

In order to have confidence about the value of Mendeley as a scholarly impact indicator it is important to understand the contexts in which it does not work. Thus, the current article reports a study of individual articles for which their Mendeley reader count does not correspond to their Scopus citation count (i.e., the number of citations found by Scopus within Scopus-indexed publications). The objective is to produce a list of reasons why articles may be extensively read in Mendeley but rarely cited in Scopus-indexed publications and vice versa. This information can help users of Mendeley reader counts to look for individual problematic cases and to understand the limitations of Mendeley reader counts as a scholarly impact indicator. Scopus was chosen in preference to the Web of Science for its greater coverage of the academic literature (Moed & Visser, 2008; Torres-Salinas, Lopez-Cózar, & Jiménez-Contreras, 2009) – for example, 97% of Web of Science publications from 2005 are also in Scopus (Moed & Visser, 2008). Google Scholar probably has wider coverage in general (Bar-Ilan, 2008; Harzing, 2014) but it is not possible to use it in large scale bibliometric studies due to the lack of facilities for providing citation counts for extensive sets of articles as well as the lack of a field classification scheme. The following specific research questions drive the study.

- Why are articles cited in Scopus-indexed publications more often than they are registered by Mendeley users?
- Why are articles registered by Mendeley users more often than they are cited in Scopus-indexed publications?

The goal is not to provide a comprehensive list of reasons in answer to the above research questions, nor to estimate the prevalence of the reasons found in any particular research area or within academia as a whole, but only to identify a set of reasons for discrepancies between reader counts and citation counts and to give evidence for each reason.

Methods

The overall research design was to identify a collection of articles that had disproportionately many or few Mendeley readers in comparison to their Scopus-indexed citation counts and then to examine these articles and produce a list of reasons for the differences found. Since reasons for the discrepancies may vary by field, fifteen different subject areas were chosen (Horticulture; Music; Molecular Medicine; Organic Chemistry; Computer Vision; Oceanography; Nuclear Energy and Engineering; Health, Toxicology and Mutagenesis; Parasitology; Health Policy; Biological Psychiatry; Instrumentation; Education; Small Animals; Pharmacy). These subjects represent a wide variety of different types of research. Each subject area was based upon an individual Scopus category. Although Scopus categories have limitations, they seem to be sufficiently coherent to be analysed separately.

For each of the fifteen selected Scopus categories, articles from 2011 were downloaded in February 2015, up to a maximum of 10,000. The year 2011 was selected to give enough time for articles to attract substantial numbers of citations and readers. The limit of 10,000 articles is a technical limitation but is large enough for the analysis for each subject area. For subjects with more than 10,000 articles published in 2011, the missing articles are those published in the middle of the year.

For each article downloaded from Scopus in each category, the Mendeley Applications Programming Interface (API) version 1 was used via the free software Webometric Analyst to download the number of readers for the article in February 2015 (for more details see: Thelwall & Wilson, in press). Each Scopus article was matched with corresponding articles in Mendeley using a DOI search, when a DOI was registered in Scopus, and with a combined title, first author and year search in Mendeley (for all articles, irrespective of whether a DOI was present or not). The following query shows the format used.

```
title:The Prevalence of Microalbuminuria and Proteinuria in Cats with Diabetes Mellitus AND author:Al-Ghazlat AND year:2011
```

All results were automatically checked for correct matches first by matching the DOIs, when present, and then by matching the year and title words to ensure a high degree of similarity in order to account for minor typographic errors without adding false matches (for more details see: Thelwall & Wilson, in press). The results for all articles that were judged to be matches were totalled. These results include the number of readers, the occupation of each reader (e.g., Doctoral student, Professor), their subject area of interest (e.g., Computer Science) and sometimes also their country of origin. This is self-reported information from Mendeley users.

The next stage was to identify articles with relatively many or few Mendeley readers compared to Scopus citations. For this, linear regression was used to regress reader counts against citation counts for the articles in each subject area in order to estimate the expected number of Mendeley readers for each article based upon its citation count. Although linear regression is unsuited to citation and readership data because it is highly skewed (Price, 1965; Thelwall & Wilson, 2014a) and the data should be given a logarithmic transformation before applying the regression in order to get a statistically robust regression fit (Thelwall & Wilson, 2014b), any transformation would affect the nature of the outliers to be examined. No transformation was used on the basis that it would be most useful to detect outliers based on the untransformed data so that an article would be an outlier based on a straightforward comparison of its numbers of readers and citations.

