

2018

Readership Count and Its Association with Citation: A Case Study of Mendeley Reference Manager Software

S Ravikumar

North-Eastern Hill University, ravikumar@nehu.ac.in

Dohtdong Basbunlang Ms

North-Eastern Hill University, basbunlangdohtdong@gmail.com

Follow this and additional works at: <http://digitalcommons.unl.edu/libphilprac>

 Part of the [Library and Information Science Commons](#)

Ravikumar, S and Basbunlang, Dohtdong Ms, "Readership Count and Its Association with Citation: A Case Study of Mendeley Reference Manager Software" (2018). *Library Philosophy and Practice (e-journal)*. 2012.
<http://digitalcommons.unl.edu/libphilprac/2012>

Readership Count and Its Association with Citation: A Case Study of Mendeley Reference Manager Software

Dr.S.Ravikumar

Assistant Professor

Dept. of Library and Information Science

North Eastern Hill University

ravikumar@nehu.ac.in

Ms.Basbunlang Dohtdong

PG Student

Dept. of Library and Information Science

North Eastern Hill University

basbunlangdohtdong@gmail.com

ABSTRACT:

The present study was carried to find out the association between Mendeley readership count and citation pattern for scholarly articles. This study was carried out with the most prolific authors of 2014 from the four subject domain “Clinical medicine, Microbiology, Molecular Biology and Neuroscience” and around 4886 papers was identified and studied their Mendeley readership count and citation count . It was found that the articles of the most prolific authors in the above subject have a strong positive correlation between Mendeley readership count and citation; ρ value is .715**. The linear relationship for individual subjects was between .626** to .789**, significant at 0.01 level.

Keywords: Citation; Bibliometrics; Altmetrics; New Media; Mendeley; Reference Manager

1.1 INTRODUCTION

Earlier, some Scientists, research administrators, and funding bodies routinely need to evaluate the effectiveness of research for promotion, tenure, funding allocation or broad policy decision. This is difficult because the value of research can take a long time to become apparent and may only be evident to small set of domain experts. An unavoidable problem with citation count, however, is that they cannot be used for recent articles because citation take time to accrue due to the time required publishing and indexing the citing research. This issue can delay research evaluation that harness citation-based indicators for one or more years, slowing down any decision making based on the results. Later through the development of new technology, some research began to promote new ideas and bring it into process in such a way that makes it easy and simple to retrieve the data or information. More recently, the introduction of social media in scholarly communication has created a new area of informetric coined “Altmetrics” (Priemet.al, 2010; Priemand Hemminger, 2010; Priem, 2014). Altmetrics are metrics and qualitative data that are complementary to traditional, citation based metrics. They can include (but are not limited to) peer reviews on faculty of 1000, citation on wikipedia and in public policy document, discussion on research blogs, mainstream media coverage, bookmarks on reference managers like

Mendeley, and mention on social network such as Twitter. Sources from the web, Altmetrics can tell us a lot about how often journal articles and other scholarly output like dataset are discussed and used around the world. Altmetric they are quicker to accumulate than citation-based metrics. They capture more diverse impact than citation based metric. According to (Piwowar, 2013) Altmetrics not only refer to the possibility of expanding the analysis of other impact of traditional scientific output, but also to broaden the study of any “research products” beyond (only) scholarly publication. In the current information age, scientific literature is downloaded in various platforms (e.g. Zotero, CiteULike, research gate, BibSomy). These download count are used to evaluate those articles at various level. In the present study Mendeley Reference Manager Tools is used to study the readership of the sample data from four subjects namely (Clinical Medicine, Neuroscience, Microbiology and Molecular Biology). Mendeley was chosen because Mendeley is a platform for users to manage scholarly references, create online profiles and communicate with peers. Mendeley is a powerful reference management tools, empowering researchers to organize their references. Read, organized and cite all of our research from one library, connect and collaborate with millions of researchers worldwide. Mendeley is a tool for every stage of our career, take control of our research. Turn even the largest set of documents into a full searchable, easy to manage library. As a social impact measure; and the use of Mendeley for article readership count evaluation, Mendeley is one of the very important topics in scientometric and Altmetric studies. Count of readers in social reference sharing sites, such as Mendeley, may be an alternative to citation count when evidence of early impact is needed. This is possible because articles reader may start to be registered in Mendeley on the day that an article is published. Out of all current Altmetrics, count of reader of articles in the reference sharing site Mendeley appear to be the most closely related to citation count.

