Universiteit Leiden
The Netherlands

A multi-dimensional analysis of usage counts, Mendeley readership, and citations for journal and conference papers

Wencan, T.; Fang, Z..; Xianwen, W.; Costas Comesana, R.

Citation

Wencan, T., Fang, Z., Xianwen, W., & Costas Comesana, R. (2024). A multi-dimensional analysis of usage counts, Mendeley readership, and citations for journal and conference papers. *Scientometrics*, *129*, 985–1013. doi:10.1007/s11192-023-04909-w

Version:	Publisher's Version
License:	Licensed under Article 25fa Copyright Act/Law (Amendment Taverne)
Downloaded from:	https://hdl.handle.net/1887/3759586

Note: To cite this publication please use the final published version (if applicable).

Scientometrics (2024) 129:985–1013
https://doi.org/10.1007/s11192-023-04909-w

A multi-dimensional analysis of usage counts, Mendeley readership, and citations for journal and conference papers

Wencan Tian[1] · Zhichao Fang[2,3] · Xianwen Wang[1] · Rodrigo Costas[3,4]

Received: 29 May 2023 / Accepted: 4 December 2023 / Published online: 10 January 2024
© Akadémiai Kiadó, Budapest, Hungary 2024

Abstract
This study analyzed 16,799 journal papers and 98,773 conference papers published by IEEE Xplore in 2016 to investigate the relationships among usage counts, Mendeley readership, and citations through descriptive, regression, and mediation analyses. Differences in the relationship among these metrics between journal and conference papers are also studied. Results showed that there is no significant difference between journal and conference papers in the distribution patterns and accumulation rates of the three metrics. However, the correlation coefficients of the interrelationships between the three metrics were lower in conference papers compared to journal papers. Secondly, funding, international collaboration, and open access are positively associated with all three metrics, except for the case of funding on the usage metrics of conference papers. Furthermore, early Mendeley readership is a better predictor of citations than early usage counts and performs better for journal papers. Finally, we reveal that early Mendeley readership partially mediates between early usage counts and citation counts in the journal and conference papers. The main difference is that conference papers rely more on the direct effect of early usage counts on citations. This study contributes to expanding the existing knowledge on the relationships among usage counts, Mendeley readership, and citations in journal and conference papers, providing new insights into the relationship between the three metrics through mediation analysis.

Keywords Altmetrics · Usage metrics · Mendeley readership · IEEE Xplore · Conference papers

✉ Xianwen Wang
xianwenwang@dlut.edu.cn

[1] WISE Lab, Institute of Science of Science and S&T Management, Dalian University of Technology, Dalian, China

[2] School of Information Resource Management, Renmin University of China, Beijing, China

[3] Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands

[4] DSI-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy, Stellenbosch University, Stellenbosch, South Africa

Introduction

The evaluation of scholarly articles has always been a fundamental element in research evaluation, being essential in evaluating researchers for career advancement, project applications, or just their academic reputation (Breitzman, 2021). Traditionally, citation-based metrics have been used to evaluate the impact of individual publications (Waltman, 2016). However, these metrics have increasingly come under scrutiny due to their complex motives and citation delay phenomena (Cui et al., 2023; Khan & Younas, 2017; Wang et al., 2016), as well as other more fundamental issues like their meaning or conceptual ambiguity (Hicks et al., 2015).

In this context, altmetrics garnered considerable attention in recent years (Chen et al., 2020; Chi et al., 2019; Costas et al., 2015; Geng et al., 2022) since they offered potentially "alternative" evaluation metrics for scientific publications by capturing social media attention, online discussions, and other forms of non-traditional impact, thus potentially supplementing traditional citation-based metrics (Erdt et al., 2016; Sugimoto et al., 2017). Altmetrics, or social media metrics (Haustein et al., 2016), were therefore expected to provide faster metrics for research evaluation and, to some extent, mitigate the citation delay issue (Khan & Younas, 2017; Wang et al., 2016). However, previous studies have demonstrated the unsuitability of social media metrics, mostly based on social media events (e.g., Twitter, Facebook), for research evaluation or even predicting citations (Fang et al., 2020a, 2020b; Haustein et al., 2014; Zahedi & Costas, 2018). The most notable exception is the case of Mendeley readership, the only altmetric source to show a moderate correlation with citations (Fang et al., 2020a, 2020b; Zahedi et al., 2014) and similar distribution properties as citations (Costas et al., 2017). Another important source of metrics for scientific publications is *usage metrics*, which have been researched for even longer than altmetrics (Glaenzel & Gorraiz, 2015) and also provide relevant data on how individual scientific publications are being viewed or downloaded (typically referring to the number of HTML views or PDF downloads) by different users.

Mendeley readership and usage counts, together with the number of (re)tweets of articles, are the metrics that have been studied more often in their relationships with citations (Fang et al., 2021, 2022; Mohammadi et al., 2015; Wang et al., 2014). However, a combined study of usage counts, Mendeley readership and citations, particularly paying attention to the potential mediation effects among them, have never been done before. This study aims at better understanding the relationship between the first two (usage counts and Mendeley readership) and citations, and comprehensively analyzes journal and conference papers' similarities and differences.

Relationship between usage counts and citations

In the extant literature on the interplay between usage counts and citations, most research focused on exploring the correlation between these two metrics. Such inquiries aimed at developing novel academic evaluation metrics (such as the ("usage impact factor" Bollen & van de Sompel, 2008; Schloegl & Gorraiz, 2010), the "usage immediacy index" (Rowlands & Nicholas, 2007), or the "download immediacy index" (Wan et al., 2010)) to assess individual researchers, academic journals, and institutions' research capabilities. Scholars have conducted a considerable amount of related research on different disciplines, which yielded different conclusions. Most previous studies pointed to a positive correlation

between usage and citation counts (Bollen & van de Sompel, 2008; Bollen et al., 2003; Chi & Glänzel, 2018; McGillivray & Astell, 2019). For instance, in the field of chemistry, the usage and citation data on the Web of Science platform showed a moderate correlation (Chi et al., 2019). Lippi and Favaloro (2013) identified a strong correlation between article rankings based on downloads and the relative number of ScienceDirect citations. Chi and Glanzel (2017) conducted a comparative analysis of usage and citation data of articles published in Web of Science from three countries, Belgium, Israel, and Iran, and found that citations and usage counts are significantly correlated, particularly in the social sciences. Furthermore, Ding et al. (2021) verified the Granger causality relationship between usage and citation counts in more than 7000 articles published in *The Lancet*.

In addition to examining the correlation between usage and citation counts at the article level, some scholars have also investigated the relationship between these two metrics at the journal level, finding that the correlation is stronger at the journal level (Guerrero-Bote & Moya-Anegón, 2014; Schloegl & Gorraiz, 2010; Vaughan et al., 2017). Furthermore, papers published in non-English-language journals exhibit a higher correlation between usage counts and citations than those published in English-language journals (Guerrero-Bote & Moya-Anegón, 2014). In addition, to address the research gap in exploring the relationship between usage counts and citations using Chinese databases, Vaughan et al. (2017) used a sample of 150 journals from the China Academic Journal Network Publishing Database and demonstrated a strong correlation between usage counts and citations at the journal level.

Relationship between Mendeley readership and citations

The relationship between Mendeley readership and citations has been investigated in several studies. It is widely acknowledged within the scientific community that a positive correlation exists between these two metrics (Zahedi & Haustein, 2018; Zahedi et al., 2017). For example, Li et al. (2012) analyzed 1613 articles published in *Nature* and *Science* and reported a significant correlation between Mendeley readership and WoS citations. In another study, Thelwall (2017b) compared the correlation between Mendeley readership and Scopus citations for journal articles in 325 narrow Scopus fields and found a strong positive correlation in most fields, with an average correlation coefficient of 0.671, and a higher correlation in social sciences than in humanities (Mohammadi & Thelwall, 2014). Notably, the proportion of researchers who actually read the articles saved in Mendeley may affect the reliability of the aforementioned results. To address this issue, Mohammadi et al. (2016) surveyed 860 Mendeley users and found that 55% of users had created a personal library in Mendeley and claimed to have read or planned to read at least half of the academic articles included in their library. Moreover, 85% of respondents indicated that using Mendeley had facilitated their future citation work. Furthermore, having an online distribution channel for the article also increases its usage on Mendeley, thereby boosting the article's citation counts (Kudlow et al., 2017).

Aduku et al. (2017) and Thelwall (2020) expanded the relationship between Mendeley readership and citations from journal articles to conference papers. Aduku et al. (2017) investigated the associations between Mendeley readership and Scopus citations for both journal articles and conference papers in Scopus from 2011, spanning four disciplines (Computer Science Applications, Computer Software, Building & Construction Engineering, and Industrial & Manufacturing Engineering). They found that conference papers in the latter two disciplines exhibited a lower correlation between Mendeley

readership and Scopus citations. However, Thelwall (2020) subsequently validated the high correlation between Mendeley readership and Scopus citation counts for conference papers in all 11 different disciplines within the field of computer science, using a larger dataset and a broader time span.

Relationship between usage counts and Mendeley readership

Compared to the two types of relationship research mentioned above, research on the relationship between usage counts and Mendeley readership has been relatively limited. Existing research suggests that the correlation between these two metrics is generally low to moderate. For instance, Schloegl et al. (2014) studied articles published in the *Journal of Strategic Information Systems* and *Information and Management* and discovered correlation coefficients between the two metrics of 0.73 and 0.66, respectively. In a separate investigation, Thelwall and Kousha (2017) revealed that the correlation coefficient between the two metrics was distributed between 0.2 and 0.4. Additionally, Wang et al. (2020) found that the correlation coefficient between usage counts and Mendeley readership did not exhibit a considerable difference between preprints and non-open access papers, with both coefficients distributed between 0.18 and 0.52.