There are many different popular methods to detect outliers in a data set. The simplest is probably the residual: the difference between the number of readers of an article and the number of readers estimated by its number of citations using regression. This is a reasonable method to detect outliers. Its main limitation is that it is biased towards large numbers in the sense that the residuals of the same size are probably inherently more important for articles with fewer citations because the ratio of the number of readers to citations would be higher. The Studentised variant of the residual was used, which calculates residuals after regressing against all data except the point considered. This variant should be more sensitive to genuine outliers.

For each subject area, the ten articles with the largest positive Studentised residuals and the ten articles with the largest negative Studentised residuals were selected for manual analysis, a total of 300 articles. This was repeated after a logarithmic transformation of the reader and citer counts with log(1+x), for example changing 5 readers to log(1+5)=log(6)=1.79 as recommended above and to reduce the impact of very high counts of readers or citations (Thelwall & Wilson, 2014b), giving a second set of 300 articles in order to catch additional outliers using a second reasonable method.

The characteristics of each article in both sets (414 in total due to overlaps) were investigated qualitatively by the author in order to detect a reason why it may have been relatively highly read or cited. The characteristics examined included the article title, journal name and author and the available reader attributes (e.g., nationality, reader occupations, reader subject area). Google Scholar was also used to cross-check the citation counts reported by Scopus in cases of low citation counts. The articles were examined both individually and collectively in order to identify patterns that would explain the results. The reasons identified were then generalised and grouped together into a list, with evidence given for each one in order to justify its inclusion. This is essentially a grounded theory approach that is reliant upon the investigator to detect patterns with the limited objective to identify plausible reasons for outliers.

Results

Table 1 summarises key information about each field analysed to give context to the main results. It confirms that the fields analysed are very different, with substantial differences in the mean number of citations per article (from 2.25 to 11.95), the mean number of readers per article (from 3.61 to 16.64) and the ratio of readers to citations (from 0.68 to 2.73 readers per citation). Reader counts correlate strongly (0.567 to 0.691) with citation counts in all subject categories except for Computer Vision (0.200). Spearman correlations were used to estimate the degree of dependence between reader counts and citation counts because the data was too skewed to use Pearson correlations. The reason for the low Computer Vision correlation might be a journal with few readers but many citations (*Journal of Information Hiding and Multimedia Signal Processing*) and several IEEE journals and conference proceedings with a high ratio of reader counts to citation counts (e.g., *IEEE Transactions on Pattern Analysis and Machine Intelligence*). These are perhaps peripheral to the subject area and hence may have a different citation/reader profile. More importantly, however, the *Annual International Conference of the IEEE Engineering in Medicine and Biology – Proceedings* has an ISSN (0589-1019) is classified as a journal in Scopus, hence its presence in the data set. It contributes 2066 (37%) of the Computer Vision articles, many of which are uncited. These articles have 30914 readers and 1907 citations, a ratio of 16.2 to 1. This extremely high ratio seems to be due to the articles occurring twice in Scopus, once as a journal article and once as a conference paper, with the conference paper version tending to be credited with the citations.

Table 1. Spearman correlations between Mendeley reader counts and citation counts for articles in Scopus from 2011 in the fifteen subjects analysed.

Scopus subject category	**Articles**	**Citation mean (median)**	**Reader mean (median)**	**Spearman correlation**
Biological Psychiatry	3214	11.95(8)	16.14(11)	0.679
Computer Vision	5620	6.07(2)	13.00(10)	0.200
Education	9950	3.16(1)	8.65(4)	0.593
Health Policy	9020	5.09(3)	8.75(6)	0.664
Health, Toxicology and Mutagenesis	9861	8.88(5)	7.86(5)	0.568
Horticulture	4539	4.63(3)	4.90(3)	0.604
Instrumentation	9993	5.22(3)	3.76(2)	0.567
Molecular Medicine	10000	11.19(7)	10.32(6)	0.570
Music	1364	1.51(0)	3.89(1)	0.618
Nuclear Energy and Engineering	7330	7.81(2)	5.61(2)	0.691
Oceanography	6814	7.17(5)	12.54(8)	0.671
Organic Chemistry	10000	9.10(6)	6.22(4)	0.605
Parasitology	6535	8.52(5)	10.79(7)	0.652
Pharmacy	727	2.25(1)	3.61(1)	0.648
Small Animals	1195	3.50(2)	8.84(6)	0.638

The outliers identified by the two different methods tend to be similar but the logarithmic method tends to produce fewer outliers with high values on both indicators (figures 1 and 2 - random jitter has been added to the reader counts and citation counts to avoid identical points overlapping).

