

# Mendeley Readership Counts: An Investigation of Temporal and Disciplinary Differences<sup>1</sup>

Mike Thelwall, Pardeep Sud

**Scientists and managers using citation-based indicators to help evaluate research cannot evaluate recent articles because of the time needed for citations to accrue. Reading occurs before citing, however, and so it makes sense to count readers rather than citations for recent publications. To assess this, Mendeley readers and citations were obtained for articles from 2004 to late 2014 in 5 broad categories (agriculture, business, decision science, pharmacy, and the social sciences) and 50 subcategories. In these areas, citation counts tended to increase with every extra year since publication, and readership counts tended to increase faster initially but then stabilise after about 5 years. The correlation between citations and readers was also higher for longer time periods, stabilising after about five years. Although there were substantial differences between broad fields and smaller differences between subfields, the results confirm the value of Mendeley reader counts as early scientific impact indicators.**

## Introduction

Scientists, research administrators, and funding bodies routinely need to evaluate the effectiveness of research for promotion, tenure, funding allocations or broad policy decisions. This is difficult because the value of research can take a long time to become apparent and may only be evident to small sets of domain experts. As a result, research evaluations may harness expert opinions, sometimes supported by quantitative indicators. Quantitative indicators may also be used in place of domain experts if the cost of harnessing appropriate expertise is high compared to the importance of the evaluation. These indicators are normally based on citation counts because these seem to give the most robust evidence for many fields (Moed, 2006), despite their known limitations (MacRoberts & MacRoberts, 1989), and are relatively easy to obtain from citation indexes, such as the Web of Science (WoS) or Scopus. An unavoidable problem with citation counts, however, is that they cannot be used for recent articles because citations take time to accrue due to the time required to publish and index the citing research. This issue can delay research evaluations that harness citation-based indicators for one or more years, slowing down any decision making based on the results. Early citation counts (e.g., 1-2 years after publication) are moderately reliable indicators of long term citation counts in some fields, except perhaps in the case of highly cited articles (Adams, 2005; Levitt & Thelwall, 2008; Wang, 2013). Longer time periods may be needed to give optimal results (Abramo, Cicero, & D'Angelo, 2012), although additional evidence may also improve the reliability of early impact estimates (Levitt & Thelwall, 2011). Nevertheless, the use of early citation counts still imposes a delay of at least a year for any evaluation.

Counts of readers in social reference sharing sites, such as Mendeley, may be an alternative to citation counts when evidence of early impact is needed. This is possible because article readers may start to be registered in Mendeley on the day that an article is published. Mendeley readership counts can also be used to give evidence of the types of

---

<sup>1</sup> Thelwall, M. & Sud, P. (2016). Mendeley readership counts: An investigation of temporal and disciplinary differences. *Journal of the Association for Information Science and Technology*, 57(6), 3036-3050. doi:10.1002/asi.23559 © copyright 2015 John Wiley & Sons, Inc.

people that read an article as well as their disciplines and countries of origin, and so can potentially provide more fine-grained impact evidence than can simple citation counts, even for older articles. Since Mendeley began in 2008, it is not clear, however, whether readership counts for articles from before 2008 will be substantial enough to justify their use in evaluations. Hence there are two important time issues for the use of Mendeley readership counts in research evaluations: do they occur more quickly than citations; and are they useful for historical analyses of older articles from before, say, 2008?

Although it seems intuitively obvious that Mendeley readership counts can appear soon after an article has been published and it also seems likely that they will be less common for articles from before 2008, only one previous paper has systematically investigated the role of time in Mendeley citations. Based upon four journals from a single discipline, it suggested that readership counts would be acceptable proxies for citation counts for articles that were 1-10 years old (Maflahi & Thelwall, in press). Nevertheless, it may be that Mendeley readership may be less common or occur later in other fields or for articles in less prestigious journals, and so it is not clear whether these recommendations can apply to other fields. This article compares Scopus citations with Mendeley readership counts in order to assess systematically how readership counts vary over time and how their relationship with citation counts changes over time.

## Background

Although research-related management decisions, such as for appointments, promotion, tenure, and funding applications, tend to rely on human judgements, this is time-consuming for the experts involved. Moreover, peer judgements may be influenced by nepotism or by conscious or unconscious sources of bias, for example relating to gender, race or nationality. In response to these issues, some research management and funding decisions are now made primarily through bibliometric indicators and others are supported by bibliometric indicators. For example Journal Impact Factors (JIFs) are important for research funding decisions in Spain (Bordons, Fernández, & Gomez, 2002; Jimenez-Contreras, Robinson-Garcia, & Cabezas-Clavijo, 2011) and have been tied directly to funding in Finland (Adam, 2002) and Flanders (Richardson, 2011). In the UK, the major periodic national Research Excellence Framework (REF) is led by peer review but is aided by citation counts and annual benchmarks for them for some fields (HEFCE, 2013). Bibliometrics have also been used in Excellence in Research for Australia (ERA) to support peer review (Kwok, 2013), for example through the provision of nationally agreed journal rankings (Haslam & Koval, 2010). The next evaluation for New Zealand's Performance-Based Research Fund (PBRF) is set to aid reviewers with impact factor and citation data (TEC, 2014, p17), and citation-based metrics are also proposed for the Valutazione Quinquennale della Ricerca (VQR) in Italy (Richardson, 2011). Bibliometrics are also used for performance monitoring (e.g., Labrosse, Ventimiglia, Campbell, Haustein, Côté, & Archambault, 2013), although this does not affect those evaluated as directly as the examples discussed above.

It is known that evaluation systems can lead to large-scale measurable changes in scholarship. In Australia, for example, a previous system that rewarded research productivity rather than quality seemed to discourage high quality research (Butler, 2003). In this context, a universal limitation of citation-based indicators is that they do not directly reflect the wider societal impacts that governments seek. As a result, evaluation systems that use bibliometrics alone may push academics towards ignoring societal impacts. Hence

it seems reasonable to look for societal impact indicators, perhaps for use in conjunction with bibliometrics for a more balanced overall range of indicators.

## **The birth of altmetrics**

Alternative metrics, such as those using Mendeley readership counts, have arisen mainly due to the inability of citation counts to reflect research impacts outside of academia. In the UK REF, wider societal impact is explicitly assessed separately through a set of evidence-based case studies but these narratives are difficult to make convincing, even in apparently straightforward cases (Parker & van Teijlingen, 2012). The need to assess wider impacts can also be a concern for research funders (Sainty, 2013) and so there is a general need for indicators that are not based on citation data. There have been several attempts to develop such indicators. Patent citations have been proposed as a way to reflect knowledge transfer to commercial applications (Meyer, 2000; Tijssen, 2001; Trajtenberg, 1990), although their value may be restricted to specific technological sectors for which patenting is common (Brouwer & Kleinknecht, 1999; Coupe, 2003). Web based indicators have also been proposed for quantifying the overall fame of individual academics, based upon web mentions of them (Cronin, Snyder, Rosenbaum, Martinson, & Callahan, 1998) and for the general impact of academic websites based upon hyperlinks to them (Ingwersen, 1998). The web has also been exploited as a de-facto citation index for academic publications by counting links or citations to them from web pages (Vaughan & Shaw, 2003, 2005), or other web-based references to articles (Kousha & Thelwall, 2006) and for alternative journal IFs (Vaughan, & Hysen, 2002). Moreover, some new web-based indicators have targeted specific types of non-scholarly impact, such as value within education (Kousha & Thelwall, 2008).

The emergence of the social web as an important phenomenon has led to the creation of a raft of new academic indicators that, in theory, could reflect wider impacts of articles. This was triggered by an explicit call for the creation and evaluation of such "altmetrics" (Priem, Taraborelli, Groth, & Neylon, 2010). The new metrics include those based on tweets (Eysenbach, 2011), Wikipedia citations (Priem, Piwowar, & Hemminger, 2012), blog citations (Shema, Bar-Ilan, & Thelwall, 2012, 2014), social references (Li, Thelwall, & Giustini, 2012) and many others (Costas, Zahedi, & Wouters, 2014a; Thelwall, Haustein, Larivière, & Sugimoto, 2013). The data for these metrics may be gathered by the researcher using specialist software or obtained from commercial (Adie & Roe, 2013) or other data providers. Nevertheless, despite the social web being largely public and free, no study so far has found clear evidence that any social web indicators reflect types of impact outside of academia. For example, the wide uptake of Twitter in many countries makes it a logical source of evidence about public interest in science. Nevertheless, although academic research is extensively tweeted about by scholars (Priem, & Costello, 2010; Weller & Puschmann, 2011), including in tweets aimed at the public (Desai, Shariff, Shariff, Kats, Fang, Christiano, & Ferris, 2012), there is no evidence that research articles are often tweeted by the public (e.g., Thelwall, Tsou, Weingart, Holmberg, & Haustein, 2013).