Objectives

Based on the literature review presented above, it is evident that most studies examining the correlation among the three metrics have primarily utilized data derived from journal articles, while limited attention has been given to conference papers. Although Aduku et al. (2017) and Thelwall (2020) have already compared the relationship between Mendeley readership and citations in journal articles and conference papers using the Scopus database, it remains unclear whether the conclusions are applicable to other databases. Moreover, it is still necessary and important to take usage counts into account and conduct further investigations into deeper path analysis or interaction mechanism analysis of the relationship among these three metrics. Therefore, this paper employs multiple perspectives, including descriptive, regression, and mediation analyses, to investigate the similarities and differences in the relationship among the three metrics in both journal and conference papers. The specific research questions in this study are as follows.

1. RQ1. Are the distribution patterns, correlations, and accumulation rates of usage counts, Mendeley readership, and citations in journal papers and conference papers similar? Furthermore, what is the relationship between funding, international collaboration, and open access with these three metrics? Does this relationship differ between journal and conference papers?
2. RQ2. Which metric, early usage counts or early Mendeley readership, performs better in predicting the future academic citation impact of articles?
3. RQ3. Does early Mendeley readership mediate the relationship between early usage counts and later citation impact? And if so, how does this mediating effect differ between journal papers and conference papers?

Data and methods

Dataset

Data were obtained from articles published in the IEEE Xplore database in 2016, including 16,799 journal papers (published in 134 distinct journals, 98% of journal publications are of the "article" type) and 98,773 conference papers. There are primarily two reasons for choosing this database. First, the IEEE Xplore database is a professional electronic, electrical, and computer engineering database (Khan & Younas, 2017; Tian et al., 2019). It encompasses a rich collection of journal articles and conference proceedings related to these fields, thereby providing us with substantial conference paper data. Second, in contrast to Web of Science and Scopus databases, which only offer the total usage count for articles, IEEE Xplore has been providing monthly usage data (total number of PDF downloads and HTML views, as seen in the example in Fig. 1) for each item since 2011 (Breitzman, 2021). This unique feature of dynamic usage data aligns more closely with the focus of our research inquiry.

Using web crawling techniques, we obtained the annual usage data for the articles mentioned above between 2016 and 2020 (specific details can be found in Sect. "The process of obtaining article usage counts"), including 83,995 usage data points generated by journal

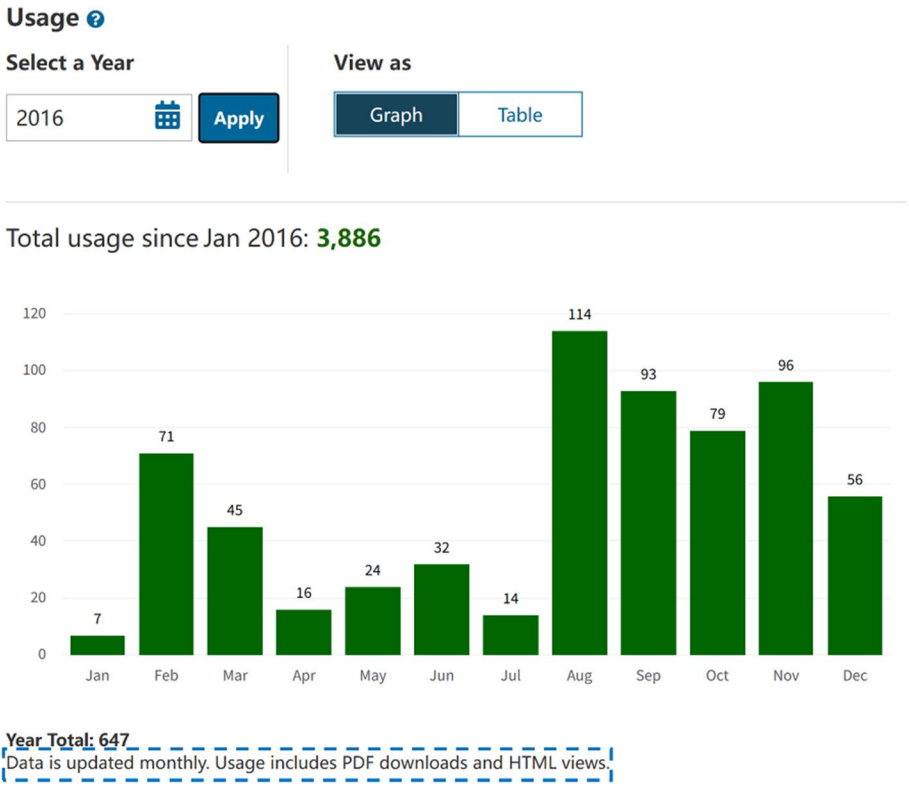

Fig. 1 A screenshot of a usage page from an article in IEEE Xplore (https://doi.org/10.1109/COMST.2016.2521642)

papers and 493,865 usage data points generated by conference papers. The data acquisition period was in May 2021, lasting approximately one month. In addition, we retrieved the annual Mendeley readership and citation counts for the publications above between 2016 and 2020 from the CWTS in-house database. The in-house Mendeley database is created based on the Mendeley readership data that were collected by using the Mendeley API on a yearly basis (in July) between 2016 and 2020. The in-house Dimensions database is created based on the snapshot files provided directly by Digital Science. The version June, 2022 of the Dimensions was used in this study to obtain the annual citation counts of the publications. As a result, for each publication in our dataset, we calculated the usage counts, Mendeley readership, and citations for each year within the observation time window of 2016–2020.

The process of obtaining article usage counts

IEEE Xplore provides a way to create retrieval queries by limiting the time range. For instance, to retrieve data added to the database on January 1, 2016, we can construct a link like this: https://ieeexplore.ieee.org/search/searchresult.jsp?action=search&newsearch=true&ranges=20160101_20160101_Search%20Latest%20Date. By simply altering the values within the "ranges" parameter, we can generate 366 retrieval links to obtain metadata for all publications added to the database in 2016.

Within the metadata of each retrieved publication, the "PDF link" field provides the "articleNumber" (equivalent to a unique ID assigned by the IEEE Xplore database to each publication). Using this "articleNumber", we can create a link to access the publication's usage data. For example, if a publication's PDF link is https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7393668, its "articleNumber" is 7,393,668, and the link to access the publication's usage data would be https://ieeexplore.ieee.org/rest/document/7393668/metrics (see Fig. 2). We access this link and save the data displayed on that page in JSON format.

Finally, parsing and processing the obtained JSON files allows us to obtain annual usage data for each publication. It is worth noting that we excluded publication types such as "Standards", "Books", "Courses", and others, retaining only data from

Fig. 2 Example of usage data for a publication (https://doi.org/10.1109/COMST.2016.2521642). The content in the first box is the access link for the usage data of the publication with the article number "7,393,668." The content in the second box represents the monthly usage data in JSON format generated by this publication since its addition to the IEEE Xplore database

publication types "Conferences" and "Journals." Following further data refinement, which included the removal of 8542 records (8422 conference papers and 120 journal papers) lacking DOIs and those that did not successfully match in both the Mendeley and Dimensions databases via DOI, and the removal of 536 records published in 12 journals lacking Journal Impact Factors (JIFs), we ultimately retained data of a total of 16,799 journal papers and 98,773 conference papers.

Introduction of variables

To address the study's second research question, which investigates the ability of early usage counts and early Mendeley readership to predict future citation counts of an article, *early usage counts* and *early Mendeley readership* are operationalized as the cumulative counts within the first two years after publication, whereas *future citations* are measured by the cumulative counts within five years after publication. Additionally, the study controlled for several potential confounding variables (Cui et al., 2023), such as whether the article received funding, involved international collaboration, was open access, and the journal impact factor (JIF). Funding, international collaboration, and open-access information were obtained directly from the Dimensions database hosted at CWTS. The JIF information was obtained from the Journal Citation Reports (JCR) 2016 provided by the Web of Science. Table 1 provides a detailed description of the variables.

It is worth noting that the counting methods for "usage counts" and "early usage counts" differ. The former refers to the total usage counts an article received from 2016 to 2020, while the latter pertains to the cumulative usage counts an article garnered in the publication year and the following year (i.e., 2016–2017). For instance, consider an article "A" published in 2016, and the usage counts it generated each year from 2016 to 2020 are 5, 4, 3, 2, and 1. The "usage counts" would be 15 (5 + 4 + 3 + 2 + 1), while the "early usage counts" would be 9 (5 + 4). To answer RQ1, in Sect. "Descriptive analysis", we primarily use the former. To answer RQ2 and RQ3, in Sect. "Regression analysis" and "Mediation analysis", we predominantly utilize the latter. The distinction between "Mendeley readership" and "early Mendeley readership" follows a similar pattern.

Table 1 Variable description

Variable	Description
Citations	It is the dependent variable in all models and represents the total number of citations an article has received within the five years since publication
Early usage counts	Cumulative usage counts of an article within the first two years after publication
Early Mendeley readership	Cumulative number of Mendeley readers of an article within the first two years after publication
Funding	Set to 1 if an article is supported for funding; otherwise 0
International collaboration	Set to 1 if an article was developed in an international collaboration; otherwise, mark 0
Open access	Set to 1 if an article is open access (as in Dimensions); otherwise 0
JIF	Journal impact factor of the journal in which an article was published. It was retrieved from the Journal Citation Reports for the year 2016

Mediation effects

Mediation analysis is a powerful research tool in social sciences that facilitates a better comprehension of the fundamental mechanisms through which variables interact with each other. The mediation effect analysis was initially developed in psychology (Baron & Kenny, 1986) and gradually extended to the fields of management and economics (Raguseo et al., 2021; Singh et al., 2023). More recently, this method has also been applied in scientometrics. For instance, Alvarez-Bornstein and Bordons (2021) examined the mediating effects of journal quartile, collaboration type, and the number of references on the relationship between funding and article impact. They found that the presence and magnitude of this effect varied by discipline. Ebrahimy et al. (2016) assessed whether three article-level metrics provided by PLOS—save, discussion, and recommendation metrics—were mediators between the visibility and citations of biomedical articles. Their findings indicated that only the save metric had a positive mediating effect in the relationship between visibility and citations, while recommendation metrics had no impact on this relationship. They also found that discussion metrics played a negative mediating role in this relationship between visibility and citations. Using articles from the biomedical field as the data sample, Vilchez-Roman and Vara-Horna (2021) used usage frequency as a mediating variable to explore the effect of social media platforms such as Twitter and Facebook on the citations. The study revealed that while the direct impact of Twitter on citations was negative, the indirect effect through usage frequency was positive and significant.