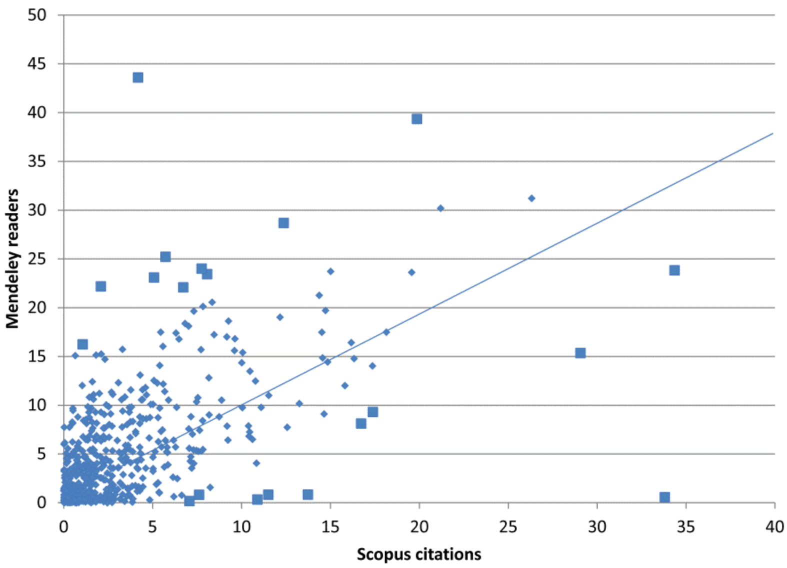

Figure 1. Reader counts compared to citation counts for Pharmacy articles. The top ten Studentised residuals according to a standard regression are marked with a square shape.

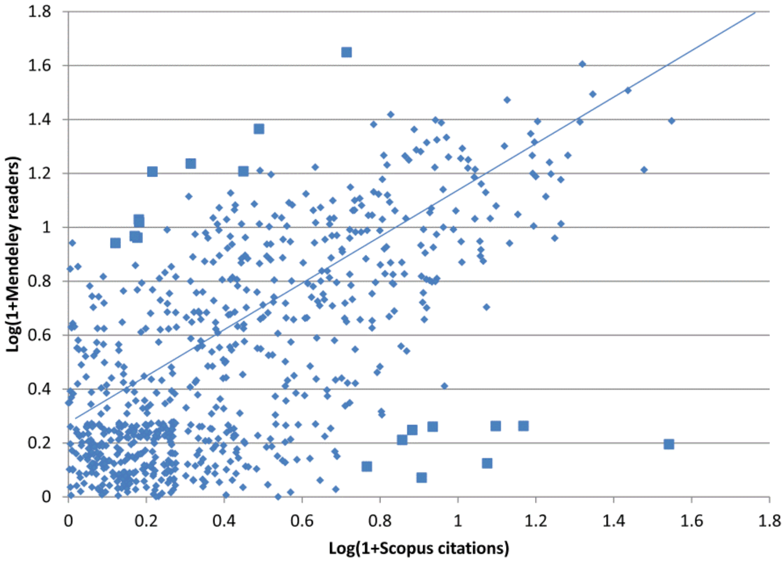

Figure 2. Reader counts compared to citation counts for Pharmacy articles after applying a log(1+x) transformation to both. The top ten Studentised residuals based on this data are marked with a square shape.

The following reasons were found for an article having many Mendeley readers compared to its number of Scopus-indexed citations.