In addition to the lack of evidence about wider social impacts, altmetrics are relatively easy to spam (e.g., Haustein, Bowman, Holmberg, Tsou, Sugimoto, & Larivière, in press) and so should not be used for evaluations in which those evaluated have an interest in achieving higher scores in order to obtain a better outcome (Wouters & Costas, 2012). Finally, altmetrics are unlikely to be exempt from the well-known limitations of citation-

based metrics (MacRoberts & MacRoberts, 1989) and so all altmetrics will have problematic aspects, if used in evaluations.

Even though altmetrics and other alternative metrics do not seem to provide evidence of wider societal impact for research, they may still be useful as early indicators of research impact within academia. Moreover, triangulating them with citations may help to remove some of the existing biases within current citation indexes, although there is no credible evidence yet to support the latter claim. In this context, social reference sharing sites seem to be the most useful. This is because they are widely used by academics compared to other sources of altmetrics, except perhaps Twitter (Costas, Zahedi, & Wouters, 2014a; Thelwall, Haustein, Larivière, & Sugimoto, 2013; Zahedi, Costas, & Wouters, 2014) and, unlike Twitter, are not easily used for publicity and so seem to be less likely to be spammed. The use of Twitter for publicity by journals and authors probably explains why tweet counts and citations have a correlation close to zero for academic articles (Thelwall, Haustein, Larivière, & Sugimoto, 2013; Haustein, Larivière, Thelwall, Amyot, & Peters, 2014), although Tweets may be more promising in specific cases, such as for online open access medical journals (e.g., Eysenbach, 2011).

### ***Mendeley readership counts***

The social reference sharing website Mendeley is designed to help users to manage and share their academic references (Henning & Reichelt, 2008). As a side-effect, it is possible to count readers in order to generate indicators for academic articles (Li, Thelwall, & Giustini, 2012), whole journals (Haustein & Siebenlist, 2011) or for knowledge transfer between disciplines (Mohammadi & Thelwall, 2014) and countries (Thelwall & Maflahi, in press), or for triangulation with citations in bibliometric studies (Sud & Thelwall, in press; Thelwall & Wilson, in press). Most people registering academic articles in Mendeley have already read them or intend to read them (Mohammadi, Thelwall, & Kousha, in press) and so it seems reasonable to use the term *reader count* to refer to the number of Mendeley users registering an article in their personal libraries. Nevertheless, readers can disappear from Mendeley and so reader counts may occasionally decrease (Bar-Ilan, 2014).

Mendeley users are asked to record their professional status when they join and from this information it seems that at least 95% are academics or students, with PhD students, postgraduates and postdocs being heavily represented (Mohammadi, Thelwall, Haustein, & Larivière, in press). Users are also asked to declare their subject areas of interest. From this information, readership seems to vary by discipline, with some areas of research tending to attract relatively many Mendeley readers compared to WoS citations (Costas, Zahedi, & Wouters, 2014b). Hence, Mendeley readership counts are indicators of predominantly academic readership of articles, with disciplinary variations, with a bias towards younger researchers (for older, highly successful researchers see also: Mas-Bleda, Thelwall, Kousha, & Aguillo, 2014) and probably also towards researchers (and research: Alperin, 2014) in wealthier nations and those that are more accustomed to computing technologies. In particular, Mendeley readership does not reflect wider impacts of research outside of academia and might perhaps be thought of instead as an academic usage metric (e.g., Haustein & Siebenlist, 2011; Tarborelli, 2008).

A number of statistically significant positive correlations between Mendeley readership counts and citations have shown that Mendeley readership data is not random but has some relationship with traditional citation impact (for a justification, see: Sud & Thelwall, 2014). In particular, Web of Science citations correlate strongly with Mendeley

reader counts for Nature (0.56) and Science (0.54) articles from 2007 (and were more common than CiteULike readers) (Li, Thelwall, & Giustini, 2012), Scopus citations correlate moderately (0.45) with Mendeley readers (and were three times more numerous than CiteULike readers) for articles from selected bibliometricians (Haustein, Peters, Bar-Ilan, Priem, Shema, & Terliesner, 2014), and Web of Science citations correlate moderately overall with Mendeley readers in clinical medicine (0.46), engineering and technology (0.33), social science (0.45), physics (0.31) and chemistry (0.37) (Mohammadi, Thelwall, Haustein, & Larivière, in press). All of the aforementioned correlations are for relatively short time windows of a few years, however, and the figures would probably be higher for longer time windows, as well as for more recent data. If the readership categories are broken down by occupation, the correlations are highest for PhD students and postdocs, and lowest for librarians, undergraduates and non-academics (the Other Professional category) (Mohammadi, Thelwall, Haustein, & Larivière, in press).

Mendeley reader counts correlate more highly with citations than do all other altmetrics tested so far (Thelwall, Haustein, Larivière, & Sugimoto, 2013), making them the most promising altmetric alternative to citation counts. Mendeley seems to be mainly useful for journal articles, however. For example, few books seem to be registered in comparison to articles in the same field (Hammarfelt, 2014), with perhaps 4%-14% of academic monographs (from selected reputable publishers) having at least one Mendeley reader (Kousha & Thelwall, in press).

One study of Mendeley has explicitly addressed the issue of changes over time. It found a low correlation of 0.2 between citations and readers in the year after publication for four library and information science journals and this steadily increased to 0.6 after seven years (Maflahi & Thelwall, in press). Since this study was restricted to one narrow discipline, it is important to assess the extent to which similar results would also occur in other areas.

## Research questions

This article assesses the influence of time on the extent to which articles attract Mendeley readers and the influence of time on the extent to which article readers correlate with article citations. Correlation tests are an important first method to evaluate altmetrics (Sud & Thelwall, 2014). The following research questions drive the study.

1. How do the correlations between citations and Mendeley readers for articles change over time?
2. How do the average numbers of Mendeley readers for articles change over time?
3. Are there broad or narrow field differences in the answers to the above questions?

These questions are addressed with analyses of publications during the period 2004-2014 from five broad fields and their subfields, to give both breadth and depth of coverage. The time period of ten years was chosen to include a substantial period of time before the public release of Mendeley in August 2008. Scopus was used because its coverage seems to be somewhat wider and more international than that of WoS (Erfanmanesh & Didegah, 2013; Li, Burnham, Lemley, & Britton, 2010; Minasny, Hartemink, McBratney, & Jang, 2013).

## Methods

Scopus categories were used as the basic source of field delimitation. Although it would have been possible to choose other categorisation systems, such as those from WoS or the National Science Foundation (NSF), or to derive categories from article text or citations (Braam, Moed, & van Raan, 1991), the use of Scopus categories has the advantages of being

transparent and reproducible. The following Scopus categories were selected: *Agricultural and Biological Sciences*; *Business, Management and Accounting*; *Decision Sciences*; *Pharmacology, Toxicology and Pharmaceuticals*; and *Social Sciences*. This gave a broad range of different research, including some from the natural, social, medical and formal sciences. This is a relatively arbitrary choice, however.

All articles in the chosen subjects and years were downloaded together with their citation counts from Scopus during November and December 2014. The articles were searched for in Mendeley, also during November and December 2014, by a query for title, year and first author last name, as well as by Digital Object Identifier (DOI), when present, in order to get the greatest possible coverage (Zahedi, Haustein, & Bowman, 2014). The two month period used to obtain citations and Mendeley readers is a limitation because data gathered later may have higher reader and citation counts, so small differences should be ignored in comparisons between different data sets, especially for articles from 2014. The query returns many false matches, so the top ten matches were retrieved and automatically checked. Articles with a DOI in both Mendeley and Scopus were kept if the DOIs matched, otherwise rejected. Articles without a DOI in one or both were kept only if their years matched, their first author last names approximately matched, and if their titles contained at least 85% of the same words. This is a conservative procedure that will occasionally reject articles, for example if they had short titles containing typographical errors in Scopus or Mendeley, or if the user recorded an incorrect year or abbreviated the article title.

For each subfield and year, the average number of citations per publication and the average number of Mendeley readers per publication were calculated in order to reveal differences over time. The geometric mean was used for the average because citation counts and Mendeley readers both follow highly skewed distributions. The geometric mean was modified by adding 1 to all the raw data before the averaging, and then subtracting 1 from the final average. This is a standard variant of the geometric mean that allows meaningful averages to be calculated even though the data contains many zeros. The result is a type of average that is not greatly changed by individual large values. Spearman correlations were calculated between the citation counts and reader counts for each subfield and year in order to assess the extent to which the two agree. Pearson correlations were not used because the data is skewed.