Building on prior research (Aduku et al., 2017; Ebrahimy et al., 2016; Thelwall, 2020; Vilchez-Roman & Vara-Horna, 2021), this article extends the data sample to the IEEE Xplore 2016 dataset and performs an examination of the mediating mechanisms between usage counts, Mendeley readership, and citations by controlling for relevant variables. Furthermore, a comparative analysis between journal and conference papers is conducted to expand the study's insights.

Mediation analysis can be conducted using two main methods: traditional regression analysis and structural equation modeling (SEM). As illustrated in Fig. 3a–c, regression analysis involves three steps. First, the coefficient a is estimated by regressing Y on X to assess the total effect. The coefficient b is then estimated by regressing M on X to examine the relationship between the explanatory variable and the mediator variable. Finally, the coefficient c is estimated by regressing Y on M while controlling for X to determine

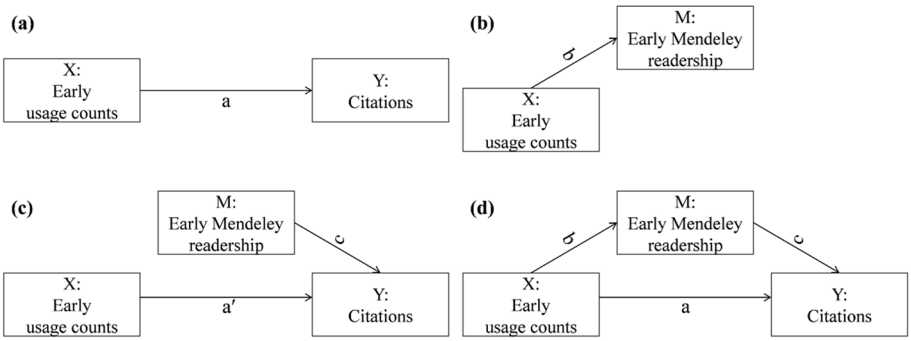

Fig. 3 Diagrammatic illustration of two mediation analysis methods: regression analysis (**a–c**) and structural equation modeling (**d**)

the relationship between the dependent variable and the mediator variable, and the estimated coefficient a' can also be obtained. When all coefficients a, b, and c are significant, a' can be used to determine the presence and extent of mediation effects. Specifically, if a' is significant and equals 0, M has a complete mediation effect, while if a' is not significant and does not equal 0, M has a partial mediation effect. Of course, the existence of partial mediation indicates a decrease in the direct effect of path a. To determine the magnitude of this decrease, further validation is required using methods such as the Sobel test. However, as shown in Fig. 3d, SEM allows for the simultaneous estimation of all model parameters, which can address the large standard errors and inaccurate parameter estimates that can occur when using regression analysis (Iacobucci et al., 2007), making it a superior method for conducting mediation analysis. Therefore, this study employs SEM for mediation analysis and follows the methods proposed by Baron and Kenny (1986) and Zhao et al. (2010) to test mediation effects. The *medsem* program developed by Mehmetoglu in 2018 is used to calculate direct and indirect effects (Mehmetoglu, 2018). Stata 17.0 was used as the analysis tool.

The mediation analysis using SEM follows two steps. Firstly, the model is fitted to estimate the coefficients of the direct effect and the mediation effect of early usage counts on citations. If both $X \rightarrow M$ and $M \rightarrow Y$ are significant, there is a mediation effect, and the research should proceed to step two. If neither coefficient is significant, there is no mediation effect, and the research should be terminated. Secondly, the size of the mediation effect relative to the direct effect is evaluated using the Sobel test, the Delta test, and the Monte Carlo simulation test. If the Z-value based on these tests is significant and the coefficient $X \rightarrow Y$ is not significant, early usage counts have a full mediation effect on citations, indicating that early usage counts cannot directly affect citations but only have an indirect effect through early Mendeley readership. If both the Z-value and coefficient $X \rightarrow Y$ are significant, early usage counts have both a direct effect on citations and an indirect effect through early Mendeley readership. Finally, the final estimation results are organized and reported, with three possible outcomes: no mediation, partial mediation, or full mediation.

Results

Descriptive analysis

As shown in Fig. 4, we present the distribution of usage counts, Mendeley readership, and citations (statistical information about the three metrics can be found in Table 6 in the Appendix). It is worth noting that the usage counts and Mendeley readership mentioned in Sect. "Descriptive analysis" are cumulative totals in the five years since the article's publication. It can be observed that the usage counts of both journal articles and conference papers exhibit a similar trend, with an initial increase followed by a subsequent decline. In contrast, citations follow a power-law distribution. Specifically, the distribution of usage counts for journal articles is mainly concentrated between 200 and 400, with only a small proportion of articles exceeding 1000. For conference papers, the distribution of usage counts is focused between 50 and 150, with a very limited number of articles surpassing 500. Furthermore, the distribution of Mendeley readership exhibits a significant difference between journal and conference papers, with the former following a pattern similar to that of usage counts and the latter following a pattern comparable to that of citations.

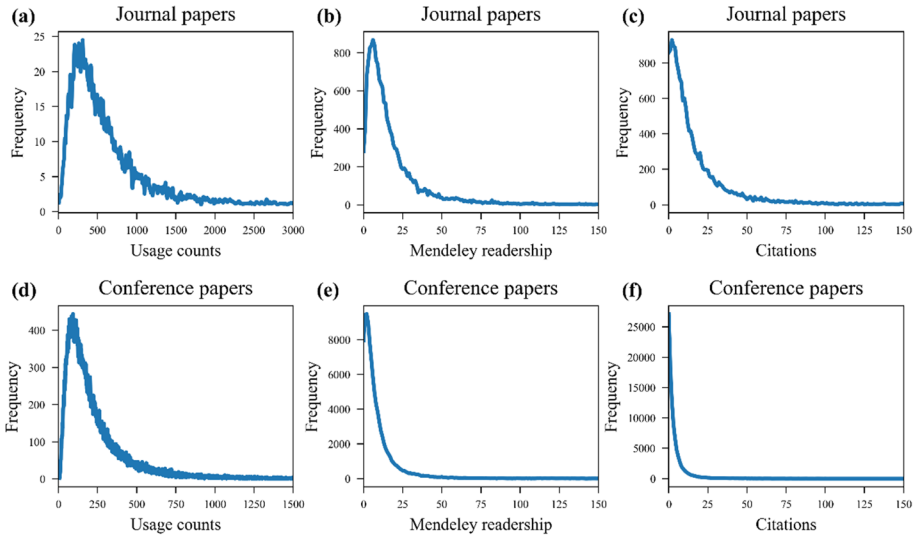

Fig. 4 Distribution of usage counts, Mendeley readership, and citations in journal (subfigures **a–c**) and conference (subfigures **b–d**) papers

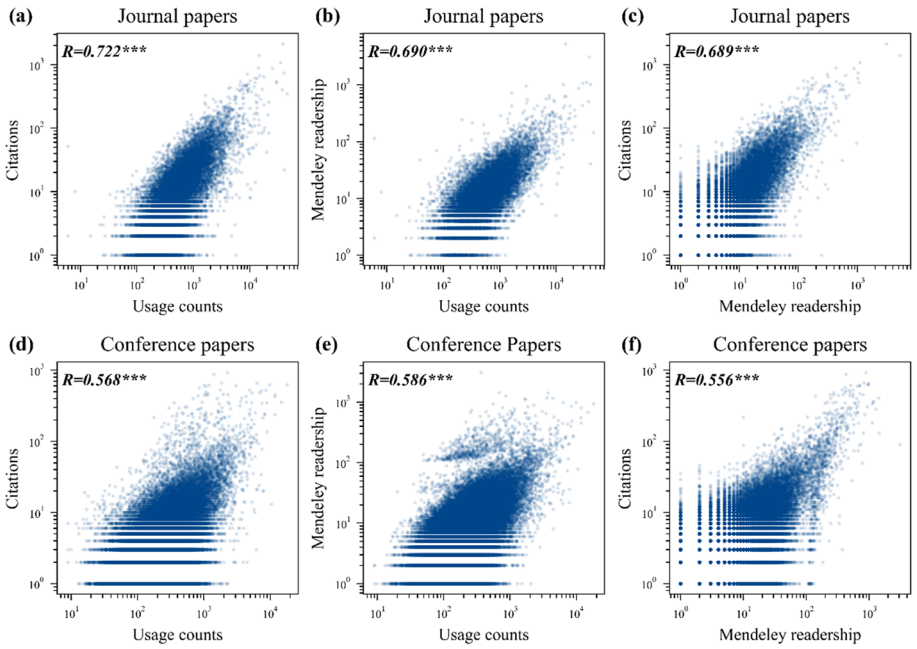

Fig. 5 Pairwise correlation between usage counts, Mendeley readership, and citations for journal (subfigures **a–c**) and conference (subfigures **b–d**) papers. ***Indicates that the correlation coefficient is statistically significant at the 0.001 level, same below

Figure 5 illustrates the pairwise correlations between the metrics of usage counts, Mendeley readership, and citations. Spearman correlation coefficients are reported in the subfigures. It can be observed that these three metrics have significant positive correlations, with the overall correlation coefficients in conference papers being lower than those in journal papers. Specifically, the correlations between the metrics in journal papers are all higher than 0.68, with the strongest correlation observed between usage counts and citations ($r=0.722$). In contrast, the correlation coefficients for conference papers are relatively lower, with the highest correlation coefficient observed between usage counts and Mendeley readership ($r=0.586$). This suggests that both usage counts and Mendeley readership are relevant predictors of citations, regardless of publication type. Additionally, the correlations among these three metrics vary depending on the publication type.