- Articles may attract disproportionately many students compared to academics, with students not producing work that is indexed in Scopus. In Horticulture, Batchelor's degree and Master's degree students formed 21% of readers overall (a high proportion in comparison to other areas) but 38% of the readers of "The role of RFID in agriculture: Applications, limitations and challenges". The article, "Early microbial biofilm formation on marine plastic debris" attracted 26% undergraduate readers compared to an Oceanography average of 6%. In Health, Toxicology and Mutagenesis, "Positive effects of vegetation: Urban heat island and green roofs" attracted 40% student readers compared to a subject average of 23%. In Music, "The implicit body as performance: Analyzing interactive art" attracted 47% Master's student readers compared to a subject average of 18%.
- Articles may attract disproportionately many professional readers that may not author articles. "Size matters: Management of stress responses and chronic stress in beaked whales and other marine mammals may require larger exclusion zones" attracted 15% Other Professional readers compared to an Oceanography average of 6%. "Improving studies of resource selection by understanding resource use" attracted 13% Other Professional readers compared to a Health, Toxicology and Mutagenesis average of 4%. "Canine intracranial neoplasia: Clinical risk factors for development of epileptic seizures" attracted 30% Other Professional readers compared to a Small Animals average of 16%.
- Articles may be multidisciplinary and attract many readers from an area that has a field norm of more readers per citation, despite being legitimately labelled with another category. In Music, four of the articles with relatively many readers and none of the articles with relatively many citations were from the journal *Psychology of Music*. In Computer Vision, six of the disproportionately read articles were from *IEEE Transactions on Pattern Analysis and Machine Intelligence*, and only one of the

disproportionately cited articles. In Nuclear Energy and Engineering, *Applied Energy* had 9 of the most read articles.

- Magazine-style journals may attract high numbers of article readers, perhaps reflecting a more casual typical reader. The top ten articles for relatively high readership in Molecular Medicine were all from *Nature Biotechnology.*
- Review articles may attract disproportionately many readers because some authors cite the reviewed articles rather than the review. *Annual Review of Marine Science* published five of the top read Oceanography articles and none of the top cited articles.
- Academic articles may attract general interest readers from the academic community that do not use the article directly in their work. "Is free will an illusion?" within Health Policy attracted 233 readers but no citations.
- Articles may support the research process without being citable. The short article, "Four principles to help non-native speakers of English write clearly" in *Fisheries Oceanography* had 181 readers but no Scopus citations.
- Article readers may be predominantly from countries with researchers that rarely publish in Scopus journals. In Oceanography, "Coral bleaching and habitat effects on colonisation of reef fish assemblages: An experimental study" by Swedish and Tanzanian authors had no Scopus citations and only five citations in Google Scholar (four reports and one book chapter) despite having 35 Mendeley readers, none of whom were from a major publishing nation (although only four registered a country of origin).

The following reasons were found for articles having many Scopus citations compared to their Mendeley readers.

- Articles may interest an academic community that does not use Mendeley due to limited internet access. All ten Horticulture articles with relatively few readers were about folk medicine in Bangladesh (e.g., "Use of inorganic substances in folk medicinal formulations: A case study of a folk medicinal practitioner in Tangail district, Bangladesh" had 58 citations but no readers) and were mainly cited by other articles about folk medicine in Bangladesh written by Bangladeshi scholars.
- Articles may be mainly of interest to a publishing author community, such as hospital doctors, that does not use Mendeley due to working practices. Two articles with similar titles (e.g., "National Healthcare Safety Network (NHSN) Report, data summary for 2010, device-associated module") in Health Policy had 110 and 48 citations but no readers. The citations in Scopus appeared to be mainly from research published by authors with hospital affiliations. Internet security in hospitals for privacy reasons may make casual internet access difficult.
- Articles may be multidisciplinary and attract many citations from one side of their focus, which has a field norm of few readers per citer, despite being legitimately labelled with another category. In Nuclear Energy and Engineering, *Energy and Environmental Science* had 8 of the least read articles.
- Mendeley users may register as readers of an update of an article rather than the original version, whilst authors may cite the original. This occurred for "Gene therapy finds its niche" and "Haplotype-resolved genome sequencing of a Gujarati Indian individual", both in Molecular Medicine. It also occurred for, '"Doctor, would you prescribe a pill to help me … ?" a national survey of physicians on using medicine for human enhancement', because Mendeley linked to the article, 'Response to Open

Peer Commentaries on "'Doctor, Would You Prescribe a Pill to Help Me … ?' A National Survey of Physicians on Using Medicine for Human Enhancement"'.

The following reasons for outliers were found that are technical limitations of the process used.