## Results

There are positive correlations between Mendeley reader counts and Scopus citation counts for all years examined. The correlations are stable or slowly decreasing between 2004 and about 2009 and then decrease with increasing speed until 2014 (figures 1-6). The graph shapes have some differences between broad categories, however. For example, *Business, Management and Accounting* (Figure 2) shows a more gradual decrease in correlations from about 2005 whereas most of the decrease in correlations for *Agricultural and Biological Sciences* (Figure 1) occurs between 2013 and 2014.

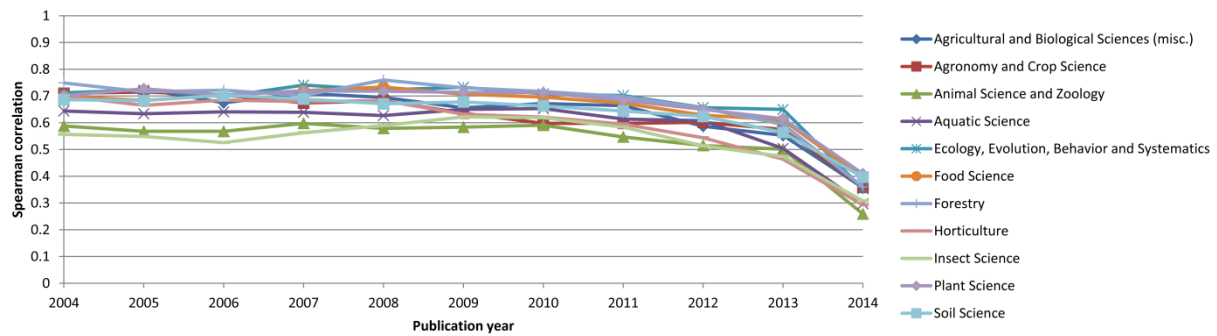


Fig. 1. Correlations between Scopus citation counts and Mendeley reader counts for subcategories of the Scopus *Agricultural and Biological Sciences* category.

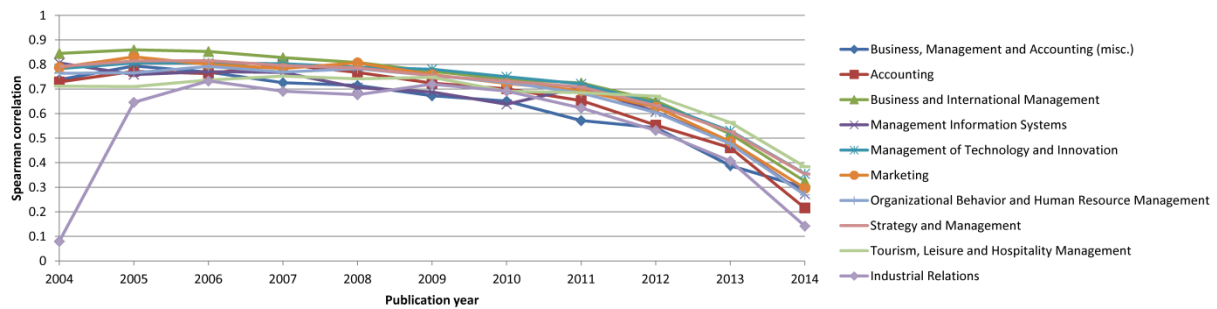


Fig. 2. Correlations between Scopus citation counts and Mendeley reader counts for subcategories of the Scopus *Business, Management and Accounting* category.

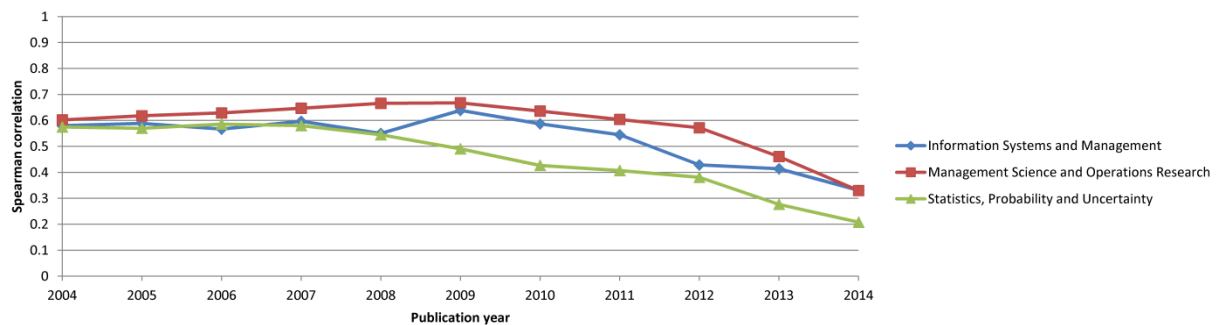


Fig. 3. Correlations between Scopus citation counts and Mendeley reader counts for subcategories of the Scopus *Decision Sciences* category.

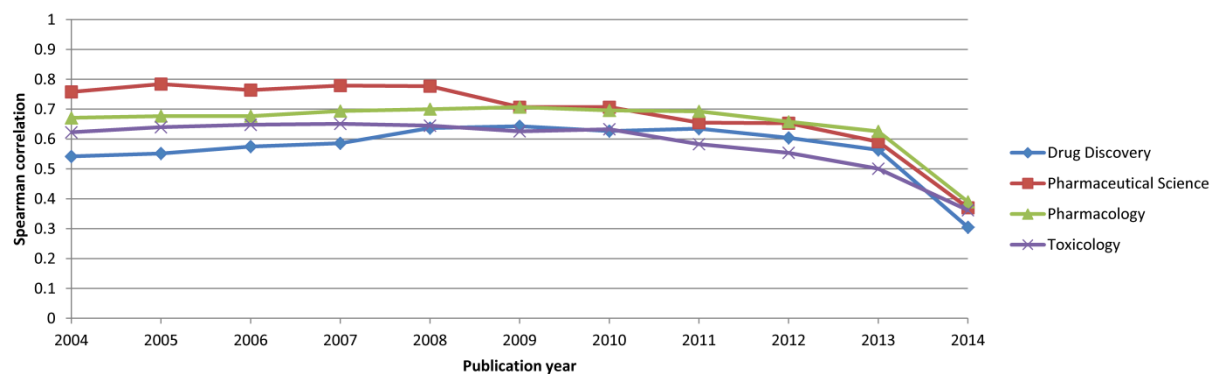


Fig. 4. Correlations between Scopus citation counts and Mendeley reader counts for subcategories of the Scopus *Pharmacology, Toxicology and Pharmaceuticals* category.

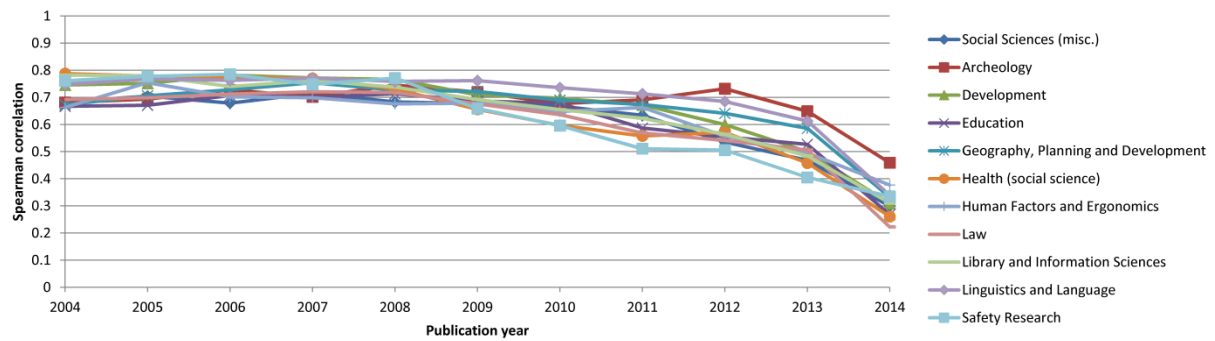


Fig. 5. Correlations between Scopus citation counts and Mendeley reader counts for subcategories of the Scopus *Social Sciences* category (first half).