1.2 METHODOLOGY

The first source that have identified for this study is from The World’s Most Influential Scientific Minds 2014 by Thomas Reuter published by Web of science in which the list of different Authors along with their field and affiliations has been taken for the collections of data (articles). This document contained a list of authors published during the year 2014. This document contained many subjects and many authors but with these we cannot take all the authors for all the subjects so, out of many subjects from the source ; four subjects were selected for the study i.e. on the basis of science subject . Since Science subject are widely downloaded in Mendeley for that we have taken some prominent science subject. From this document the Author’s name, field and affiliation of the particular author were selected and then from Google Scholar (www.google scholar.com) that particular author has been searched for their articles which were published in the year 2014 by different publishers. The articles of the particular authors were taken along with the Citation score and Mendeley readership count from www.altmetrics.com. Manual search results were saved in Excel and results was analyzed using SPSS.

1.3 OBJECTIVES

- To study the association between Mendeley readership count and Citation count.
- To do a comparative study between four subjects (i.e. Clinical Medicine, Molecular Biology, Neuroscience and Microbiology).

1.4 LIMITATION

This study has some limitations which are as follows:

- This study is limited only to one year i.e., 2014.

- This study is limited only to four disciplines (Clinical medicine, Microbiology, Molecular biology and Neuroscience).
- Most of the prolific authors were taken from only one document .i.e. “World’s Most Influential Scientific Mind 2014.

LITERATURE REVIEW

A literature review is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contribution to a particular topic. 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 impact of articles. Altmetrics includes data about usage (e. g. pdf downloads); captures (e. g. Bookmarks); mentions (e. g. in Blogs); social media (e.g. shares on Facebook) and citation (e. g. Scopus) (Cave, 2012). According to Galloway, Pease, and Rauh (2013) “Altmetrics is a fast-moving and dynamic area”. However, the use of alternative metrics to evaluate research is not new. It has a long tradition in scientometrics with the analysis of acknowledgement, patents, mentorships, news articles, and usage in syllabi (Priem, 2014). The use of the Internet for alternative metrics began with “webometrics” (or “cybermetrics”) whereby the number of times a paper was mentioned on the web was counted (Roemer & Borchardt, 2012). These mentions were called “web citations” (Shema et al., 2014b). A collection of Altmetrics from different sources may help to reveal different aspects of the impact of academic articles (Priem, Piwowar, & Hemminger, 2012). This paper builds on a previous study of Zahedi, Costas & Wouters (2013). Mendeley has also become an interesting and rich Altmetrics data source (Zahedi, Costas & Wouters, 2014). Hence, Mendeley is a free social reference sharing site that allows users to register and upload documents that they are interested in and creates reference lists for them (Gunn, 2013; Henning & Reichelt, 2008; Zaugg West, Tateishi, & Randall, 2011). 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 citation count (Costas, Zahedi, & Wouters, 2014). First studies have shown that a large share of users consists of students and postdoctoral researchers; so that a bias towards young scientists exists (Mohammadi et al., in press; Zahedi, Costas & Wouters, 2013). It also recommends relevant articles to its users (Beel, Gipp, Langer, & Breitingner, 2016) and supports information seeking (Alhoori & Furuta, 2011). Out of all current Altmetrics, count of reader of articles in the reference sharing site Mendeley appear to be the most closely related to citation count. The social reference sharing website Mendeley is designed to help users to manage and share their academic references (Henning & Reichelt, 2008). The finding of the literature illustrates that research articles are read by various users inside and outside academia using Mendeley. The variation in correlations between Mendeley readership count and citation received for different types of reader suggested that the meaning of Mendeley readership count depends upon the readers’ occupations. This implies that in some cases Mendeley readership may reflect traditional citation impact but in other cases it may reflect educational uses or impact on applied context. Therefore, Mendeley readership is a promising data source that is different from both citation and raw usage data. However, Mendeley is only one of many reference manager tools and other reference managers (e.g. Endnote, Refworks, and Zotero) also have many users but their data are not publically

available. Thus, Mendeley seems to be the only choice to reveal aspects of the readership of research articles.