- An irrelevant journal had been included in the Scopus category, and articles in this journal may have been from a field with a different normal ratio of reader counts to citation counts. This applied to *Academic Medicine* and *Resonance* within Education and to *Experimental Diabetes Research* within Music.
- An article may lack a DOI and may have a title that is not searchable in Mendeley. The article "Applications of objective image quality assessment methods [Applications Corner]" is in Mendeley but is not returned by a search for it, with or without the square brackets. *Morbidity and Mortality Weekly Report* had all of the ten least read articles in Health, Toxicology and Mutagenesis. These were cited an average of 177.5 times and read an average of 0.8 times each, compared to subject averages of 8.9 and 15.7 respectively. These articles had titles that, in their Scopus format, were difficult to search for in Mendeley due to the inclusion of double hyphens, as in "Vital signs: Overdoses of prescription opioid pain relievers --- United States, 1999--2008 ". The underlying cause seems to be the use of long hyphens in the source journal.
- An article may have an incorrect DOI in Scopus. This applied to "Trophic cycling and carbon export relationships in the California current ecosystem" in Oceanography, which had a DOI of 10.3354/meps010257, which was highly read, but the correct DOI is 10.4319/lo.2011.56.5.1866. "Study to Assess Health Effects of Railyard - Elders Volunteer for Fukushima Daiichi Plant Cleanup" in Health Policy also had an incorrect DOI (10.1289/ehp.119-a289) in Scopus.
- An article may have an incorrect DOI in Mendeley. This applied to "Understanding and projecting Level Change" and "A complete high-resolution Paleocene benthic stable isotope record for the central Pacific (ODP Site 1209)" in Oceanography.
- An article may have an incorrect title in Mendeley. "Transnational caregiving: Part 2, caring for family relations across nations" in Health Policy had a different part number (1) in Mendeley.
- An article may have an incorrect year in Mendeley. "Methemoglobinemia and hemolytic anemia after rasburicase administration in a child with leukemia" within Pharmacy is registered as published in 2013 instead of 2011 in Mendeley.
- An article may have multiple valid DOIs, with different versions in Mendeley and Scopus. This applied to "Preliminary estimation of release amounts of 131I and 137 Cs accidentally discharged from the Fukushima Daiichi Nuclear power plant into the atmosphere" and "JENDL-4.0: A new library for nuclear science and engineering" within Nuclear Energy and Engineering.
- An article may appear more than once in Scopus, reducing the Scopus-indexed citation count of both versions. This occurred for "The neural basis of video gaming" within Biological Psychiatry.
- A non-citable item may be incorrectly classified as a journal article in Scopus. The podcast, "Climate change, crop yields, and undernutrition, with Sari Kovats by Ashley Ahearn" published online by *Environmental Health Perspectives* was recorded as an object of type journal article in Scopus.

- Apparent accidental misclassifications may occur in Scopus. Conference papers from the *Annual International Conference of the IEEE Engineering in Medicine and Biology – Proceedings* are duplicated in Scopus, with half classed as journal articles, but their conference paper duplicates tend to be credited with the citations.

Limitations

An important limitation of the method used here is that it is reliant on human judgement and important patterns may have been missed. In addition, some important reasons may not be evident from the data. For example, it seems likely that articles in some cases had many Mendeley readers compared to Scopus citations because the articles were cited in documents not covered by Scopus, such as many non-English journals (particularly for Horticulture), as well as many conference papers (particularly for Computer Vision) or books (particularly for Music). Another limitation is that the anomalies have been extracted from large collections of up to 10,000 articles and in any collection of such a size it is statistically likely for substantial outliers to appear due to random variations even in the absence of external causes. This particularly applies to the cases above where the evidence is based exclusively on the share of one category of user (e.g., Master's students).

Conclusions

Although Mendeley reader counts correlate strongly and positively with Scopus-indexed citation counts in most subject areas, some articles receive disproportionately many or few Mendeley readers in comparison to Scopus-indexed citations. There are many different reasons why this can happen. Although some of the reasons are technical, such as mistakes in DOIs or article titles, others are more fundamental. These fundamental reasons probably do not just cause a few outliers but also affect many other articles to a lesser extent. Perhaps the most important reason is that the ratio of reader counts to citation counts varies substantially by field and that interdisciplinary research articles may attract a ratio mainly from one of their constituent fields whilst legitimately being classified in another field. Related to this, an article may be used by a community (e.g., students, professionals) that do not cite it because they do not publish research or create publications that are not indexed by Scopus. Similarly an article may be cited by a community that tends not use Mendeley.

With the exception of the technical issues, the reasons for the differences between the Mendeley and Scopus data all suggest that the lower figure may underestimate the wider impact of an article. A logical conclusion from the results is that combining Mendeley readership data with Scopus-indexed citations in a way that uses the maximum rather than the minimum or average would give the most reasonable indicator for article impact. This maximum should take into account the differing average number of readers and citations. This suggestion does not seem to have been made before in published research but is worth considering in applications where it would be important to avoid even small numbers of anomalies, such as when evaluating individual academics. The results also confirm the importance of resolving technical issues with data collection and ensuring that the subject categories used are as homogeneous as possible.