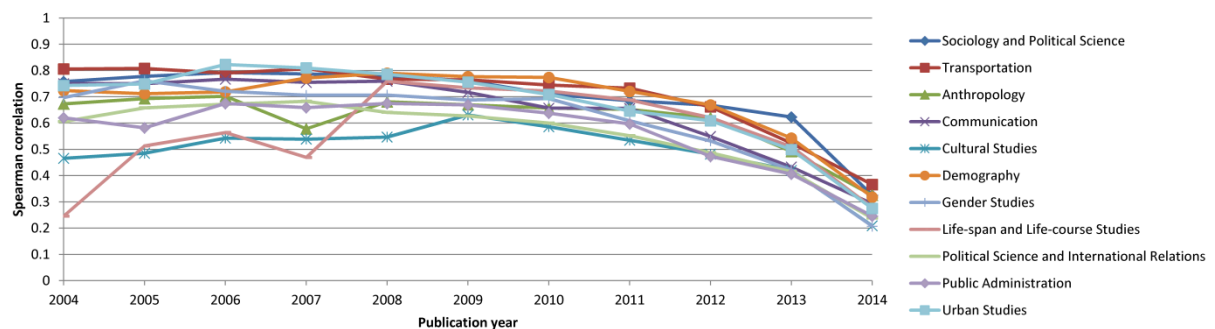


Fig. 6. Correlations between Scopus citation counts and Mendeley reader counts for subcategories of the Scopus *Social Sciences* category (second half).

## Citation counts

Average citation counts are reported for the articles in the data set in order to contrast with reader counts. The citation counts seem to decrease almost linearly over time in most, but not all, categories (figures 7-12), although they would presumably tend to stabilise at some earlier point in time, due to obsolescence, although researchers are increasingly willing to cite older papers (Larivière, Archambault, & Gingras, 2008). The geometric mean citation counts close to zero in 2013 and 2014 explain the low correlations with Mendeley reader counts in these years and confirm the need to seek alternatives to citation counts for recently-published articles.

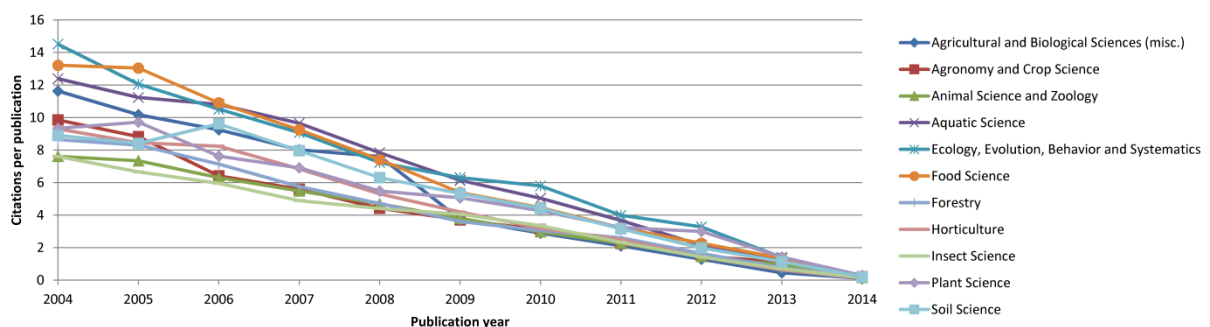


Fig. 7. Scopus geometric mean citations per publication for subcategories of the Scopus *Agricultural and Biological Sciences* category.



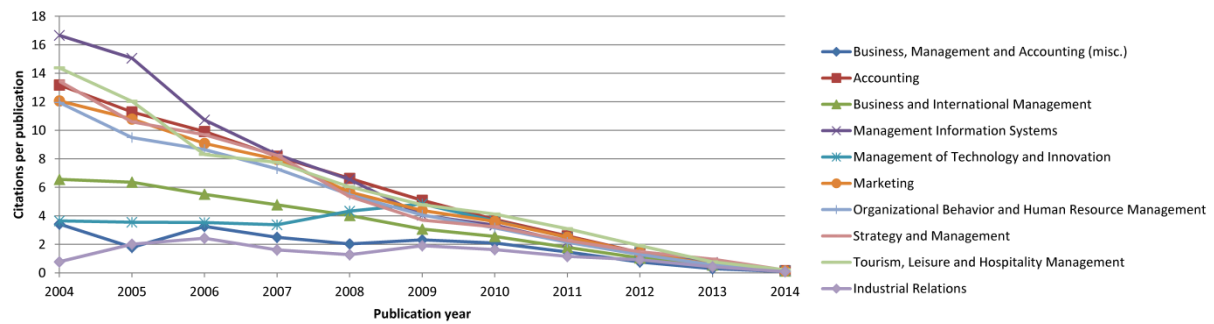


Fig. 8. Scopus geometric mean citations per publication for subcategories of the Scopus *Business, Management and Accounting* category.

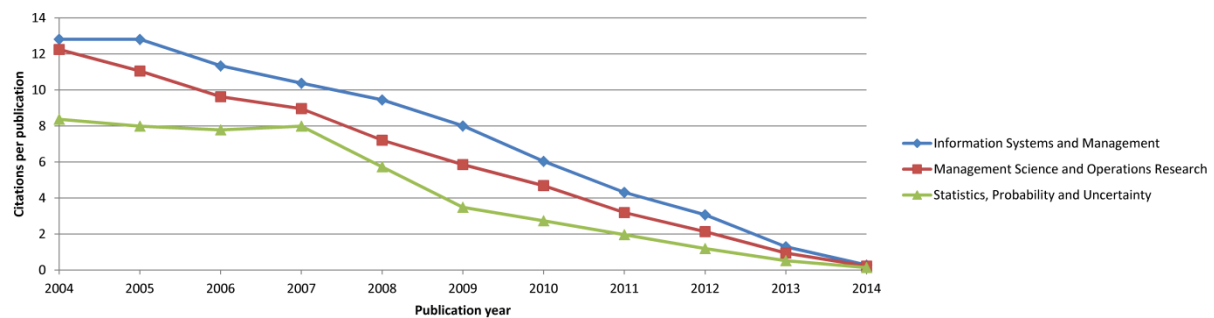


Fig. 9. Scopus geometric mean citations per publication for subcategories of the Scopus *Decision Sciences* category.

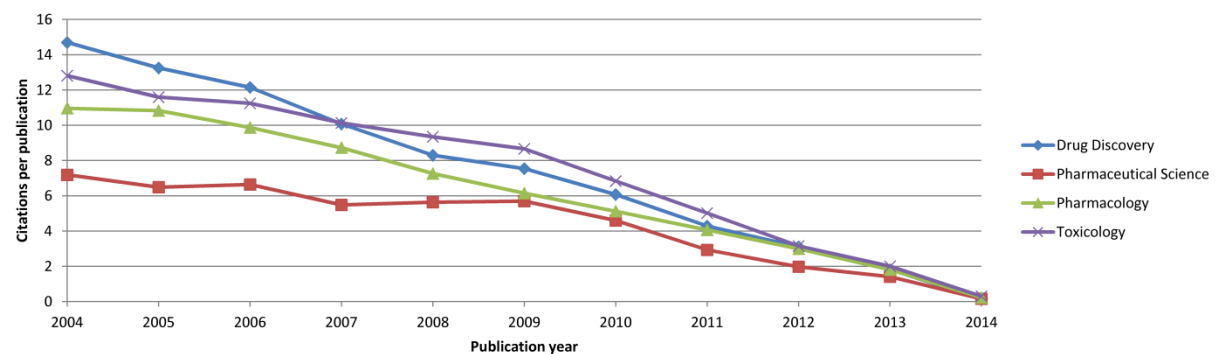


Fig. 10. Scopus geometric mean citations per publication for subcategories of the Scopus *Pharmacology, Toxicology and Pharmaceuticals* category.

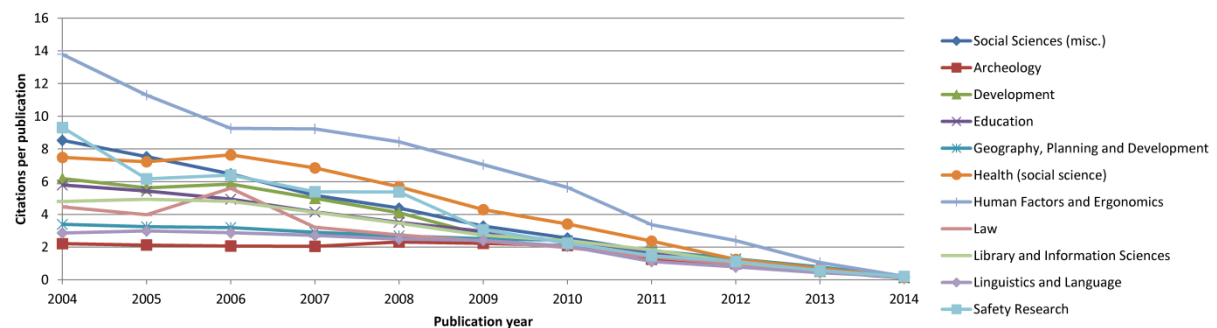


Fig. 11. Scopus geometric mean citations per publication for subcategories of the Scopus *Social Sciences* category (first half).