DATA ANALYSIS AND FINDINGS:

Data analysis is the process of transforming raw data into usable information in order to add value to the statistical output. All the data that have been extracted from the articles of the World’s most influential scientific minds 2014 found in Web of Science associated with citation and Mendeley, they have been processed and generate in the form of table and graph. For this analysis, Spearman correlations were used to assess the strength of association between the citation count and the readership count. For this analysis, 4886 articles written by 718 authors were analyzed using computational tools.

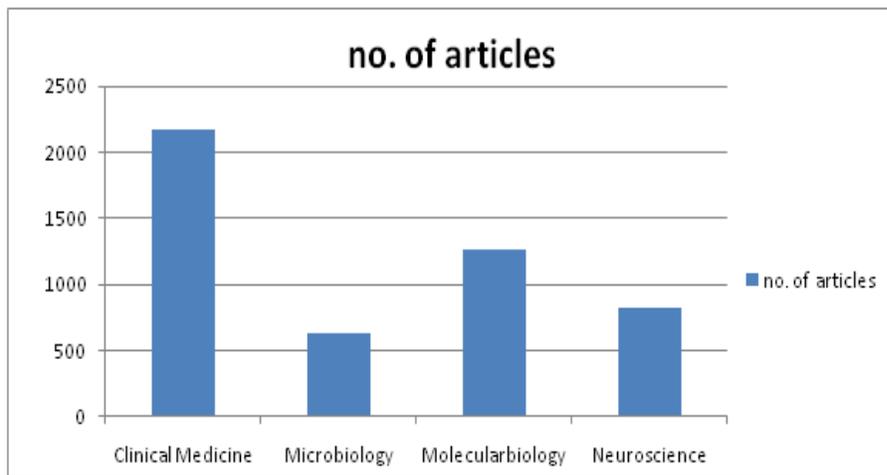


Fig 1: No of articles shared by four disciplines

The above graph shows the number of articles shared by different four subjects (i.e. Clinical medicine, Molecular Biology, Neuroscience, Microbiology). Clinical Medicine has 2169 number of articles followed by molecular biology 1264 articles, Neuroscience 822 articles and Microbiology 626 articles. Among the four subjects Clinical medicine share the highest number of articles and Microbiology share the lowest number of articles.

Spearman Correlation		
	Citation	Mendeley
Correlation Coefficient	1	.715**
Sig. (2-tailed)		0
N	4881	4881

** . Correlation is significant at the 0.01 level (2-tailed)

Table.1: Association between Mendeley readership count and citation count.

The above table explains the correlation between two variables (i.e. Mendeley and citation count) which is significant at 0.01 levels. The ρ value = .715**, which clearly explain that there is a strong correlation between an article getting citation and number of Mendeley readership count it received. The prolific author’s article received 4,56,886 download in Mendeley for total 4886 article and the same article had been cited by 3,33,784 papers.