References

Bar-Ilan, J. (2008). Which h-index?—A comparison of WoS, Scopus and Google Scholar. Scientometrics, 74(2), 257-271.

Bar-Ilan, J. (2012). JASIST@ Mendeley. Presented at ACM Web Science Conference Workshop on Altmetrics, Evanston, IL, 21 June 2012. http://altmetrics.org/altmetrics12/bar-ilan/

Costas, R., Zahedi, Z., & Wouters, P. (in press). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. Journal of the Association for Information Science and Technology.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. Journal of Medical Internet Research, 13(4).

Gunn, W. (2013). Social signals reflect academic impact: What it means when a scholar adds a paper to Mendeley. Information Standards Quarterly, 25(2), 33-39.

Harzing, A. (2014). A longitudinal study of Google Scholar coverage between 2012 and 2013. Scientometrics, 98(1), 565-575.

Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. Journal of Informetrics, 5(3), 446–457.

Haustein, S., Larivière, V., Thelwall, M., Amyot, D., & Peters, I. (2014). Tweets vs. Mendeley readers: How do these two social media metrics differ? IT-Information Technology, 56(5), 207-215.

Henning, V., & Reichelt, J. (2008). Mendeley - A last.fm for research? IEEE Fourth International Conference on eScience (eScience '08) (pp. 327-328). Los Alamitos: IEEE.

Kousha, K., & Thelwall, M. (2008). Assessing the impact of disciplinary research on teaching: An automatic analysis of online syllabuses. Journal of the American Society for Information Science and Technology, 59(13), 2060-2069.

Li, X, & Thelwall, M. (2012). F1000, Mendeley and Traditional Bibliometric Indicators. 17th International Conference on Science and Technology Indicators (Vol. 3, pp. 1–11).

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

Maflahi, N. & Thelwall, M. (in press). When are readership counts as useful as citation counts? Scopus vs. Mendeley for LIS journals. Journal of the Association for Information Science and Technology.

Mas-Bleda, A., Thelwall, M., Kousha, K. & Aguillo, I.F. (2014). Do highly cited researchers successfully use the Social Web? Scientometrics, 101(1), 337-356.

Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. Scientometrics, 49(1), 93-123.

Moed, H. F., & Visser, M. S. (2008). Appraisal of citation data sources. Centre for Science and Technology Studies, Leiden University. http://www.hefce.ac.uk/media/hefce/content/pubs/indirreports/2008/missing/Appraisal%20of%20Citation%20Data%20Sources.pdf

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (in press). Who reads research articles? An altmetrics analysis of Mendeley user categories. Journal of the Association for Information Science and Technology.

Mohammadi, E., Thelwall, M. & Kousha, K. (in press). Can Mendeley bookmarks reflect readership? A survey of user motivations. Journal of the Association for Information Science and Technology.

Price, D. D. (1965). Networks of scientific papers. Science, 149, 510–515.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2011). Altmetrics: A manifesto. http://altmetrics.org/manifesto.

Torres-Salinas, D., Lopez-Cózar, E. D., & Jiménez-Contreras, E. (2009). Ranking of departments and researchers within a university using two different databases: Web of Science versus Scopus. Scientometrics, 80(3), 761-774.

Sud, P. & Thelwall, M. (2014). Evaluating altmetrics. Scientometrics, 98(2), 1131-1143.

Thelwall M, Haustein S, Larivière V, Sugimoto CR (2013). Do Altmetrics Work? Twitter and ten other social web services. PLoS ONE, 8(5), e64841.

Thelwall, M. & Wilson, P. (2014a). Distributions for cited articles from individual subjects and years. Journal of Informetrics, 8(4), 824-839.

Thelwall, M. & Wilson, P. (2014b). Regression for citation data: An evaluation of different methods. Journal of Informetrics, 8(4), 963–971.

Thelwall, M. & Wilson, P. (in press). Mendeley readership Altmetrics for medical articles: An analysis of 45 fields, Journal of the Association for Information Science and Technology.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. The Rand Journal of Economics, 21(1), 172-187.

Wouters, P., Costas, R. (2012). Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century. Utrecht: SURF foundation. http://www.surffoundation.nl/nl/publicaties/Documents/Users%20narcissism%20and%20control.pdf

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. Scientometrics, 101(2), 1491-1513.