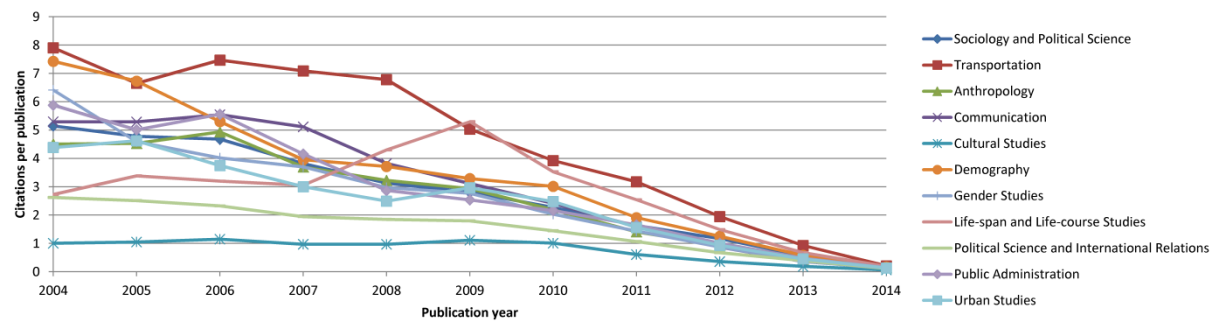


Fig. 12. Scopus geometric mean citations per publication for subcategories of the Scopus *Social Sciences* category (second half).

## Readership counts

The shapes of the Mendeley readership count graphs are substantially different to the shapes of the citation count graphs. Although there are many exceptions, the graphs tend to be relatively stable from 2004 to about 2011 before starting to decrease more rapidly until 2014, when they are mostly above 1 (figures 13-18). There are clearly differences between and within broad fields in the total number of readers per publication and in the overall shapes of the graphs.

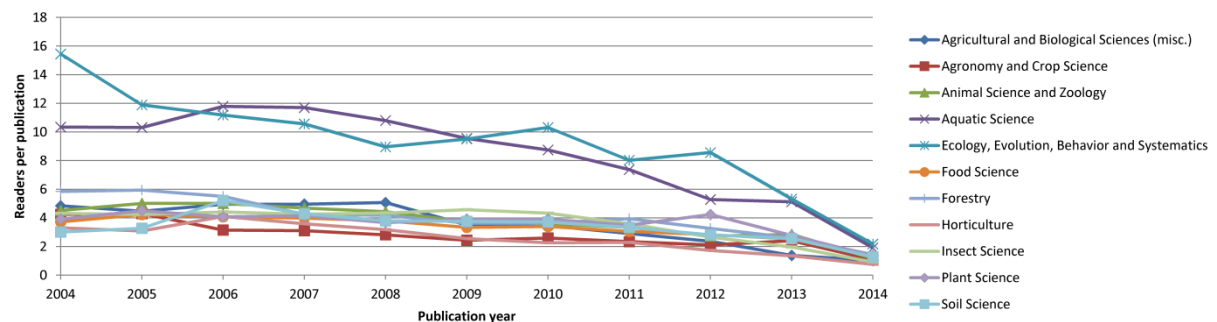


Fig. 13. Geometric mean Mendeley readers per publication for subcategories of the Scopus *Agricultural and Biological Sciences* category.

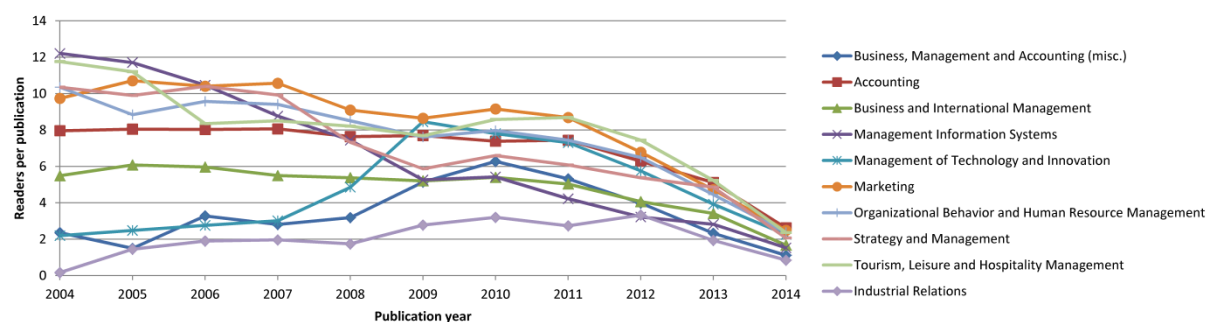


Fig. 14. Geometric mean Mendeley readers per publication for subcategories of the Scopus *Business, Management and Accounting* category.

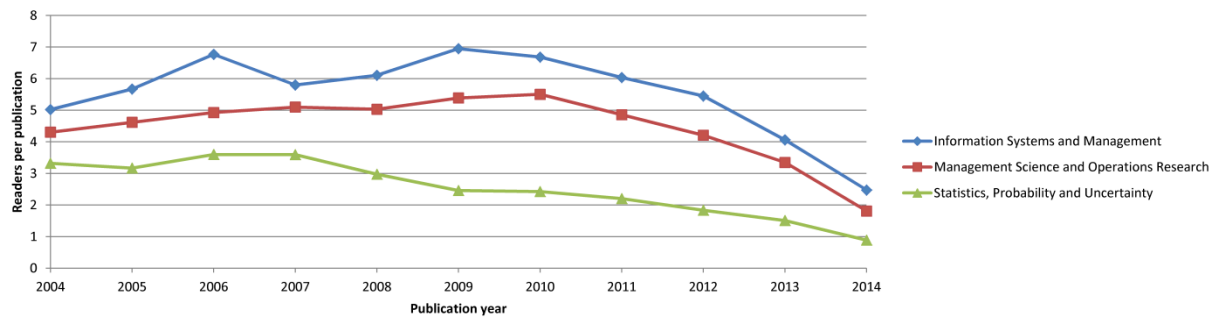


Fig. 15: Geometric mean Mendeley readers per publication for subcategories of the Scopus *Decision Sciences* category.

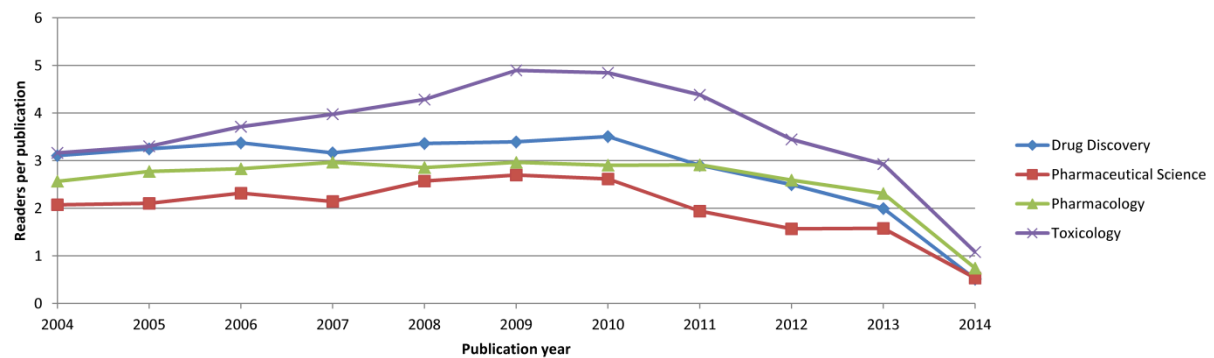


Fig. 16: Geometric mean Mendeley readers per publication for subcategories of the Scopus *Pharmacology, Toxicology and Pharmaceutics* category.

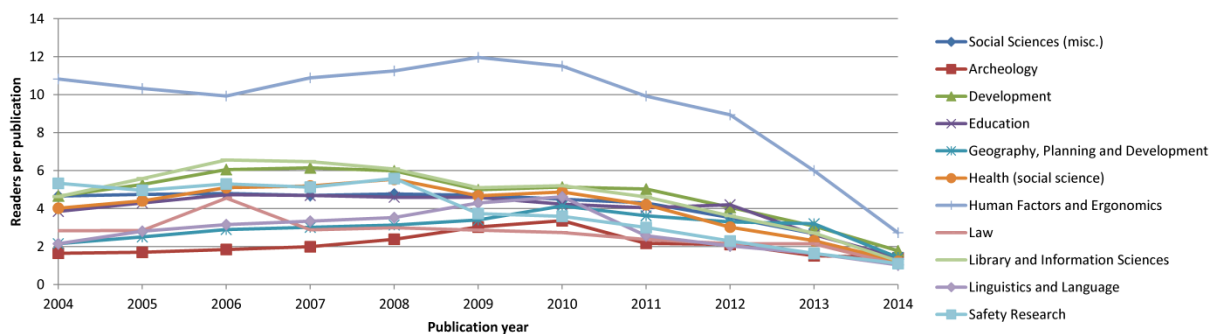


Fig. 17: Geometric mean Mendeley readers per publication for subcategories of the Scopus *Social Sciences* category (first half).