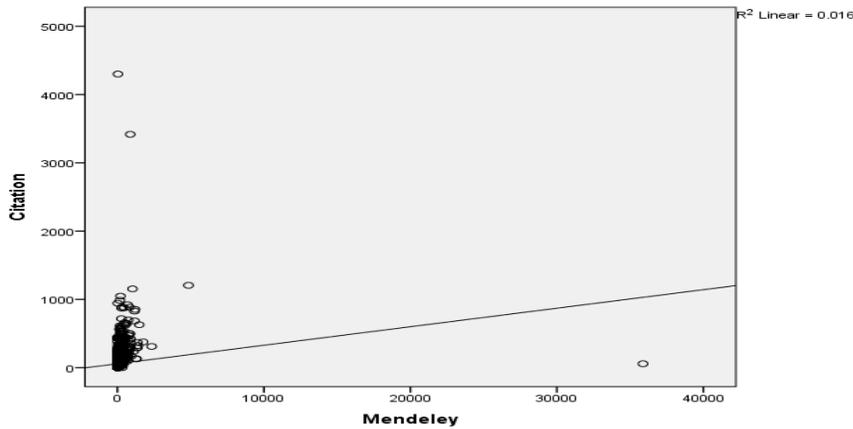


Fig.2: Graph showing the association between Mendeley readership and citation count

The graph above shows the linear relationship between Mendeley and citation count. Most of the data (i.e. the number of Mendeley and number of citations) are close to the fitted regression line and only few are scattered away from the line. There is a positive linear correlation i.e. $R^2=0.016$ suggesting that Mendeley and citation are concentrated in the regression line. It represents the positive linearity between two variables.

2. Comparative study between four subjects (Clinical Medicine, Molecular Biology, Neuroscience and Microbiology).

Spearman’s rank correlation	
Citation Vs Mendeley	
SUBJECTS	CORRELATIONS (ρ)
CLINICAL MEDICINE	.738**
MOLECULAR BIOLOGY	.789**
NEUROSCIENCE	.743**
MICROBIOLOGY	.626**

** . Correlation is significant at the 0.01 level (2-tailed).

Table.2: Spearman rank correlations between Mendeley and citation counts per subjects (i.e Clinical Medicine, Microbiology, Molecular Biology and Neuroscience).

From the above empirical data it was found that the correlations between Mendeley readership count and citation count for all four subjects are statistically significant at 0.01 levels. The correlation for all the disciplines are similar but stronger in Molecular biology which has correlation coefficient value=.789** .It can be seen that from the above table there were statistically significant between Mendeley and citation count for Molecular Biology area which had 1264 papers. Molecular biology received 1,63,874mendeley for total 1264 articles and the same articles cited by 1,01,477papers.Microbiology has the lower correlation coefficient value=.639**.Microbiology had 626 papers, received 39,391Mendeley for total 622 articles and the same articles cited by 27,047 papers. Compare to Clinical Medicine which has correlation coefficient value=.738**, Clinical Medicine had 2169 papers, received 2,15,715 Mendeley for total 2169 articles and the same article cited by 1,67,354 papers. Neuroscience correlation coefficient value=.743**. Citation count for Neuroscience which had 822 papers, received 75,932 download in Mendeley and for the same article received 37,906 citations. Among the four subjects, the correlation in the fields of Microbiology is lesser compared to the correlation in Molecular Biology and others. In most sub dataset, the correlation coefficient ranges from .639** to .789**. This study reveal that the correlation between citation count and Mendeley readership count mentioned is significantly correlated around 0.715**, which is almost the same when compared in general.