We conducted further investigation into the temporal dynamics of usage counts, Mendeley readership, and citations during the 5-year period after publication. As depicted in Fig. 6a and b, the usage counts of journal articles reached their peak in the first year, subsequently declining rapidly, whereas for conference papers, they peaked in the second year before decreasing. On the other hand, Mendeley readership experienced its highest point in the second year, followed by a downward trend that was more conspicuous for conference papers. Meanwhile, citations tended to stabilize from the third year onward. As demonstrated in Fig. 6c and d, irrespective of the publication type, the dissemination speed of these metrics seems to follow the order of "usage counts > Mendeley

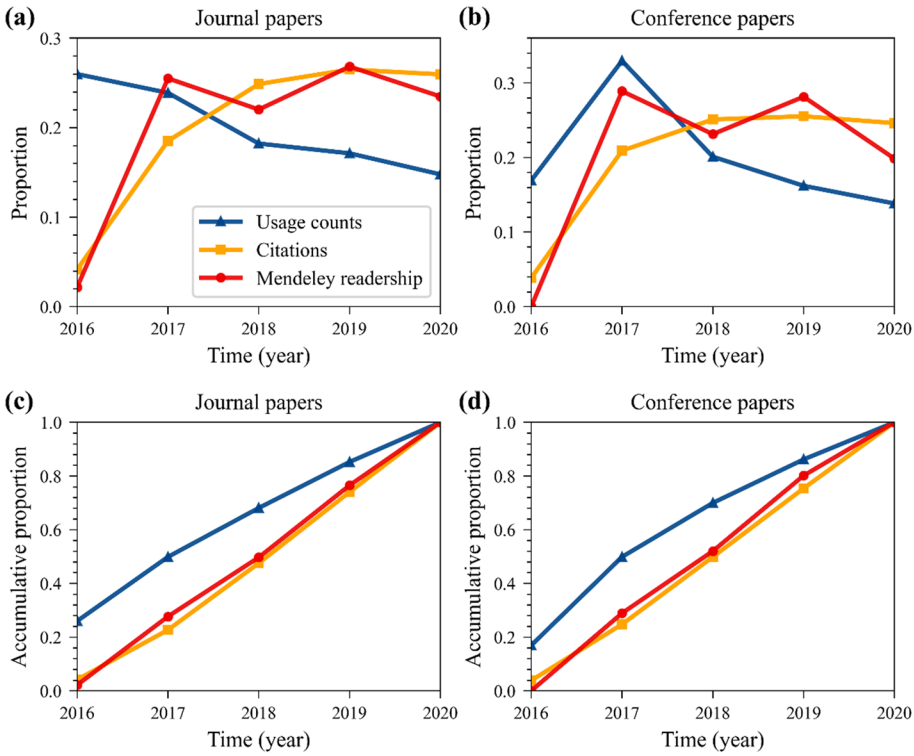

Fig. 6 Changes in the usage counts, Mendeley readership, and citations of journal papers (subfigures **a** and **c**) and conference papers (subfigures **b** and **d**) over time

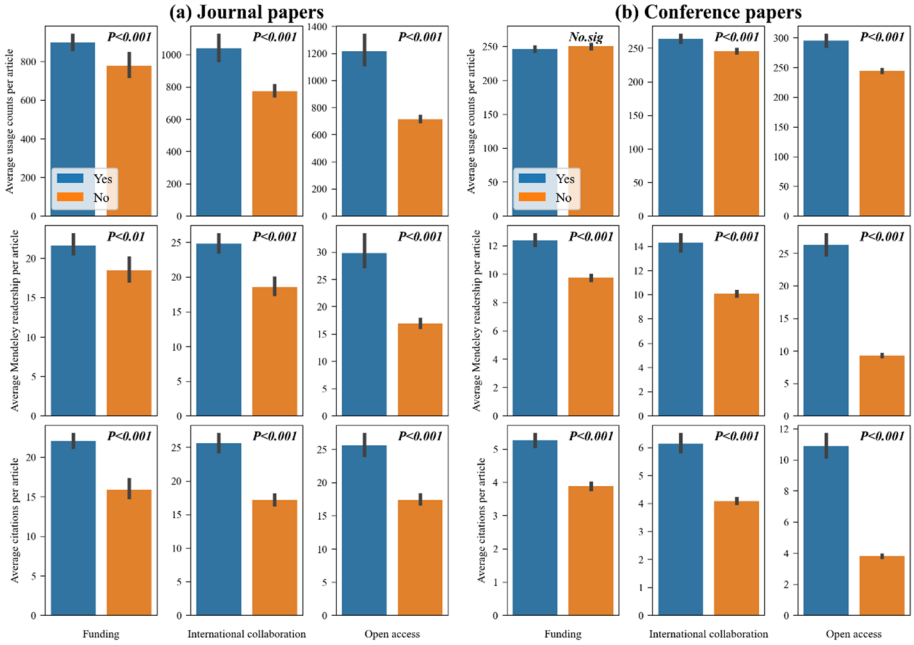

Fig. 7 Relationship between funding/international collaboration/open access and usage counts/Mendeley readership/citations for **a** journal papers and **b** conference papers. "Yes" means that articles in this group were funded by grants or developed international collaborations or were open access. "No" means the opposite

readership > citations". Notably, the dissemination speed of the Mendeley readership is nearly indistinguishable from that of citations.

Finally, we explored the relationship between other bibliometric indicators – the presence of funding, international collaboration, and open access—and usage counts, Mendeley readership, and citations (e.g., do funded articles have higher usage counts, Mendeley readership, or citations compared to unfunded articles?). In this exploration, we initially grouped the journal/conference articles based on whether they received funding (or not), following the same grouping approach for publications with/out international collaboration and with/out open access. Subsequently, we aggregated the means of usage counts/Mendeley readership/citations within each group (as shown in the bars in all subfigures in Fig. 7, where the error bars represent 95% confidence intervals). Finally, employing the *statsmodels* package in Python (https://www.statsmodels.org/stable/index.html), we conducted a one-way analysis of variance (ANOVA) test to examine whether a statistically significant difference exists in the usage counts/Mendeley readership/citations between the two groups (the statistical test result is marked in the upper right corner of all subfigures in Fig. 7).

As shown in Fig. 7, the findings reveal that funding, international collaboration, and open access are positively associated with all three indicators, except for the insignificant relationship of funding with the usage counts of conference papers. Furthermore, the relationship of open access with the three indicators is greater than that of funding and international collaboration, as previously demonstrated in existing research on the advantages of open access in terms of usage counts, Mendeley readership and citations (Holmberg et al.,

2020; Taylor, 2023; Wang et al., 2015). Additionally, while international collaboration and funding have a similar relationship with citations and Mendeley readership, international collaboration outperforms funding in terms of usage counts, particularly in journal articles. Specifically, in journal articles, international collaboration (*vs.* non-international cooperation) could increase the usage counts of an article by 266 (average), but funding (*vs.* non-funding) could only increase it by 120 (average). It is worth mentioning that more detailed information about the statistics of Fig. 7 can be found in Table 7 in the Appendix.

For journal papers, we conducted further investigations into the associations between usage counts, Mendeley readership, citations, and JIF at the journal-level (including 134 journals). All three indicators were significantly and positively correlated with JIF (see Fig. 8) at the journal-level. Specifically, the Spearman correlation coefficients between usage counts and JIF, Mendeley readership and JIF, and citations and JIF were 0.570, 0.584, and 0.775, respectively. Notably, the robust positive correlation between citations and JIF suggested that papers published in high-impact journals tended to garner more citations (Cui et al., 2023; Vaughan et al., 2017). Moreover, the moderate correlation between usage counts/Mendeley readership and JIF implied that articles published in high-impact journals were also more likely to attract more readers on average, although with a weaker relationship than in the case of citations. This suggests that the impact of a journal cannot just be captured by their JIF, but that metrics like usage counts or Mendeley readership can

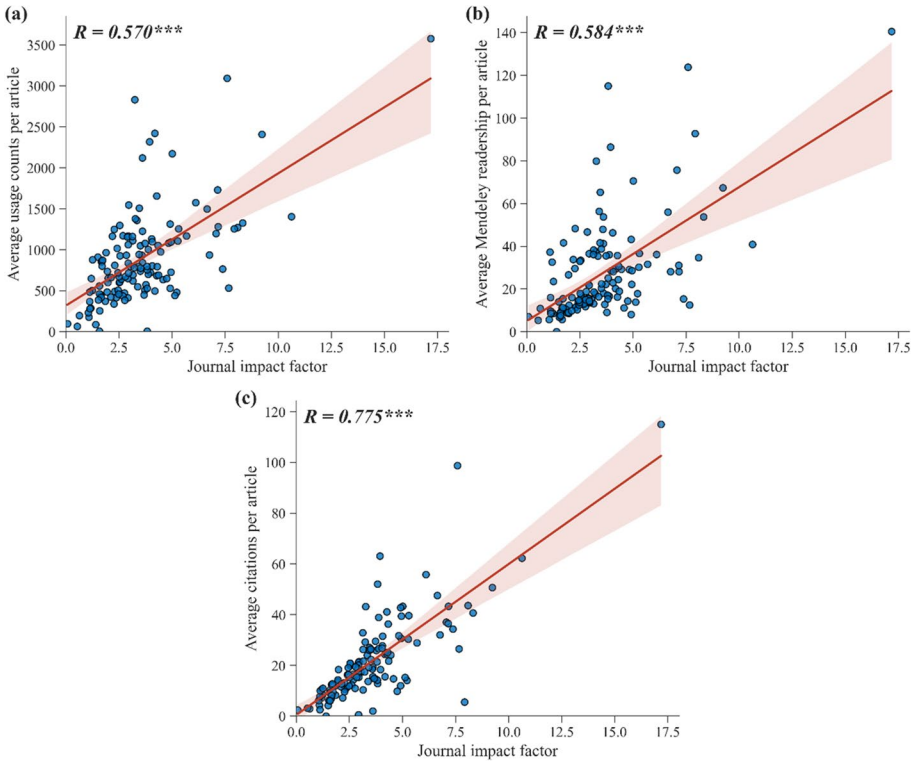

Fig. 8 Spearman correlation between journal impact factor and **a** average usage counts, **b** Mendeley readership, and **c** citations at the journal-level. $N = 134$ unique journals

contribute to providing a more comprehensive and nuanced reflection of an actual journal's impact.