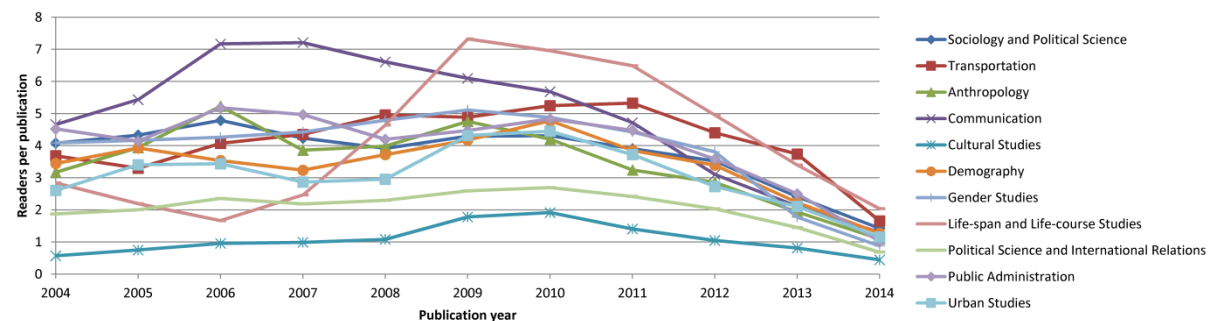


Fig. 18: Geometric mean Mendeley readers per publication for subcategories of the Scopus *Social Sciences* category (second half).

## Discussion

The results above should be interpreted cautiously because they are limited to the broad Scopus categories chosen and may be different for other sets of articles. In particular, medical research areas with fast growth rates for citations would presumably have steeper lines close to 2014 in the citation count and correlation graphs. Perhaps slower moving humanities fields might have more shallow lines and lower correlations, similar to that of archaeology in the figures above. It is also possible that Mendeley is almost ignored in some fields, giving very low reader counts and correlations, although this seems unlikely given the universally positive findings in previous research covering many different research areas. Another limitation is that the 2014 data is only partial since it was collected at the end of the year and the correlations from this year, although low, may be artificially inflated by older articles in the same year being more likely to have citations and readers. The same is true for all years to some extent but this issue is strongest for 2014. A more fundamental limitation is that the Mendeley reader counts are based on searches for articles in addition to DOI matches, when the DOI is available. The reader counts will therefore tend to be more comprehensive for articles with DOIs. Presumably, older articles are less likely to have DOIs in Scopus and so the shape towards the left of the Mendeley reader graphs and correlation graphs may underestimate to a small extent the continued growth in Mendeley reader counts and correlations over time (i.e., due to Mendeley records with typographical errors not matching relevant Scopus articles).

Even the relatively narrow Scopus sub-categories are likely to include specialisms with particularly high or low citation cultures and this may exaggerate the extent of the correlations found. To test for this, the correlations were calculated separately for each journal in each subcategory and then the journal correlations were averaged for each subcategory. This method takes specialism into account to some extent because many journals have a much narrower focus than the broad Scopus category in which they sit. This journal-based approach reduced the correlations by about 0.2 (see figures 19 and 20 – the graphs for the other subject areas are not shown) but the reduction is likely to be at least partly due to journal stratification rather than specialisation. For example, articles in high impact general journals would tend to attract more citations and more readers than would articles in low impact general journals but separating the two in the correlation calculation would reduce the overall correlation between citation counts and reader counts. Most importantly, however, the overall shapes of the lines in the figures are broadly similar.

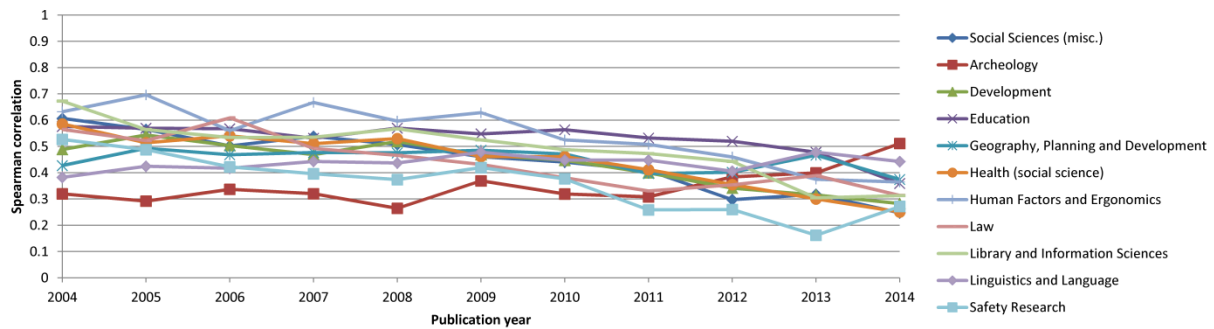


Fig. 19. Mean Spearman correlations between citation counts and readership counts by publication year for the social science (A) set of articles. The correlations were calculated separately for each journal with at least 2 articles published in the year and then the journal-level correlations were averaged over the year.

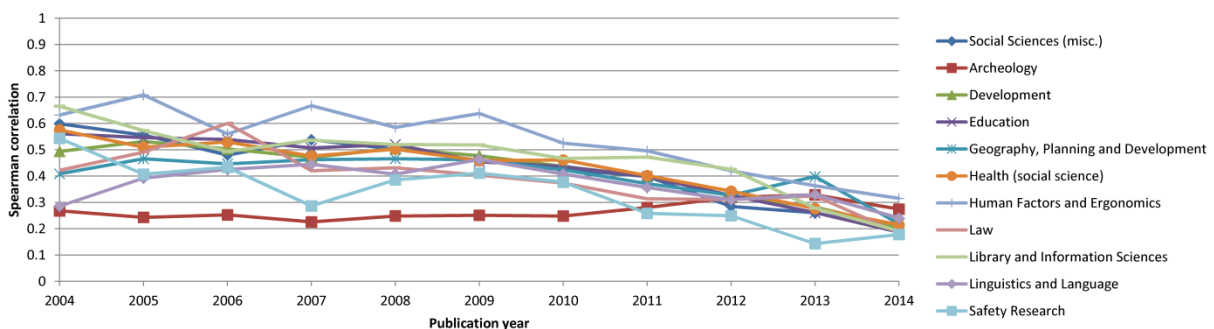


Fig. 20. Mean Spearman correlations between citation counts and readership counts by publication year for the social science (A) set of articles. The correlations were calculated separately for each journal with at least 10 articles published in the year and then the journal-level correlations were averaged over the year.

The individual categories, such as Archaeology, that do not follow the general trends in figures 1-6 are of particular interest. The two largest deviations are probably due to the low numbers of articles in the early years: Industrial Relations (Figure 2) has only 70 articles in 2004, rising to 1441 in 2009, and Life-span and Life-course Studies (Figure 6) has only 66 articles in 2004, rising to 1363 in 2013. The decreasing correlation from 2009 back to 2004 for Cultural Studies (Figure 6) is also unusual and is not caused by low numbers because each year has a minimum of 3891 articles. The explanation is probably the low number of Mendeley readers per article, peaking in 2010 (Figure 18 – see also the discussion below).

The differing average citation counts per publication and different graph shapes in figures 7-18 are not surprising because reference list sizes and citation counts are known to vary between disciplines, as is the speed with which articles attract citations and the rate of obsolescence of the academic literature. The average (geometric mean) citation counts decrease for older articles within Industrial Relations (Figure 8) and Life-span and Life-course Studies (Figure 12), which is presumably due to the low numbers of articles in the early years in both cases. The same trend is evident for Management of Technology and Innovation (Figure 8), which has the highest average (geometric mean) citation count in 2009. This may be an artefact of a change in the set of journals included within the category from 2009, dropping a large set of management science publications (e.g., *Management Science*, *Journal of the Operational Research Society*) and some general journals from

Chinese technical universities (e.g., *Journal of Wuhan University of Technology*, *Journal of Huazhong University of Science and Technology (Natural Science Edition)*).

