Mendeley reader counts offer early evidence of the scholarly impact of academic articles

blogs.lse.ac.uk/impactofsocialsciences/2017/01/17/mendeley-reader-counts-offer-early-evidence-of-the-scholarly-impact-of-acader

1/17/2017

Although the use of citation counts as indicators of scholarly impact has well-documented limitations, it does offer insight into what articles are read and valued. However, one major disadvantage of citation counts is that they are slow to accumulate. ***Mike Thelwall*** *has examined reader counts from Mendeley, the academic reference manager, and found them to be a useful source of early impact information. Mendeley reader data can be recorded from the moment an article appears online and so avoids the publication cycle delays that so slow down the visibility of citations.*

Counts of citations to academic articles are widely used as evidence to inform estimates of the impact of academic publications. This is based on the belief that scientists often cite works that have influenced their thinking and therefore that citation counts are indicators of influence on future scholarship. In the UK's REF2014 research assessment exercise, 11 of the 36 subject panels drew upon citation counts to inform their judgements of the quality of academic publications, for example by arbitrating when two expert reviewers gave conflicting judgements. Citation counts are also widely used internationally for hiring, promotion, and grant applications and aggregated citation-based statistics are used to assess the impact of the work of large groups of scholars in departments, universities and even entire countries. On top of this, there are many informal uses of citation counts by individual scholars looking to assess whether their work is having an impact or to decide which of their outputs is having the most impact.

Image credit: Mendeley Desktop and iOS by Team Mendeley. This work is licensed under a CC BY 2.0 license.

Despite their many limitations, such as obvious cases where they are misleading and entire fields for which they are

almost meaningless, citation counts can support the onerous task of peer review and even substitute for it in certain cases where the volume of outputs is such that peer review judgements are impractical. At the level of the individual scholar, citation counts can be useful to indicate whether papers are read and valued. This gives outputs a visible afterlife once they have been published and helps to identify avenues of research that have been unexpectedly successful, motivating future similar work. It also gives scholars a sometimes-needed incentive to look outwards at the wider community when writing an article and consider how it might attract an audience that might cite it. Of course, uncited does not equate to irrelevant and James Hartley has recently listed his rarely cited articles that he values, which is a useful reminder of this. Nevertheless, even though I have little idea why my most cited article has attracted interest, the knowledge that it has found an audience has motivated me to conduct follow-up studies and to fund PhDs on the subject, whilst dropping lines of research that have disappointingly flown under the radar and (so far) avoided notice.

One major disadvantage of citation counts is that they are slow to accumulate. Once an article has been published, even if someone reads it on the first day that it appears and immediately uses it to inform a new study, it is likely to be 18 months (depending on the discipline) before that study is complete, written up, submitted to a journal, peer reviewed, revised, accepted and published so that its citations appear in Google Scholar, Web of Science or Scopus. Uses of citation counts in formal or informal research evaluations may therefore lag by several years. This delay is a major disadvantage for most applications of citation counts. There is a simple solution that is effective in some contexts: Mendeley reader counts (Figure 1).

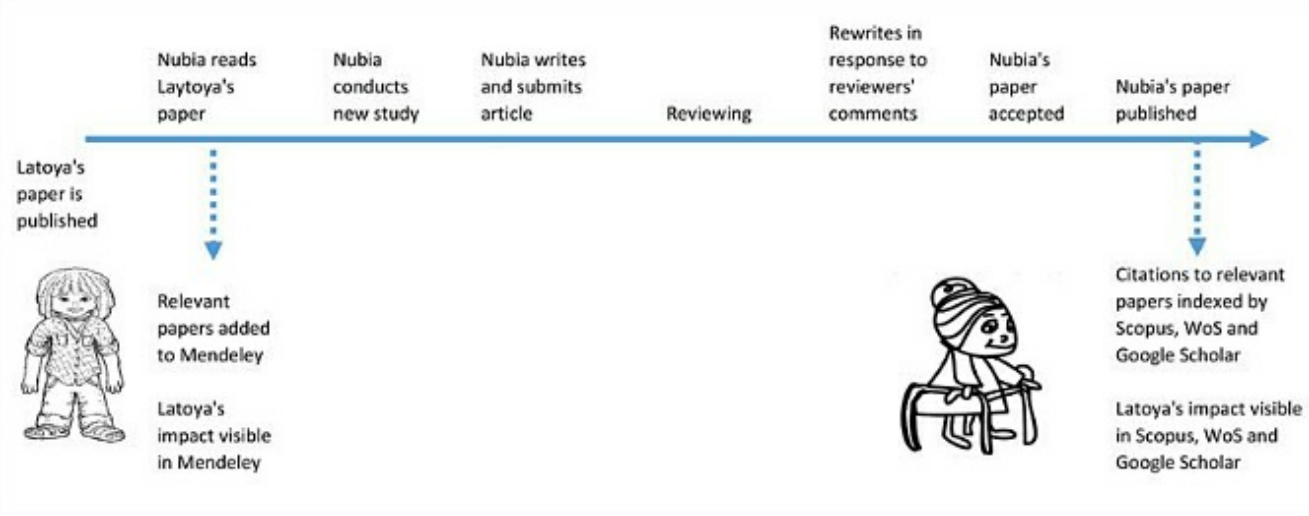

Figure 1: Mendeley readers typically appear at least a year before citations due to delays between other researchers reading a paper and their new study being published.

