



# On the relationships between bibliographic characteristics of scientific documents and citation and Mendeley readership counts: A large-scale analysis of Web of Science publications<sup>☆</sup>



Zohreh Zahedi<sup>a,\*</sup>, Stefanie Haustein<sup>b</sup>

<sup>a</sup> Centre for Science and Technology Studies, Leiden University, P.O. Box 905, 2300 AX Leiden, The Netherlands

<sup>b</sup> École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, H3C 3J7, Canada

## ARTICLE INFO

### Article history:

Received 16 June 2017

Received in revised form 9 December 2017

Accepted 12 December 2017

### Keywords:

Mendeley readership

WoS citation

Bibliographic characteristics

Document types

## ABSTRACT

In this paper we present a first large-scale analysis of the relationship between Mendeley readership and citation counts with particular documents' bibliographic characteristics. A data set of 1.3 million publications from different fields published in journals covered by the Web of Science (WoS) has been analyzed. This work reveals that document types that are often excluded from citation analysis due to their lower citation values, like editorial materials, letters, news items, or meeting abstracts, are strongly covered and saved in Mendeley, suggesting that Mendeley readership can reliably inform the analysis of these document types. Findings show that collaborative papers are frequently saved in Mendeley, which is similar to what is observed for citations. The relationship between readership and the length of titles and number of pages, however, is weaker than for the same relationship observed for citations. The analysis of different disciplines also points to different patterns in the relationship between several document characteristics, readership, and citation counts. Overall, results highlight that although disciplinary differences exist, readership counts are related to similar bibliographic characteristics as those related to citation counts, reinforcing the idea that Mendeley readership and citations capture a similar concept of impact, although they cannot be considered as equivalent indicators.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

### 1.1. Effect of document characteristics on citation impact

Measuring research impact using citation analysis has a long tradition in the field of scientometrics. Today, citation-based indicators are widely used and play a central role in the evaluation of scientific works. Despite their *de facto* use as proxies of scientific quality, citations are not able to fully capture the use and influence of scientific papers (MacRoberts & MacRoberts, 2017; Moed, 2005). Bibliometric research has also shown that a variety of factors can influence citation counts (Larivière & Gingras, 2011; Opthof & Leydesdorff, 2010; Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011). Such factors

<sup>☆</sup> The peer review process of this paper was handled by Staša Milojević, Associate Editor of Journal of Informetrics.

\* Corresponding author.

E-mail addresses: [z.zahedi.2@cwts.leidenuniv.nl](mailto:z.zahedi.2@cwts.leidenuniv.nl) (Z. Zahedi), [stefanie.haustein@umontreal.ca](mailto:stefanie.haustein@umontreal.ca) (S. Haustein).

include, the document types and age of publications, their number of pages, the length of their titles and reference lists (Bornmann & Leydesdorff, 2015; Bornmann, Leydesdorff, & Wang, 2014; Vieira & Gomes, 2010); their different theoretical or methodological approaches (Antonakis, Bastardoz, Liu, & Schriesheim, 2014); whether they are open access (Hajjem, Harnad, & Gingras, 2006); the citation propensity of their fields and their interdisciplinarity (Yegros-Yegros, Rafols, & D'Este, 2015); or the Impact Factor of their publication journal (Boyack & Klavans, 2005).

Numerous previous studies have analyzed whether citation impact is affected by various document characteristics. These studies have explored different characteristics at the article, journal, and author levels using correlation and regression analyses. For example, in the Natural, Life, and Health sciences (Thelwall, 2017a), papers with unusual and obscure titles were associated with lower citation impact. Mixed results were found regarding the effect of title length (Jacques & Sebire, 2010; Stremersch, Camacho, Vanneste, & Verniers, 2015), or titles that included non-alphanumeric characters such as hyphens or colons (Buter & van Raan, 2011; Haslam et al., 2008; Nair & Gibbert, 2016). Based on the assumption that longer articles with longer reference lists may reflect in-depth analysis and diversity of ideas, the number of pages and references have also been analyzed as factors that may affect citation counts (Fox & Boris, 2016). The results showed that papers with more references and more pages tended to get more citations (Ajiferuke & Famoye, 2015; Davis, 2011). Similarly, the number of authors, institutes, and countries involved in a given publication may indicate the extent of collaboration, which is again assumed to increase citation impact. However, results regarding the effect of collaboration on citation rates are mixed (for an overview see Onodera & Yoshikane, 2015) as regards variations by country of collaboration (Thelwall & Sud, 2016), level of collaboration (e.g. whether national, international, intra/inter institutional) (Leimu & Koricheva, 2005), or authors and disciplines (Williams, Stevenson, Nicholas, Watkinson, & Rowlands, 2009). For a recent review of studies analyzing factors affecting citation counts we refer to Tahamtan, Safipour Afshar, & Ahamdzadeh (2016).