The shapes of figures 13-18 are likely to be influenced by the release date of Mendeley in late 2008, assuming that users tend to record recently published articles disproportionately often. In this context, it would be reasonable for average reader counts to peak in 2008 and decrease steadily for years before 2008. This shape can be seen in many of the graphs, such as most lines in Figure 15 and Figure 16 and some of the lines in all the other figures. There are more variations in the shapes of the Mendeley graphs than in the shapes of the citation graphs, however, and some of the averages fairly steadily increase for older articles. Examples of this include Statistics, Probability and Uncertainty (Figure 15) and Ecology, Evolution, Behavior and Systematics (Figure 13). It is possible that older articles in more slowly-moving subjects continue to be recorded by Mendeley readers at a relatively high rate. In Figure 13, both Aquatic Science and Ecology, Evolution, Behavior and Systematics seem to be particularly extensively recorded by Mendeley users in comparison to citations (compare with Figure 7), giving evidence of disciplinary differences in uptake of Mendeley, even within a single broad category.

To give some context to the results, the overall level of citing and reading articles varies substantially by field, as does the correlation between citations and readers (Table 1). Reader and citation counts for fields correlate strongly (0.582) but the correlation is not high enough to suggest that the differences are irrelevant (Table 2). Moreover, the higher correlation between mean Mendeley readers and the overall readers/citers correlation (0.410) than the correlation between mean citations and the overall readers/citers correlation (0.214) suggests that the uptake of Mendeley within a field is an important variable pointing to the usefulness of Mendeley reader counts.

Table 1. Summary statistics for the entire period 2004-2014.

Subject	Mean citations per pub.	Mean readers per pub.	Nonzero citations	Nonzero Mendeley readers	Correlation between citations & readers
Agricultural and Biological Sciences (misc.)	5.2	3.5	82%	69%	0.64
Agronomy and Crop Science	4.3	2.8	75%	67%	0.62
Animal Science and Zoology	3.9	3.7	76%	69%	0.54
Aquatic Science	6.4	8.5	85%	87%	0.59
Ecology, Evolution, Behavior & Systematics	6.8	9.3	81%	82%	0.67
Food Science	6.4	3.3	80%	74%	0.66
Forestry	4.2	4.0	71%	68%	0.67
Horticulture	4.6	2.6	74%	63%	0.60
Insect Science	3.8	3.6	75%	75%	0.54
Plant Science	5.1	3.6	80%	71%	0.67
Soil Science	5.2	3.4	79%	68%	0.64
Business, Management & Accounting (misc.)	1.8	3.4	55%	70%	0.62
Accounting	5.7	6.9	73%	86%	0.65
Business and International Management	3.3	4.8	61%	69%	0.72
Management Information Systems	6.3	6.6	68%	75%	0.65
Management of Technology & Innovation	2.9	4.6	65%	64%	0.71
Marketing	5.3	8.3	71%	85%	0.69
Org. Behav. & Human Resource Management	4.9	7.5	69%	84%	0.67
Strategy and Management	5.4	7.2	69%	77%	0.70
Tourism, Leisure & Hospitality Management	5.8	8.0	72%	88%	0.67
Industrial Relations	1.3	2.0	50%	61%	0.54
Decision Sciences (miscellaneous)	1.7	0.8	48%	50%	0.30
Information Systems and Management	7.3	5.5	80%	84%	0.53
Management Science & Operations Research	6.0	4.5	76%	80%	0.58
Statistics, Probability and Uncertainty	4.4	2.5	71%	72%	0.46
Pharmacology, Toxicology & Pharmaceuticals	0.9	1.4	34%	45%	0.49
Drug Discovery	7.4	2.8	84%	77%	0.57
Pharmaceutical Science	4.4	2.0	72%	60%	0.69
Pharmacology	6.2	2.6	79%	70%	0.65
Toxicology	7.4	3.6	86%	83%	0.59
Social Sciences (misc.)	3.8	4.1	69%	79%	0.61
Archeology	1.6	2.1	54%	49%	0.68
Development	3.2	4.7	66%	78%	0.66
Education	3.0	3.9	67%	72%	0.61
Geography, Planning and Development	2.2	3.0	64%	66%	0.66
Health (social science)	4.3	4.1	69%	77%	0.63
Human Factors and Ergonomics	6.5	9.5	79%	93%	0.63
Law	2.5	2.7	59%	63%	0.61
Library and Information Sciences	2.8	4.7	60%	72%	0.65
Linguistics and Language	1.9	2.8	54%	60%	0.70

Safety Research	3.8	3.8	60%	74%	0.62
Sociology and Political Science	2.7	3.7	64%	70%	0.70
Transportation	4.6	4.1	73%	69%	0.71
Anthropology	2.6	3.5	61%	69%	0.61
Communication	3.1	4.9	58%	75%	0.64
Cultural Studies	0.8	1.1	40%	45%	0.49
Demography	3.4	3.4	64%	71%	0.68
Gender Studies	2.7	3.9	60%	82%	0.61
Life-span and Life-course Studies	2.8	4.1	66%	82%	0.55
Political Science & International Relations	1.5	2.1	53%	63%	0.56
Public Administration	2.9	4.0	61%	75%	0.57
Urban Studies	2.4	3.1	57%	66%	0.67

Table 2. Pearson correlations between the columns of Table 1.

	Mean citations per pub.	Mean readers per pub.	Non-zero citations	Non-zero readers	Citations / readers correlation
Mean citations per pub.	1	0.582	0.895	0.694	0.214
Mean readers per pub.		1	0.466	0.794	0.410
Non-zero citations			1	0.674	0.265
Non-zero readers				1	0.275

## Conclusions

The results confirm a pattern found in a previous study of four library and information science journals (Maflahi & Thelwall, in press): correlations between citation counts and reader counts for journal articles tend to increase over five years and then stabilise. The evidence presented here suggests that this is a phenomenon that is widespread across academic subjects, although perhaps not quite universal (e.g., cultural studies, and perhaps the arts and humanities in general). Moreover, the cause of the low correlations in early years appears to be the very low numbers of citations, since there tend to be substantially more readers than citers in the few years after publication. Hence, the results suggest that Mendeley reader counts are a good substitute for citation counts for scientometric applications involving articles that are younger than five years old, and especially when the articles are in their first few years, and especially in fields with higher levels of Mendeley use. Nevertheless, Mendeley counts should not be used for evaluations with stakeholders because they can be manipulated (Wouters & Costas, 2012). Disciplinary differences in uptake of Mendeley should also be taken into account in evaluations because of substantial differences in the average numbers of readers per article (figures 13-18).

An important limitation for the conclusion about the value of Mendeley readers for recent papers is that it has not been proven yet that early Mendeley readers correlate with (i.e., can be used to predict) long term citations, in contrast to the case of Tweets for one specialist journal (Eysenbach, 2011). Whilst it seems likely that this would be true and that the low correlations with citation counts for recent articles are due to the low numbers of citations, future work with early Mendeley reader counts and later citation counts for the same set of articles is urgently needed to check this hypothesis.



## References

- Abramo, G., Cicero, T., & D'Angelo, C. A. (2012). A sensitivity analysis of research institutions' productivity rankings to the time of citation observation. *Journal of Informetrics*, 6(2), 298-306.
- Adam, D. (2002). Citation analysis: The counting house. *Nature*, 415(6873), 726-729.
- Adams, J. (2005). Early citation counts correlate with accumulated impact. *Scientometrics*, 63(3), 567-581.
- Adie, E., & Roe, W. (2013). Altmetric: enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, 26(1), 11-17.
- Alperin, J.P. (2014). Exploring altmetrics in an emerging country context. Figshare. <http://dx.doi.org/10.6084/m9.figshare.1041797>
- Bar-Ilan, J. (2014). JASIST@Mendeley Revisited. [fighttp://dx.doi.org/10.6084/m9.figshare.1031681](http://dx.doi.org/10.6084/m9.figshare.1031681)
- Bordons, M., Fernández, M. T., & Gomez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics*, 53(2), 195-206.
- Braam, R. R., Moed, H. F., & van Raan, A. F. (1991). Mapping of science by combined co-citation and word analysis, I. Structural aspects. *Journal of the American Society for Information Science*, 42(4), 233-251.
- Brouwer, E., & Kleinknecht, A. (1999). Innovative output, and a firm's propensity to patent: An exploration of CIS micro data. *Research Policy*, 28(6), 615-624.
- Butler, L. (2003). Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts. *Research Policy*, 32(1), 143-155.
- Coupe, T. (2003). Science is golden: Academic R&D and university patents. *The Journal of Technology Transfer*, 28(1), 31-46.
- Costas, R., Zahedi, Z., & Wouters, P. (2014a). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *arXiv preprint arXiv:1401.4321*.
- Costas, R., Zahedi, Z., & Wouters, P. (2014b). Disentangling the meaning of 'altmetrics': content analysis of Web of Science scientific publications. Figshare. <http://dx.doi.org/10.6084/m9.figshare.1041770>
- Cronin, B., Snyder, H. W., Rosenbaum, H., Martinson, A., & Callahan, E. (1998). Invoked on the Web. *Journal of the American Society for Information Science*, 49(14), 1319-1328.
- Desai, T., Shariff, A., Shariff, A., Kats, M., Fang, X., Christiano, C., & Ferris, M. (2012). Tweeting the meeting: an in-depth analysis of Twitter activity at Kidney Week 2011. *PLoS ONE*, 7(7), e40253.
- Erfanmanesh, M. A., & Didegah, F. (2013). A Comparison of Web of Science and Scopus for Iranian Publications and Citation Impact. *International Journal of Information Science & Management*, 11(1), 11-27.
- 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).
- Haslam, N., & Koval, P. (2010). Possible research area bias in the Excellence in Research for Australia (ERA) draft journal rankings. *Australian Journal of Psychology*, 62(2), 112-114.
- Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. *Scientometrics*, 101(2), 1419-1430.

- Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (in press). Tweets as impact indicators: Examining the implications of automated bot accounts on Twitter. *Journal of the Association for Information Science and Technology*. <http://arxiv.org/abs/1410.4139>
- 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.
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101(2), 1145-1163.
- Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, 5(3), 446-457.
- HEFCE (2013). Citation data: REF 2014. <http://www.ref.ac.uk/subguide/citationdata/>
- Henning, V., & Reichelt, J. (2008). Mendeley - A Last.fm for research? In: *IEEE Fourth International Conference on eScience (eScience'08)*. (pp. 327-328). Menlo Park: IEEE.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of documentation*, 54(2), 236-243.
- Jimenez-Contreras, E., Robinson-Garcia, N., & Cabezas-Clavijo, A. (2011). Productivity and impact of Spanish researchers: reference thresholds within scientific areas. *Revista Española de Documentación Científica*, 34(4), 505-525. <http://eprints.rclis.org/16784/1/Productivity%20and%20impact%20of%20Spanish%20researchers.pdf>
- Kousha, K., & Thelwall, M. (2006). Motivations for URL citations to open access library and information science articles. *Scientometrics*, 68(3), 501-517.
- Kousha, K., & Thelwall, M. (2008). Assessing the impact of disciplinary research on teaching: An automatic analysis of online syllabuses. *Journal of the American Society for Information Science and Technology*, 59(13), 2060-2069.
- Kousha, K. & Thelwall, M. (in press). Can Amazon.com reviews help to assess the wider impacts of books? *Journal of the Association for Information Science and Technology*.
- Kwok, J.T. (2013). Impact of ERA Research Assessment on University Behaviour and their Staff. [http://apo.org.au/files/Resource/nteu\\_impactofera\\_april2013.pdf](http://apo.org.au/files/Resource/nteu_impactofera_april2013.pdf)
- Labrosse, L., Ventimiglia, A., Campbell, D., Haustein, S., Côté, G., & Archambault, É (2013). Scientific Output and Collaboration of European Universities, EC Directorate-General for Research and Innovation [http://ec.europa.eu/research/innovation-union/pdf/scientific\\_output\\_collaboration\\_european\\_univ.pdf#view=fit&pagemode=none](http://ec.europa.eu/research/innovation-union/pdf/scientific_output_collaboration_european_univ.pdf#view=fit&pagemode=none)
- Larivière, V., Archambault, É., & Gingras, Y. (2008). Long-term variations in the aging of scientific literature: From exponential growth to steady-state science (1900–2004). *Journal of the American Society for Information Science and technology*, 59(2), 288-296.
- Levitt, J. M., & Thelwall, M. (2008). Patterns of annual citation of highly cited articles and the prediction of their citation ranking: A comparison across subjects. *Scientometrics*, 77(1), 41-60.
- Levitt, J. M., & Thelwall, M. (2011). A combined bibliometric indicator to predict article impact. *Information Processing & Management*, 47(2), 300-308.
- Li, J., Burnham, J. F., Lemley, T., & Britton, R. M. (2010). Citation Analysis: Comparison of Web of Science, Scopus, SciFinder, and Google Scholar. *Journal of Electronic Resources in Medical Libraries*, 7(3), 196-217.

- Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement, *Scientometrics*, 91(2), 461-471.
- 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.
- Maflahi, N. & Thelwall, M. (in press). When are readership counts as useful as citation counts? Scopus versus 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.
- Minasny, B., Hartemink, A. E., McBratney, A., & Jang, H. J. (2013). Citations and the h index of soil researchers and journals in the Web of Science, Scopus, and Google Scholar. *PeerJ*, 1, e183.
- Moed, H. F. (2006). *Citation analysis in research evaluation*. Berlin: Springer.
- 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*.
- Mohammadi, E. & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology*, 65(8), 1627-1638.
- Parker, J., & van Teijlingen, E. (2012). The Research Excellence Framework (REF): assessing the impact of social work research on society. *Practice*, 24(1), 41-52.
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv preprint arXiv:1203.4745*.
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto. <http://altmetrics.org>
- Priem, J., & Costello, K. L. (2010). How and why scholars cite on Twitter. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1-4.
- Richardson, M. (2011). The value of bibliometrics. *The Research Excellence Framework: Revisiting the RAE*. *Research Trends*, 22. <http://www.researchtrends.com/issue22-march-2011/the-research-excellence-framework-revisiting-the-rae-2/>
- Sainty, M. (2013). Research impact: a United Kingdom Occupational Therapy Research Foundation perspective. *The British Journal of Occupational Therapy*, 76(12), 528-537.
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2012). Research blogs and the discussion of scholarly information. *PloS ONE*, 7(5), e35869.
- Shema, H., Bar-Ilan, J., & 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 Association for Information Science and Technology*, 65(5), 1018-1027.
- Sud, P. & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131-1143.
- Sud, P. & Thelwall, M. (in press). Not all international collaboration is beneficial: The Mendeley readership and citation impact of biochemical research collaboration. *Journal of the Association for Information Science and Technology*.

- Taraborelli, D. (2008). Soft peer review: social software and distributed scientific evaluation. In: Hassanaly, P., Ramrajsingh, A., & Randall, D. and Salembier, P. and Tixier, M., (eds.) Proceedings of the 8th International Conference on the Design of Cooperative Systems, Carry-le-Rouet, 20-23 May 2008. (pp. pp. 99-110). Institut d'Etudes Politiques d'Aix-en-Provence: Aix-en-Provence, France.
- TEC (2014). Performance-Based Research Fund. Sector Reference Group – Consultation Paper #1: Approach to the design of the 2018 Quality Evaluation. <http://www.tec.govt.nz/Documents/Reports%20and%20other%20documents/PBRF-2018-QE-design.pdf>
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. PloS ONE, 8(5), e64841.
- Thelwall, M. & Maflahi, N. (in press). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. Journal of the Association for Information Science and Technology.
- Thelwall, M., Tsou, A., Weingart, S., Holmberg, K., & Haustein, S. (2013). Tweeting links to academic articles, Cybermetrics, 17(1), <http://cybermetrics.cindoc.csic.es/articles/v17i1p1.html>.
- Thelwall, M. & Wilson, P. (in press). Does research with statistics have more impact? The citation rank advantage of structural equation modelling. Journal of the Association for Information Science and Technology.
- Tijssen, R. J. (2001). Global and domestic utilization of industrial relevant science: patent citation analysis of science–technology interactions and knowledge flows. Research Policy, 30(1), 35-54.
- Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. The Rand Journal of Economics, 172-187.
- Vaughan, L., & Hysen, K. (2002). Relationship between links to journal Web sites and impact factors. Aslib Proceedings 54(6), 356-361.
- Vaughan, L., & Shaw, D. (2003). Bibliographic and web citations: what is the difference?. Journal of the American Society for Information Science and Technology, 54(14), 1313-1322.
- Vaughan, L., & Shaw, D. (2005). Web citation data for impact assessment: A comparison of four science disciplines. Journal of the American Society for Information Science and Technology, 56(10), 1075-1087.
- Wang, J. (2013). Citation time window choice for research impact evaluation. Scientometrics, 94(3), 851-872.
- Weller, K., & Puschmann, C. (2011). Twitter for scientific communication: How can citations/references be identified and measured. Proceedings of the ACM WebSci11, 1-4.
- Wouters, P., & Costas, R. (2012). Users, narcissism and control: tracking the impact of scholarly publications in the 21st century. *Proceedings of the 17th International Conference on Science and Technology Indicators* (Vol. 2, pp. 487-497).
- 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.
- Zahedi, Z., Haustein, S. & Bowman, T (2014). Exploring data quality and retrieval strategies for Mendeley reader counts. Presentation at SIGMET Metrics 2014 workshop, 5 November 2014. <http://www.slideshare.net/StefanieHaustein/sigme-tworkshop-assist2014>