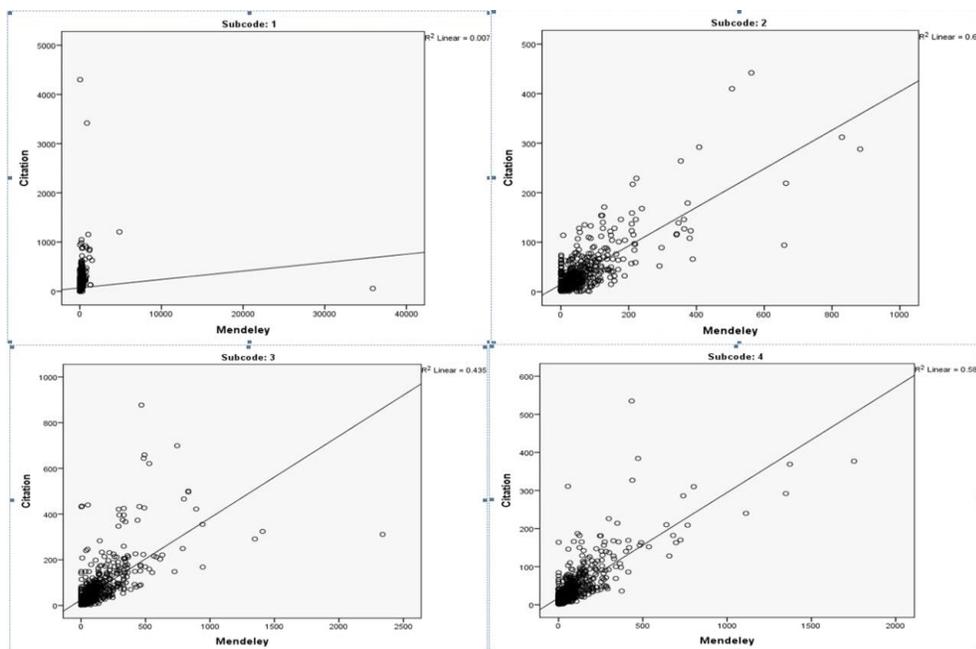


Fig 3: Graph showing Spearman linear correlations

between Mendeley and citation counts for individual subjects (i.e Clinical Medicine, Microbiology, Molecular biology and Neuroscience).

From the above graph of all the four subject linear relationship between Mendeley and citation count is explained .The highest linearity was in Molecular Biology ($R^2=0.601$) followed by Neuroscience ($R^2=0.589$), Microbiology ($R^2=0.435$), and Clinical Medicine ($R^2=0.007$) have lowest number of citation

and Mendeley. In Clinical Medicine the number of Mendeley and citation are close to the regression line and only few numbers are in the periphery. There is a weak linear relationship but statistically significant i.e. $R^2=0.007$. In Molecular Biology most of the number of Mendeley and citation are focus to the regression line and few are scattered away from the line. As $R^2=0.601$ which mean a strong positive linear relationship is existing.

5.2 DISCUSSION

From the sample data it was found that Mendeley readership count is more than citations per subjects. It is seen that there were strong and positive correlation between Mendeley readership count and citation count for all subject domain (i.e. Clinical medicine, Molecular biology, Neuroscience, Microbiology) which was statistically significant. Results above show that the density of Mendeley readership is higher than that of citation for the most recent years and for most of the disciplines. These results suggest the potential advantage of Mendeley readership over citation for the analysis of impact of the most recent publications and particularly in the field of Clinical medicine. (Thelwall & Sud, 2015) suggested that the faster uptake and the stronger density of Mendeley reader count for the most recent years could be seen as a good proxy for “early scientific impact” for articles from recent years and also for fields with higher levels of Mendeley use. These results suggest that Mendeley readership count can work as an important source to reflect evidence of “early impact” of scientific publications since, as shown in this study as well as in a previous analysis (Thelwall & Sud, 2015), readership occur and are available earlier than citation during the first years after publication.

6.1 CONCLUSION

From the present study it is concluded that Mendeley is widely accepted Almetric variable by major readers and it was found that the respective or irrespective of different subject domain downloaded in Mendeley superior than the citation count. From the data it is seen that more the readership in Mendeley will leads to more citation for a document or articles it is clear that if an article is downloaded and read in Mendeley the chances of citing the article is more. So, from the study it is concluded that more the readership in Mendeley will leads to more citation.

6.2 SUGGESTION

From the present study it is found that Mendeley readership count and citation count is strongly correlated but there are good numbers of reference manager software are available in public domain like CiteUlike, Refworks, Bibsomy. Till now there is no major comparative study has been done between Mendeley and other reference manager software like CiteUlike, Refworks, and Endnote etc. We suggest that a comparative study between all the reference manager software and its association with citation count will give a clear picture about utilizations and acceptance of Reference Manager tools.

REFERENCE

- Adams, J. (2005). Early citation counts correlate with accumulated impact. *Scientometrics*, 63(3), 567-581.
- 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.