Regression analysis

To investigate the ability of early usage counts and early Mendeley readership to predict future citations of an article. Six linear regression models were constructed for both journal papers (see Table 2) and conference papers (see Table 3). Model 1 and Model 7 were developed using early usage data to predict citations, while Model 2 and Model 8 included additional control variables. Model 3 and Model 9 used early Mendeley data to forecast citations, with Model 4 and Model 10 incorporating control variables. Model 5 and Model 11 integrated early usage data and early Mendeley data to predict citations, while Model 6 and Model 12 added control variables.

The estimation results for journal papers indicate that after controlling for confounding variables, Model 6 has the best fit with an R^2 of 0.478. Thus, we focus on the estimation results of Model 6 in this study. In Model 6, early usage counts have a regression coefficient of 0.233, which is significant at the 0.001 level, indicating that an increase of one unit in early usage counts leads to a 0.233 unit increase in citations while holding all other variables constant. Similarly, early Mendeley readership has a regression coefficient of 0.529, which is significant at the 0.001 level, indicating that an increase of one unit in early Mendeley readership leads to a 0.529 unit increase in citations while all other variables remain constant.

For conference papers, after controlling for confounding variables, Model 12 has the best fit with an R^2 of 0.427. Therefore, we analyze the estimation results of Model 12. In Model 12, the regression coefficient of early usage counts is 0.25, which is significant at the 0.001 level, indicating that an increase of one unit in early usage counts leads to a 0.25 unit increase in citations while all other variables remain constant. Similarly, the estimation result for early Mendeley readership is 0.540, which is significant at the 0.001 level, indicating that an increase of one unit in early Mendeley readership leads to a 0.540 unit increase in citations while all other variables remain constant.

Consistent estimation results for journal and conference papers provide empirical evidence for the predictive ability of early usage counts and early Mendeley readership for citations. Furthermore, the results suggest that differences in early Mendeley readership have a greater relationship with citations than early usage counts in the journal and conference papers. Additionally, further empirical research is necessary to investigate the mechanisms underlying the potential effects of early usage counts, early Mendeley readership, and citations. Thus, we continue to explore the mediating effects in the third part.

Mediation analysis

Assuming a rather stepwise literature retrieval process followed by scientists, which involves first browsing and accessing papers (e.g., by viewing the metadata or downloading the PDF version), then saving them to literature management software (e.g., like Mendeley), and then citing them, this study hypothesized that early Mendeley readership acts as a mediator between usage behavior and citation behavior. To test this assumption, we constructed a mediation model using SEM, the results of which are presented in Table 4. The findings demonstrate that early usage counts have a significant positive effect on citations for both journal and conference papers, as does early Mendeley readership. Moreover, the

Table 2 Regression models for predicting citations by early usage counts or early Mendeley readership (journal papers)

Variables	Predicting citations by early usage counts		Predicting citations by early Mendeley readership		Predicting citations by early usage counts and early Mendeley readership	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Early usage counts	0.499*** (0.007)	0.461*** (0.007)	–	–	0.237*** (0.007)	0.233*** (0.007)
Early Mendeley readership	–	–	0.656*** (0.006)	0.635*** (0.006)	0.541*** (0.007)	0.529*** (0.007)
Funding	–	YES	–	YES	–	YES
International collaboration	–	YES	–	YES	–	YES
Open access	–	YES	–	YES	–	YES
JIF	–	YES	–	YES	–	YES
Adjusted R^2	0.249	0.277	0.430	0.437	0.473	0.478

*Standard errors are given in parentheses

Table 3 Regression models for predicting citations by early usage counts or early Mendeley readership (conference papers)

Variables	Predicting citations by early usage counts		Predicting citations by early Mendeley readership		Predicting citations by early usage counts and early Mendeley readership	
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Early usage counts	0.380*** (0.003)	0.379*** (0.003)	–	–	0.249*** (0.003)	0.250*** (0.003)
Early Mendeley readership	–	–	0.606*** (0.003)	0.601*** (0.003)	0.547*** (0.003)	0.540*** (0.003)
Funding	–	YES	–	YES	–	YES
International collaboration	–	YES	–	YES	–	YES
Open access	–	YES	–	YES	–	YES
Adjusted R^2	0.144	0.162	0.367	0.368	0.426	0.427

*Standard errors are given in parentheses

Table 4 Parameter estimation of SEM

Variables	Journal papers		Conference papers	
	Early Mendeley readership	Citations	Early Mendeley readership	Citations
Early usage counts	0.484*** (0.007)	0.233*** (0.007)	0.240*** (0.004)	0.250*** (0.003)
Early Mendeley readership	–	0.529*** (0.007)	–	0.540*** (0.003)
Funding	YES	YES	YES	YES
International collaboration	YES	YES	YES	YES
Open access	YES	YES	YES	YES
JIF	YES	YES	–	–

*Standard errors are given in parentheses

Table 5 Mediation effect test

	Journal papers			Conference papers		
	Delta	Sobel	Monte Carlo	Delta	Sobel	Monte Carlo
Indirect effect	0.257	0.257	0.257	0.130	0.130	0.130
SE	0.004	0.004	0.004	0.002	0.002	0.002
z-value	60.205	60.950	60.864	66.314	65.548	65.808
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Conf. interval	[0.249, 0.266]	[0.249, 0.266]	[0.249, 0.266]	[0.126, 0.133]	[0.126, 0.133]	[0.126, 0.133]

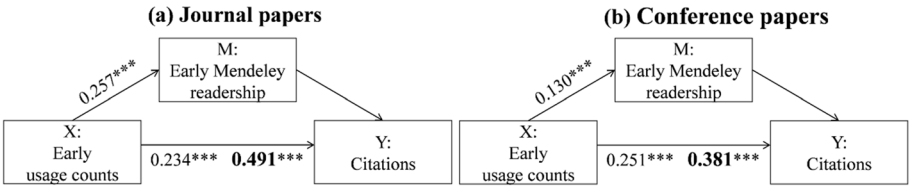

Fig. 9 Path coefficients for the mediation effects models of **a** journal papers and **b** conference papers. The total effects are displayed in a larger font and in bold, and the direct and indirect effects are in smaller font sizes

relationship between early usage counts and early Mendeley readership is also positive and significant. Although the first two relationships were reported in the regression analysis, they are included in Table 4 for the completeness of the SEM analysis. These results suggest that early Mendeley readership mediates the relationship between early usage counts and citations. Furthermore, we observed that the effect of early usage counts on early Mendeley readership is stronger for journal papers (coefficient of 0.484) than for conference papers (coefficient of 0.24).

In order to obtain a more accurate estimate of the mediating effect value of early Mendeley readership and to establish its confidence interval, a comprehensive approach was utilized in this study, which included the Delta, Sobel, and Monte Carlo simulation tests. The results, which are presented in Table 5 and Fig. 9, indicate that early Mendeley readership partially mediates the relationship between early usage counts and citations in both journal and conference papers. Specifically, for journal papers, the mediating effect value of early Mendeley readership on citations through early usage counts is 0.257, with a direct effect value of 0.234, and both are statistically significant at the 0.001 level. Similarly, for conference papers, the mediating effect value of early Mendeley readership on citations through early usage counts is 0.130, with a direct effect value of 0.251, and both are also statistically significant at the 0.001 level.

The finding that early Mendeley readership acts as a bridge between early usage counts and citations is in line with previous research. Some studies have shown that Mendeley readership is a useful indicator for predicting citation counts (Thelwall, 2018; Zahedi et al., 2017), which is also consistent with our regression results in Sect. "Regression analysis". However, not all researchers use Mendeley for literature reading and management, and users typically only record the articles they plan to read or have already read and intend to cite (Mohammadi et al., 2016). There are still many researchers who save and cite articles

through other literature management software, academic paper search platforms, or academic social networking sites. Consequently, early Mendeley readership only plays a partial mediating role.

Furthermore, the mediating effect of early Mendeley readership on journal papers is more significant than that on conference papers, as conference papers rely more on the direct effect of early usage counts on citations. In journal papers, the value of indirect effects/direct effects is 1.1, while in conference papers, the value is only 0.516. This difference may be attributed to the fact that journal papers are typically longer and more complex in their research content, requiring scholars to spend more time reading and engaging with them, thus relying more on Mendeley for literature management and annotation recording. Conference papers are typically shorter in content; therefore, potential citers may not need to manage and engage with them through a reference manager but just view or download them directly from the publishers' website, thus explaining the lower mediating role of the Mendeley readership.