### 1.2. Effect of document characteristics on social media visibility

In the context of recently introduced altmetrics—or, more specifically, its subset of social media based metrics (e.g., Facebook, Twitter, blogs, Wikipedia, Mendeley)—the effect to which some factors influence social media activity remains understudied. One large-scale study examining the effect on social media metrics of typical document characteristics (including document type, discipline, number of pages, title length, number of references, and collaboration patterns) conducted by Haustein, Costas, & Larivière (2015). This study was based on Altmetric.com and Web of Science data and found that although effects were weaker than for citations, documents were more likely to be tweeted if they had longer reference lists and involving a greater number of authors, institutes, and countries. Correlations between social media metrics and document characteristics were, however, quite low to non-existent, which was mostly due to the skewed nature of social media events related to journal articles, with most of them having no metrics at all. Social media metrics (particularly Facebook and Twitter counts) correlated mostly among each other, indicating a circular relationship (Bourdieu, 1998), meaning that being picked up by one social media increases the chances of being picked up by another one. Haustein, Costas et al. (2015) also found that news items and editorials were among the most tweeted document types, which indicates that outputs that contain more condensed, novel, opinion-based and easy-to-understand pieces tend to be more popular on Twitter. The results contrast with the citation patterns for these types of documents, which are substantially less cited than articles and reviews. Overall, the study by Haustein, Costas et al. (2015) showed that characteristics that typically are related to higher citation counts had a smaller relationship with social media counts, sometimes even in an entirely different manner (for instance, longer titles were associated with higher citation counts but with lower Twitter mentions).

### 1.3. Mendeley readership and citation counts

Mendeley is an online reference manager that allows users to save documents in their own libraries and share their libraries with others. Statistics about how often a particular document is saved are made available via the Mendeley API as 'readership' counts. While this count is described by Mendeley as 'readership', it does not actually indicate that the user who saved the document has actually 'read' it, but simply that the user has saved the reference in the library. As such, Mendeley 'saves' are seen more as acts of access to documents than of their appraisal (Haustein, Bowman & Costas, 2016), indicating that the level of engagement captured by these acts is very low.

However, Mendeley has been identified as the most prevalent and noteworthy altmetric source. It has been found that readership counts often exceed citations, and that there is a high representation of recent publications on the platform (Thelwall & Sud, 2015). Compared to other altmetric indicators, Mendeley readership counts were shown to have moderate to strong correlations with citation counts (for a review see Sugimoto, Work, Larivière, & Haustein, 2017), which reflects a greater similarity with citations than other altmetric indicators (Costas, Zahedi, & Wouters, 2015a). This can be explained by the large numbers of academic users in Mendeley, and the frequent use of Mendeley in a pre-citation context (Mohammadi, Thelwall, & Kousha, 2016). The number of Mendeley users who have added an article to their libraries has been suggested as an early indicator of citation impact (Thelwall & Sud, 2015), and Mendeley itself has been identified as a relevant tool with which to identify highly cited publications (Zahedi, Costas, & Wouters, 2017). Mendeley readership distributions have also been shown to be very similar to citation distributions (Costas, Haustein, Zahedi, & Larivière, 2016), and it has been suggested that field-normalized readership scores could be calculated in a similar fashion as for citations (Bornmann & Haunschild, 2016).

Download English Version:

<https://daneshyari.com/en/article/6934196>

Download Persian Version:

<https://daneshyari.com/article/6934196>

[Daneshyari.com](https://daneshyari.com)