- Alhoori, H., & Furuta, R. (2011). Understanding the dynamic scholarly research needs and behavior as applied to social reference management. In *International Conference on Theory and Practice of Digital Libraries* (pp. 169-178). Springer Berlin Heidelberg.
- Beel, J., Gipp, B., Langer, S., & Breitinger, C. (2016). Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 17(4), 305-338.
- Cronin, B., Sugimoto, C. R. (ed.), (2014). *Beyond bibliometrics: harnessing multidimensional indicators of scholarly impact*. Cambridge/ London: MIT Press.
- 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.
- Galloway, L. M., Pease, J. L., & Rauh, A. E. (2013). Introduction to altmetrics for science, technology, engineering, and mathematics (STEM) Librarians. *Science & Technology Libraries*, 32(4), 335–345, doi: 10.1080/0194262X.2013.829762.
- Haustein, S., & Larivière, V. (2014). A multidimensional analysis of Aslib proceedings – using everything but the impact factor. *Aslib Journal of Information Management*, 66(4), 358–380. <https://doi.org/10.1108/AJIM-11-2013-0127>.
- Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, 5(3), 446-457.
- Henning, V., & Reichelt, J. (2008). Mendeley—A Last.fm for research? In G. Fox (Ed.), *IEEE Fourth International Conference on eScience (eScience'08)* (pp. 327–328). Menlo Park, CA: IEEE.
- Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement, *Scientometrics*, 91(2), 461-471.
- Levitt, J. M., & Thelwall, M. (2011). A combined bibliometric indicator to predict article impact. *Information Processing & Management*, 47(2), 300-308.
- Maleki, A. (2015). Mendeley readership impact of academic articles of Iran. In *Proceedings of the Society for Scientometrics and Informetrics (ISSI2015)*. Boğaziçi University Printhouse: Istanbul, Turkey (pp. 109-110).
- 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.

- 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*.
- Mohammadi, E., Thelwall, M. &Kousha, K. (2016). Can Mendeley bookmarks reflect readership? A survey of user motivations. *Journal of the Association for Information Science and Technology*. 67(5), 1198-1209. doi:10.1002/asi.23477.
- Mas-Bleda, A., Thelwall, M., Kousha, K. and Aguillo, I.F. (2014) .Do highly Cited researchers successfully use the social web? *Scientometrics*, Vol101 (1): 337-356.
- 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.
- Ortega, J. L. (2015). Disciplinary differences in the use of academic social networking sites.*Online Information Review*, 39(4): 520 - 536.
- Priem, J., Taraborelli, D., Groth, P. &Neylon, C. (2010). *Alt-metrics: A manifesto*. Available from<http://altmetrics.org/manifesto>.
- Priem, J., Piwowar, H., &Hemminger, B. (2012b). *Altmetrics in the wild: Using social media to explorescholarly impact*. Retrieved from <http://arxiv.org/html/1203.4745v1>.
- 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*.
- Roemer, R. C., &Borchardt, R. (2012). From bibliometrics to altmetrics: A changing scholarly landscape. *College & Research Library News*, 39(November), 8–9. https://doi.org/10.1300/J477v01n02_11.
- 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*.
- Thelwall, M., Haustein, S., Lariviere, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. *PLOS ONE*, 8(5),e64841. doi:10.1371/journal.pone.0064841.
- Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, 94(3), 851-872.
- Zahedi, Z., Costas, R., &Wouters, P. (2014). Assessing the Impact of Publications Saved by Mendeley Users: Is There Any Different Pattern Among Users? *Proceedings of the IATUL Conferences*, 1–13. <https://doi.org/10.13140/2.1.1528.1280>.
- Zahedi, Z., Costas, R., &Wouters, P. (2017). Mendeley readership as a filtering tool to identify highly cited publications. *Jasist*, 2017, 1–18. <https://doi.org/10.1002/asi.23883>.
- 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.
- Zaugg, H., West, R. E., Tateishi, I., & Randall, D. L. (2011). Mendeley: Creating communities of scholarly inquiry through research collaboration. *Tech Trends*, 55(1), 32-36.