Discussion

Firstly, we did not conduct a detailed analysis of disciplinary heterogeneity. Thelwall (2020) demonstrated a high correlation between Scopus citation counts and Mendeley readership in 11 computer science fields. Nevertheless, whether this conclusion of no disciplinary heterogeneity holds true for conference papers in the IEEE Xplore database requires further exploration. To address this, we performed heterogeneity analysis on the top 10 disciplines with the highest number of conference papers (specific results can be found in Tables 8, 9, and Fig. 10 in the Appendix). The results align with Thelwall's previous findings, indicating that in any given discipline, there is a moderate to strong correlation between usage counts, Mendeley readership, and citations. Additionally, early usage counts and early Mendeley readership serve as early indicators of future scholarly impact across all disciplines. Furthermore, the mediating role of early Mendeley readership between early usage counts and citations exists in all disciplines.

Secondly, previous research has compared the significance of journal articles and conference papers in the field of computer science based on citation counts (Freyne et al., 2010; Thelwall, 2020; Vrettas & Sanderson, 2015). On average, conference papers have lower citation counts than journal papers. As evident from Fig. 7 and Table 6, the findings of this study are consistent with prior research. It is worth noting that the field of computer science places a higher emphasis than any other discipline on conferences as publication venues. Moreover, in this field, the quality of papers published at a select few elite conferences is comparable to that of high-level journal papers (Freyne et al., 2010; Vrettas & Sanderson, 2015).

Usage counts, as a product of the digital age and the flourishing open access movement, have driven the development of "Scientometrics 2.0" (Glänzel & Chi, 2020). Their characteristics of large data volumes and rapid accumulation make them valuable for research evaluation. However, when used as evaluation metrics, their susceptibility to manipulation is challenging to detect, and Mendeley readership shares similar limitations (Thelwall, 2017a). Furthermore, specific biases inherent to Mendeley should be taken into consideration, such as variations in reading patterns across countries (Thelwall & Maflahi, 2015) and its relatively young user base (Mohammadi et al., 2015). Additionally, whether an article possesses an online distribution channel will

also impact its Mendeley readership (Kudlow et al., 2017). Therefore, using usage data and Mendeley data as reference metrics in research evaluation can be considered, but caution should be exercised when directly applying them to scientific assessment.

Conclusion

This study employs a multidimensional analysis, including descriptive, regression, and mediation analyses, of IEEE Xplore-published journal and conference papers from 2016 as data samples to compare the relationships among usage counts, Mendeley readership, and citations. Building upon previous research contributions (as mentioned in the introduction), this study can be regarded as a complement to the research on usage counts. There are two specific research contributions. Firstly, this study utilizes mediation techniques to offer novel insights into the relationships between early usage counts, early Mendeley readership, and citations. Secondly, we provide a valuable exploration of the IEEE Xplore database, which distinguishes this work from studies that rely on commonly used databases such as Web of Science and Scopus.

Results indicate no significant difference in the distribution patterns and accumulation rates between the two types of papers (journal and conference papers). Specifically, usage counts follow a pattern of first increasing and then decreasing, while Mendeley readership and citations follow a power-law distribution. Furthermore, the dissemination speed of the three indicators follows the order of "usage counts > Mendeley readership > citations." Regarding correlation, the correlation coefficient between these three indicators of conference papers (at around 0.56) is lower than that of journal papers (at around 0.70). In addition, funding, international collaboration, and open access have a positive relationship with all three metrics, with the only exception of the relationship between funding and the usage counts of conference papers. Notably, open access has the greatest association with the three metrics, suggesting the important role that open access may have in facilitating the process of accessing, reading, and eventually citing scientific papers. However, prior research suggests that only articles with immediate open access can benefit from these advantages, while delayed open access policies are ineffective in promoting knowledge dissemination in certain emerging and developing countries (Wang et al., 2015; Zhang et al., 2021).

The results of the regression analysis show that early Mendeley readership has better predictive power for citations compared to early usage counts, particularly in journal articles. Secondly, the combination of early usage counts and early Mendeley readership leads to the best predictive performance for citations. Furthermore, the mediation analysis demonstrates that early Mendeley readership partially mediates between early usage counts and citations in the journal and conference papers, with mediation values of 0.257 and 0.130, respectively. Additionally, compared to the direct effect of early usage counts on citations (0.234 and 0.251), conference papers rely more on this direct effect. This suggests that the form of engagement of users with the publications is similar for both journal and conference papers, although with some differences, with the usage counts of conference papers being more strongly associated with citations without the mediation of readership.

Limitations and future work

This study still has certain limitations. Firstly, the sample size of conference papers is approximately six times that of journal papers, which may introduce some degree of bias into our research results. Additionally, in the case of journal papers, about 90% of them are sourced from journals published by IEEE publishers, which may also influence the representation of journal articles within the database to a certain extent.

Furthermore, the funding information for the publications in our dataset was obtained from the Dimensions database. Haunschild and Bornmann (2023) have reported that the Dimensions database provides the most reliable annotation of funding information for institutions within the United States. However, extending this reliability to non-US authors may be constrained by the database. Given this situation, we did not conduct a further review of the reliability of funding information for non-US authors in the Dimensions database, acknowledging it as one of the limitations of this study.

This study presents several avenues for further research. First, additional altmetric indicators, like tweets, news media mentions, or Wikipedia citations, could also be included in a path analysis to better understand the mechanisms underlying their relationship with eventual citation counts and among themselves. Secondly, causality could be further explored to move beyond correlational analyses, which would enable a more accurate interpretation of the actual effects of different usage and readership events on citation counts. Finally, finer-grained time-series data, such as monthly data, could be constructed for early usage counts and early Mendeley readership, and time-series forecasting methods combined with machine learning methods could be used to predict future citations with greater precision.

Appendix

See Tables 6, 7, 8 and 9; Fig. 10.

Table 6 Descriptive statistics for usage counts, Mendeley readership and citations

Publication type	Metrics	N	Min	Max	Mean	Std
Journal	Usage counts	16,799	6	140,544	851.835	2163.630
	Mendeley readership	16,799	0	5256	20.383	62.112
	Citations	16,799	0	2103	19.646	42.312
Conference	Usage counts	98,773	7	18,103	248.755	331.802
	Mendeley readership	98,773	0	3072	10.756	29.124
	Citations	98,773	0	922	4.412	14.540

Table 7 The specific information statistics for Fig. 7

Indicators	Metrics	*N*		Value		F value	*p* value
		(Yes)	(No)	(Yes)	(No)		
Funding (journal)	Usage counts	10,222	6577	898.99	778.55	1241	$p<0.001$
	Mendeley readership	10,222	6577	21.62	18.46	1041	$p<0.01$
	Citations	10,222	6577	22.04	15.92	8424	$p<0.001$
International collaboration (journal)	Usage counts	4899	11,900	1040.06	774.35	5250	$p<0.001$
	Mendeley readership	4899	11,900	24.81	18.56	3522	$p<0.001$
	Citations	4899	11,900	25.67	17.17	14,118	$p<0.001$
Open access (journal)	Usage counts	4535	12,264	1218.90	716.10	18,071	$p<0.001$
	Mendeley readership	45,35	12,264	29.82	16.89	14,453	$p<0.001$
	Citations	45,35	12,264	25.59	17.45	12,355	$p<0.001$
Funding (conference)	Usage counts	37,695	61,078	246.41	250.20	304	*No.sig*
	Mendeley readership	37,695	61,078	12.40	9.74	19,416	$p<0.001$
	Citations	37,695	61,078	5.27	3.88	21,286	$p<0.001$
International collaboration (conference)	Usage counts	15,450	83,323	264.38	245.86	4063	$p<0.001$
	Mendeley readership	15,450	83,323	14.29	10.10	27,095	$p<0.001$
	Citations	15,450	83,323	6.15	4.09	26,216	$p<0.001$
Open access (conference)	Usage counts	8340	90,433	295.53	244.44	18,137	$p<0.001$
	Mendeley readership	8340	90,433	26.31	9.32	266,682	$p<0.001$
	Citations	8340	90,433	10.89	3.82	184,156	$p<0.001$

*"Yes" means that articles in this group were funded by grants or developed international collaborations or were open access. "No" means the opposite

Table 8 The Spearman correlation between usage counts, Mendeley readership, and citations for conference papers in various disciplines

Disciplines	Number of publications	Usage counts & citations	Usage counts & Mendeley readership	Citations & Mendeley readership
Artificial Intelligence and Image Processing	26,479	0.613***	0.635***	0.619***
Communications Technologies	13,689	0.533***	0.555***	0.553***
Electrical and Electronic Engineering	10,509	0.556***	0.571***	0.522***
Information Systems	8929	0.607***	0.695***	0.572***
Materials Engineering	6407	0.543***	0.591***	0.508***
Applied Mathematics	4415	0.503***	0.542***	0.511***
Computer Software	4320	0.553***	0.673***	0.583***
Data Format	2368	0.588***	0.685***	0.572***
Statistics	2366	0.562***	0.594***	0.546***
Numerical and Computational Mathematics	2214	0.565***	0.592***	0.547***

***Indicates that the correlation coefficient is statistically significant at the 0.001 level. Furthermore, when a paper belongs to multiple disciplines (disciplinary data sourced from CWTS in-house database), we utilized the full counting approach, which means that each discipline was counted once. Therefore, the total number of publications in Table 8 may exceed the sample size. Same below

Table 9 Regression models for predicting citations by early usage counts or early Mendeley readership for conference papers in various disciplines