Mendeley is a social reference sharing website that is free to join and acts as a reference manager and sharer for academics and students. Those using it can enter reference information for articles that they are reading or intend to read (and this is what most users do, as shown by Ehsan Mohammadi, whose PhD focused on Mendeley) and then Mendeley will help them to build reference lists for their papers. As spotted by York University (Toronto) librarian Xuemei Li, it is then possible to count the number of registered Mendeley readers for any given article and use it as impact evidence for that article. This reader count acts like a citation count in that it gives evidence of (primarily academic) interest in articles but readers accrue about a year in advance of citation counts, as shown by a recent article (Figure 2 – see also: Maflahi and Thelwall, 2016; Thelwall and Sud, 2016). Mendeley data is available earlier as scholars can register details of an article they are reading in Mendeley whilst they are reading it, and so this information bypasses the publication cycle delays (Figure 1). An article may even start to accumulate evidence of interest in Mendeley in the week it is published if people recognise it as important and immediately record it in Mendeley for current or future use.

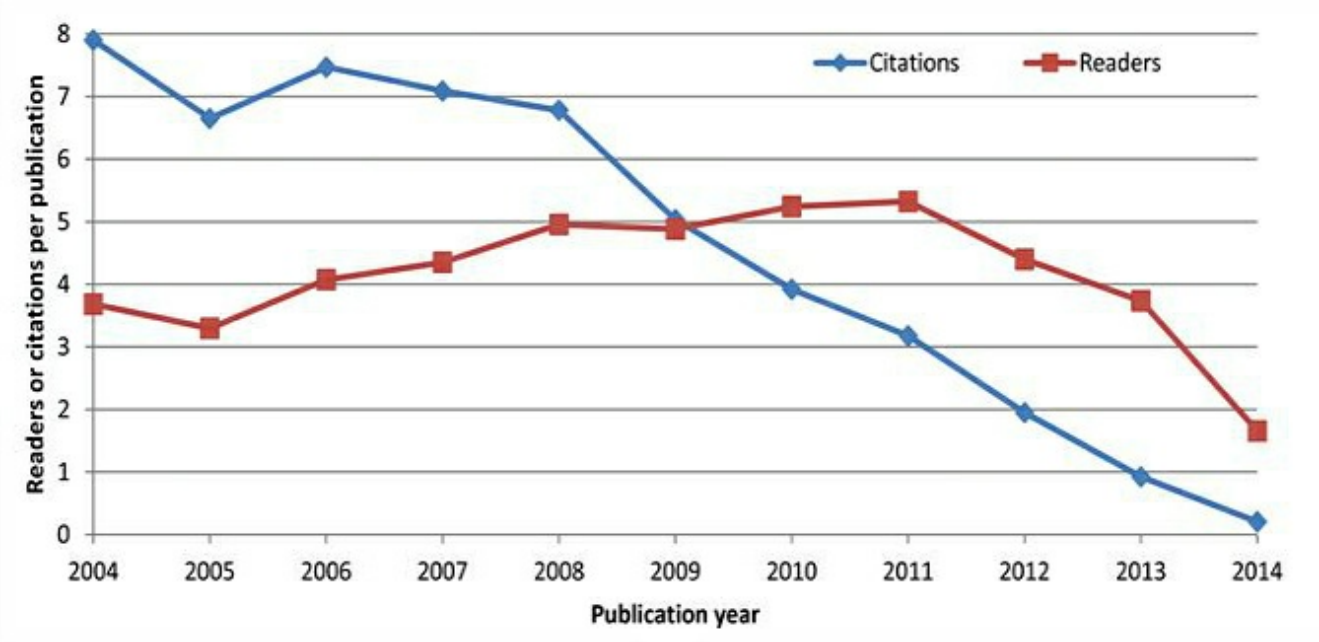

Figure 2: A comparison between average Scopus citations and Mendeley readers for articles from journals in the Scopus Transportation category, as recorded in November/December 2014. Mendeley reader counts are much higher than Scopus citations for more recent articles, with Scopus citations lagging by at least 18 months. Citation counts are higher than reader counts for older articles, probably due to citations from older articles that were written before Mendeley was widely used. Geometric means are used because citation counts are highly skewed (data from Maflahi and Thelwall, 2016).

Mendeley is by far the best general source of early scholarly impact information. Download counts are not widely available, counts of Tweets are very unreliable as an impact indicator and other early impact indicators are much scarcer. The main drawback is that, at present, anyone can set up multiple accounts and register as a reader of selected articles, making it possible to spam Mendeley. For this reason, Mendeley reader counts cannot be used in the UK REF or any other research evaluation that includes stakeholders with time to manipulate the outcomes. An additional limitation is that Mendeley reader counts are biased towards articles that attract the Mendeley user demographic, which has international and seniority/age imbalances. It is therefore tricky to use Mendeley for international impact comparisons.

It is not hard to obtain evidence of Mendeley readers for an article – just search for it by title in Mendeley (e.g. try the query 'Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows') or look for the Mendeley segment within the Altmetric.com donut for the article (as in this example; to find a page like this, Google the article and add 'site:altmetric.com' to the end of your query). For large groups of articles, the free Mendeley API can also be used to automatically download reader counts for large sets of articles via the (also free) software Webometric Analyst. If you already have a set of articles with citation counts, then it is simple to add Mendeley reader count data to it using this software.

This blog post is based on the author's article, co-written with Pardeep Sud, 'Mendeley readership counts: An investigation of temporal and disciplinary differences', published in the Journal of the Association for Information Science and Technology (DOI: 10.1002/asi.23559).

Note: This article gives the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our comments policy if you have any concerns on posting a comment below.

About the author

Mike Thelwall *is Professor of Information Science at the School of Mathematics and Computing, University of Wolverhampton. His research interests include big data: webometrics, social media metrics, and sentiment analysis; developing quantitative web methods for Twitter, social networks, YouTube, and various types of link and impact metrics; conducting impact assessments for organisations, such as the UNDP. His ORCID iD is: 0000-0001-6065-205X.*