Variables	Artificial intelligence and image processing	Communications technologies	Electrical and electronic engineering	Information system	Materials engineering
Early usage counts	0.933*** (0.022)	0.765*** (0.020)	0.669*** (0.022)	0.603*** (0.033)	0.517*** (0.029)
Early Mendeley readership	2.491*** (0.022)	0.665*** (0.021)	0.484*** (0.022)	1.248*** (0.033)	0.503*** (0.030)
Funding	YES	YES	YES	YES	YES
International collaboration	YES	YES	YES	YES	YES
Open access	YES	YES	YES	YES	YES
Adjusted R^2	0.478	0.307	0.220	0.286	0.205
Variables	Applied mathematics	Computer software	Data format	Statistics	Numerical and computational mathematics
Early usage counts	0.483*** (0.033)	0.741*** (0.049)	0.538*** (0.071)	0.522*** (0.045)	0.907*** (0.049)
Early Mendeley readership	0.627*** (0.033)	1.297*** (0.049)	1.074*** (0.071)	0.735*** (0.046)	0.652*** (0.050)
Funding	YES	YES	YES	YES	YES
International collaboration	YES	YES	YES	YES	YES
Open access	YES	YES	YES	YES	YES
Adjusted R^2	0.231	0.360	0.219	0.248	0.285

Standard errors are given in parentheses

***Indicates that the correlation coefficient is statistically significant at the 0.001 level, same below

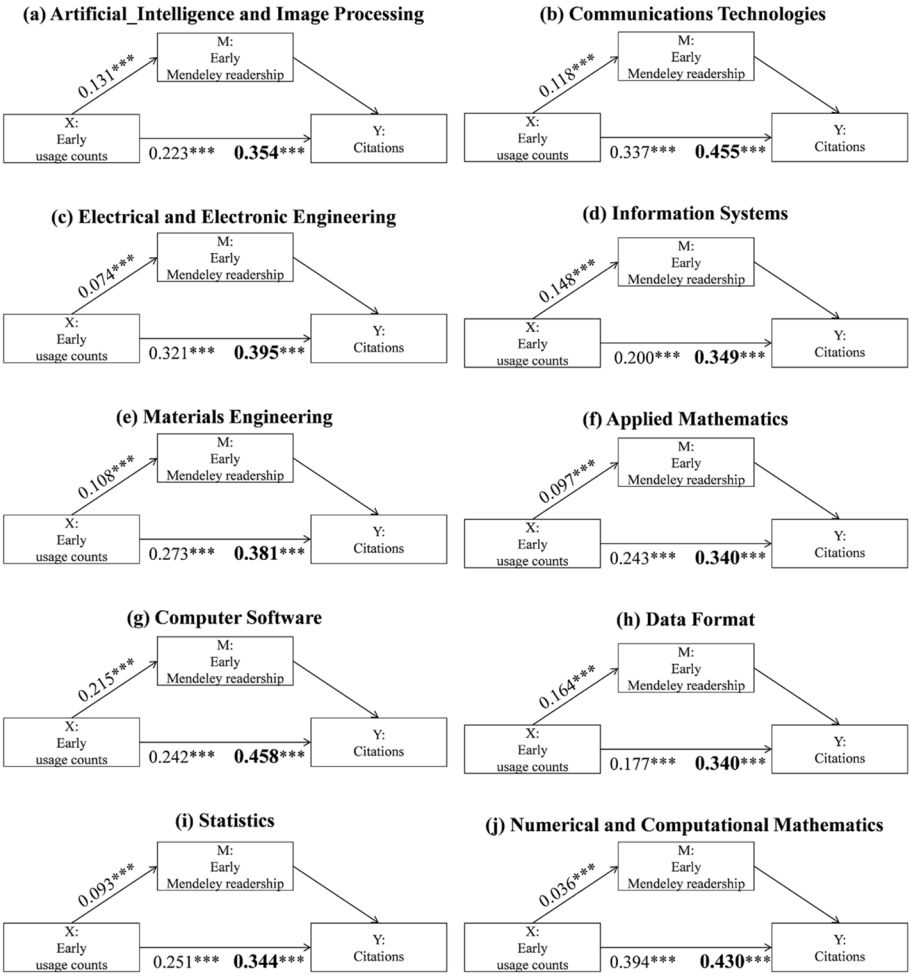

Fig. 10 Path coefficients for the mediation effects models for conference papers in various disciplines. The total effects are displayed in a larger font and in bold, and the direct and indirect effects are in smaller font sizes

Acknowledgements The authors are grateful to the anonymous reviewers for their helpful comments and suggestions.

Funding This study is supported by the National Natural Science Foundation of China (Grant No. 71974029). Rodrigo Costas is partially funded by the South African DSI-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy (SciSTIP). Wencan Tian is financially supported by the China Scholarship Council (Grant No. 202106060134). Zhichao Fang is funded by the Scientific Research Funding of Renmin University of China (Grant No. 23XNF037).

Data availability Data and codes can be accessed from this URL: https://github.com/Tianwencan/IEEE_usage

Declarations

Conflict of interest One of the authors (Rodrigo Costas) is a member of the Distinguished Reviewers Board of the journal *Scientometrics*.

References

Aduku, K. J., Thelwall, M., & Kousha, K. (2017). Do Mendeley reader counts reflect the scholarly impact of conference papers? An investigation of computer science and engineering. *Scientometrics, 112*(1), 573–581. https://doi.org/10.1007/s11192-017-2367-1

Alvarez-Bornstein, B., & Bordons, M. (2021). Is funding related to higher research impact? Exploring its relationship and the mediating role of collaboration in several disciplines. *Journal of Informetrics, 15*(1), 101102. https://doi.org/10.1016/j.joi.2020.101102

Baron, R., & Kenny, D. (1986). The moderator mediator variable distinction in social psychological-research—Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bollen, J., Luce, R., Vemulapalli, S. S., et al. (2003). Usage analysis for the identification of research trends in digital libraries. *D-Lib Magazine, 9*(5), e1395. https://doi.org/10.1045/may2003-bollen

Bollen, J., & van de Sompel, H. (2008). Usage impact factor: The effects of sample characteristics on usage-based impact metrics. *Journal of the American Society for Information Science and Technology, 59*(1), 136–149. https://doi.org/10.1002/asi.20746

Breitzman, A. (2021). The relationship between web usage and citation statistics for electronics and information technology articles. *Scientometrics, 126*(3), 2085–2105. https://doi.org/10.1007/s11192-020-03851-5

Chen, W. M. Y., Bukhari, M., Cockshull, F., et al. (2020). The relationship between citations, downloads and alternative metrics in rheumatology publications: A bibliometric study. *Rheumatology, 59*(2), 277–280. https://doi.org/10.1093/rheumatology/kez163

Chi, P.-S., & Glanzel, W. (2017). An empirical investigation of the associations among usage, scientific collaboration and citation impact. *Scientometrics, 112*(1), 403–412. https://doi.org/10.1007/s11192-017-2356-4

Chi, P.-S., & Glänzel, W. (2018). Comparison of citation and usage indicators in research assessment in scientific disciplines and journals. *Scientometrics, 116*(1), 537–554. https://doi.org/10.1007/s11192-018-2708-8

Chi, P.-S., Gorraiz, J., & Glanzel, W. (2019). Comparing capture, usage and citation indicators: An altmetric analysis of journal papers in chemistry disciplines. *Scientometrics, 120*(3), 1461–1473. https://doi.org/10.1007/s11192-019-03168-y

Costas, R., Perianes-Rodríguez, A., & Ruiz-Castillo, J. (2017). On the quest for currencies of science: Field "exchange rates" for citations and Mendeley readership. *Aslib Journal of Information Management, 69*(5), 557–575. https://doi.org/10.1108/AJIM-01-2017-0023

Costas, R., Zahedi, Z., & Wouters, P. (2015). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology, 66*(10), 2003–2019. https://doi.org/10.1002/asi.23309

Cui, Y., Wang, Y., Liu, X., et al. (2023). Multidimensional scholarly citations: Characterizing and understanding scholars' citation behaviors. *Journal of the Association for Information Science and Technology, 74*(1), 115–127. https://doi.org/10.1002/asi.24709

Ding, Y., Dong, X., Bu, Y., et al. (2021). Revisiting the relationship between downloads and citations: A perspective from papers with different citation patterns in the case of the Lancet. *Scientometrics, 126*(9), 7609–7621. https://doi.org/10.1007/s11192-021-04099-3

Ebrahimy, S., Mehrad, J., Setareh, F., & Hosseinchari, M. (2016). Path analysis of the relationship between visibility and citation: The mediating roles of save, discussion, and recommendation metrics. *Scientometrics, 109*(3), 1497–1510. https://doi.org/10.1007/s11192-016-2130-z

Erdt, M., Nagarajan, A., Sin, S.-C.J., et al. (2016). Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics, 109*(2), 1117–1166. https://doi.org/10.1007/s11192-016-2077-0

Fang, Z., Costas, R., Tian, W., et al. (2020a). An extensive analysis of the presence of altmetric data for Web of Science publications across subject fields and research topics. *Scientometrics, 124*(3), 2519–2549. https://doi.org/10.1007/s11192-020-03564-9

Fang, Z., Costas, R., Tian, W., et al. (2021). How is science clicked on Twitter? Click metrics for Bitly short links to scientific publications. *Journal of the Association for Information Science and Technology, 72*(7), 918–932. https://doi.org/10.1002/asi.24458

Fang, Z., Costas, R., & Wouters, P. (2022). User engagement with scholarly tweets of scientific papers: A large-scale and cross-disciplinary analysis. *Scientometrics, 127*(8), 4523–4546. https://doi.org/10.1007/s11192-022-04468-6

Fang, Z., Dudek, J., & Costas, R. (2020b). The stability of Twitter metrics: A study on unavailable Twitter mentions of scientific publications. *Journal of the Association for Information Science and Technology, 71*(12), 1455–1469. https://doi.org/10.1002/asi.24344

Freyne, J., Coyle, L., Smyth, B., & Cunningham, P. (2010). Relative status of journal and conference publications in computer science. *Communications of the ACM, 53*(11), 124–132. https://doi.org/10.1145/1839676.1839701

Geng, Y., Cao, R., Han, X., et al. (2022). Scientists are working overtime: When do scientists download scientific papers? *Scientometrics, 127*(11), 6413–6429. https://doi.org/10.1007/s11192-022-04524-1

Glaenzel, W., & Gorraiz, J. (2015). Usage metrics versus altmetrics: Confusing terminology? *Scientometrics, 102*(3), 2161–2164. https://doi.org/10.1007/s11192-014-1472-7

Glänzel, W., & Chi, P.-S. (2020). The big challenge of Scientometrics 2.0: Exploring the broader impact of scientific research in public health. *Scientometrics, 125*(2), 1011–1031. https://doi.org/10.1007/s11192-020-03473-x

Guerrero-Bote, V. P., & Moya-Anegón, F. (2014). Relationship between downloads and citations at journal and paper levels, and the influence of language. *Scientometrics, 101*(2), 1043–1065. https://doi.org/10.1007/s11192-014-1243-5

Haunschild, R., & Bornmann, L. (2023). Identification of potential young talented individuals in the natural and life sciences: A bibliometric approach. *Journal of Informetrics, 17*(3), 101394. https://doi.org/10.1016/j.joi.2023.101394

Haustein, S., Bowman, T. D., & Costas, R. (2016). Interpreting 'altmetrics': Viewing acts on social media through the lens of citation and social theories. In S. R. Sugimoto (Ed.), *Theories of informetrics and scholarly communication* (pp. 372–406). De Gruyter Saur.

Haustein, S., Peters, I., Sugimoto, C. R., et al. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology, 65*(4), 656–669. https://doi.org/10.1002/asi.23101

Hicks, D., Wouters, P., Waltman, L., et al. (2015). Bibliometrics: The Leiden manifesto for research metrics. *Nature, 520*(7548), 429–431. https://doi.org/10.1038/520429a

Holmberg, K., Hedman, J., Bowman, T. D., et al. (2020). Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. *Scientometrics, 122*(1), 645–659. https://doi.org/10.1007/s11192-019-03301-x

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology, 17*(2), 139–153. https://doi.org/10.1016/S1057-7408(07)70020-7

Khan, M. S., & Younas, M. (2017). Analyzing readers behavior in downloading articles from IEEE digital library: A study of two selected journals in the field of education. *Scientometrics, 110*(3), 1523–1537. https://doi.org/10.1007/s11192-016-2232-7

Kudlow, P., Cockerill, M., Toccalino, D., et al. (2017). Online distribution channel increases article usage on Mendeley: A randomized controlled trial. *Scientometrics, 112*(3), 1537–1556. https://doi.org/10.1007/s11192-017-2438-3

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics, 91*(2), 461–471. https://doi.org/10.1007/s11192-011-0580-x

Lippi, G., & Favaloro, E. J. (2013). Article downloads and citations: Is there any relationship? *Clinica Chimica Acta, 415*, 195–195. https://doi.org/10.1016/j.cca.2012.10.037

McGillivray, B., & Astell, M. (2019). The relationship between usage and citations in an open access mega-journal. *Scientometrics, 121*(2), 817–838. https://doi.org/10.1007/s11192-019-03228-3

Mehmetoglu, M. (2018). Medsem: A Stata package for statistical mediation analysis. *International Journal of Computational Economics and Econometrics, 8*(1), 63–78. https://doi.org/10.1504/IJCEE.2018.088321

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. https://doi.org/10.1002/asi.23071

Mohammadi, E., Thelwall, M., Haustein, S., et al. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. *Journal of the Association for Information Science and Technology, 66*(9), 1832–1846. https://doi.org/10.1002/asi.23286

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. https://doi.org/10.1002/asi.23477

Raguseo, E., Pigni, F., & Vitari, C. (2021). Streams of digital data and competitive advantage: The mediation effects of process efficiency and product effectiveness. *Information & Management, 58*(4), 103451. https://doi.org/10.1016/j.im.2021.103451

Rowlands, I., & Nicholas, D. (2007). The missing link: Journal usage metrics. *Aslib Proceedings, 59*(3), 222–228. https://doi.org/10.1108/00012530710752025

Schloegl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: The case of oncology journals. *Scientometrics, 82*(3), 567–580. https://doi.org/10.1007/s11192-010-0172-1

Schloegl, C., Gorraiz, J., Gumpenberger, C., et al. (2014). Comparison of downloads, citations and readership data for two information systems journals. *Scientometrics, 101*(2), 1113–1128. https://doi.org/10.1007/s11192-014-1365-9

Singh, R., Charan, P., & Chattopadhyay, M. (2023). Relational capabilities and performance: Examining the moderation-mediation effect of organisation structures and dynamic capability. *Knowledge Management Research & Practice, 21*(1), 92–106. https://doi.org/10.1080/14778238.2020.1843984

Sugimoto, C. R., Work, S., Larivière, V., et al. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology, 68*(9), 2037–2062. https://doi.org/10.1002/asi.23833

Taylor, M. (2023). Slow, slow, quick, quick, slow: Five altmetric sources observed over a decade show evolving trends, by research age, attention source maturity and open access status. *Scientometrics, 128*(4), 2175–2200. https://doi.org/10.1007/s11192-023-04653-1

Thelwall, M. (2017a). Are Mendeley reader counts high enough for research evaluations when articles are published? *Aslib Journal of Information Management, 69*(2), 174–183. https://doi.org/10.1108/AJIM-01-2017-0028

Thelwall, M. (2017b). Are Mendeley reader counts useful impact indicators in all fields? *Scientometrics, 113*(3), 1721–1731. https://doi.org/10.1007/s11192-017-2557-x

Thelwall, M. (2018). Early Mendeley readers correlate with later citation counts. *Scientometrics, 115*(3), 1231–1240. https://doi.org/10.1007/s11192-018-2715-9

Thelwall, M. (2020). Mendeley reader counts for US computer science conference papers and journal articles. *Quantitative Science Studies, 1*(1), 347–359. https://doi.org/10.1162/qss_a_00010

Thelwall, M., & Kousha, K. (2017). ResearchGate articles: Age, discipline, audience size, and impact. *Journal of the Association for Information Science and Technology, 68*(2), 468–479. https://doi.org/10.1002/asi.23675

Thelwall, M., & Maflahi, N. (2015). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. *Journal of the Association for Information Science and Technology, 66*(6), 1124–1135. https://doi.org/10.1002/asi.23252

Tian, W., Hu, Z., & Wang, X. (2019). Upgrading from 3G to 5G: Topic evolution and persistence among scientists. In: *Proceedings of the 17th international conference on scientometrics and informetrics* (pp. 1156–1165).

Vaughan, L., Tang, J., & Yang, R. (2017). Investigating disciplinary differences in the relationships between citations and downloads. *Scientometrics, 111*(3), 1533–1545. https://doi.org/10.1007/s11192-017-2308-z

Vilchez-Roman, C., & Vara-Horna, A. (2021). Usage, content and citation in open access publication: Any interaction effects? *Scientometrics, 126*(12), 9457–9476. https://doi.org/10.1007/s11192-021-04178-5

Vrettas, G., & Sanderson, M. (2015). Conferences versus journals in computer science. *Journal of the Association for Information Science and Technology, 66*(12), 2674–2684. https://doi.org/10.1002/asi.23349

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007

Wan, J., Hua, P., Rousseau, R., et al. (2010). The journal download immediacy index (DII): Experiences using a Chinese full-text database. *Scientometrics, 82*(3), 555–566. https://doi.org/10.1007/s11192-010-0171-2

Wang, X., Fang, Z., & Sun, X. (2016). Usage patterns of scholarly articles on Web of Science: A study on Web of Science usage count. *Scientometrics, 109*(2), 917–926. https://doi.org/10.1007/s11192-016-2093-0

Wang, X., Liu, C., Mao, W., et al. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics, 103*(3), 1149–1149. https://doi.org/10.1007/s11192-015-1589-3

Wang, X., Mao, W., Xu, S., et al. (2014). Usage history of scientific literature: Nature metrics and metrics of Nature publications. *Scientometrics, 98*(3), 1923–1933. https://doi.org/10.1007/s11192-013-1167-5

Wang, Z., Chen, Y., & Glanzel, W. (2020). Preprints as accelerator of scholarly communication: An empirical analysis in Mathematics. *Journal of Informetrics, 14*(4), 101097. https://doi.org/10.1016/j.joi.2020.101097

Zahedi, Z., & Costas, R. (2018). General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PLoS ONE, 13*(5), e0197326. https://doi.org/10.1371/journal.pone.0197326

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. https://doi.org/10.1007/s11192-014-1264-0

Zahedi, Z., Costas, R., & Wouters, P. (2017). Mendeley readership as a filtering tool to identify highly cited publications. *Journal of the Association for Information Science and Technology, 68*(10), 2511–2521. https://doi.org/10.1002/asi.23883

Zahedi, Z., & Haustein, S. (2018). On the relationships between bibliographic characteristics of scientific documents and citation and Mendeley readership counts: A large-scale analysis of Web of Science publications. *Journal of Informetrics, 12*(1), 191–202. https://doi.org/10.1016/j.joi.2017.12.005

Zhang, G., Wang, Y., Xie, W., et al. (2021). The open access usage advantage: A temporal and spatial analysis. *Scientometrics, 126*(7), 6187–6199. https://doi.org/10.1007/s11192-020-03836-4

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206. https://doi.org/10.1086/651257

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.