Library and Information Science Research 45 (2023) 101239

Contents lists available at ScienceDirect

Library and Information Science Research

journal homepage: www.elsevier.com/locate/lisres

Motivation for downloading academic publications

Zequan Xiong [a], Xia Peng [a], Li Yang [a], Wen Lou [b,c,d,*], Star X. Zhao [e,f]

[a] Library, East China Normal University, 500 Dongchuan Road, Shanghai, China
[b] Faculty of Economics and Management, East China Normal University, 500 Dongchuan Road, Shanghai, China
[c] Key Laboratory of Knowledge Mining and Knowledge Services in Agricultural Converging Publishing, National Press and Publication Administration, Agricultural Information Institute, Chinese Academy of Agricultural Sciences, China
[d] Key Laboratory of Advanced Theory and Application in Statistics and Data Science (East China Normal University), Ministry of Education of China, China
[e] Institute of Big Data (IBD), Fudan University, 220 Handan Road, Shanghai, China
[f] State Base of Intelligent Evaluation & Governance Experiment, Fudan University, 220 Handan Road, Shanghai, China

ARTICLE INFO

Keywords:
Motivation
Download
Usage
Technology acceptance model

ABSTRACT

Downloads have been considered as supplemental to citations in reflecting the impact of scientific research activities and scientific output, yet the motivations to download a specific publication has not been fully explored. In scientific evaluation practice, unclear motivations could lead to difficulties for evaluating the impact of academic literature without providing a cogent interpretation of downloads as an alternative metric. To fill this gap, an expanded Technology Acceptance Model (TAM) to investigate motivations for downloading academic literature was proposed and the effectiveness verified using questionnaire data containing 480 respondents. The results show that the degree of usefulness of literature to users and the degree of relevance of literature to users were the primary factors that drive users to download specific literature. Due to the consistency between downloading and citing in reflecting the usefulness, downloads is an effective metric to supplement citation metrics in evaluating the impact of academic literature.

1. Introduction

The emergence of the big data era has led to the development of various new disciplines. In the fields of scientometrics and science of science, the monopolistic status of traditional indicators based on citations has been challenged gradually by a multidimensional evaluation system. Altmetrics or usage metric indicators draw the most interest in this new evaluation system. These indicators focus primarily on the big data generated by academic publication usage, including measures of downloading, sharing, bookmarking, and discussing on social media. These types of data have been considered supplements to citation data in reflecting the impact of scientific research activities and scientific output (Kousha, Thelwall, & Abdoli, 2017; Torres-Salinas, Robinson-Garcia, & Gorraiz, 2017; Xiong & Duan, 2019), including the effect on a wider range of nonacademic audiences (Aung et al., 2019).

Previous research on altmetrics or usage metrics has focused primarily on external attributes, such as the characteristics of indicators (Duan & Xiong, 2017; Moed & Halevi, 2016; Wang, Fang, & Sun, 2016), or on correlation among different indicators (Lippi & Favaloro, 2013; Moed & Halevi, 2016; Schloegl, Gorraiz, Gumpenberger, Jack, & Kraker, 2014), especially the correlation between citation and altmetrics. Weak or moderate correlations between altmetrics and citation metrics have been found from these studies using different sets of publications (Costas, Zahedi, & Wouters, 2015; Thelwall, Haustein, Larivière, Sugimoto, & Bornmann, 2013; Xiong & Duan, 2019). In the last few years, some scholars paid attention to the intrinsic motivations of publication usage, such as motivations for bookmarking in Mendeley (Mohammadi, Thelwall, & Kousha, 2016), motivations for sharing articles on Twitter (Htoo & Na, 2017), and motivations for self-archiving on ResearchGate (Lee, Oh, Dong, Wang, & Burnett, 2019). However, researchers have limited understanding of the motivations for downloading, the most important step in the literature-use process.

2. Problem statement

Although scholars have conducted a series of studies on usage metrics, existent studies have not conducted an in-depth analysis on their intrinsic motivation. While usage data are assumed to almost always be generated by people who use the literature, there is no evidence that people selected a particular document for its impact or usefulness. Users

* Corresponding author at: Faculty of Economics and Management, East China Normal University, 500 Dongchuan Road, Shanghai, China.
E-mail address: wlou@infor.ecnu.edu.cn (W. Lou).

https://doi.org/10.1016/j.lisr.2023.101239
Received 8 July 2022; Received in revised form 9 January 2023; Accepted 13 March 2023
Available online 28 March 2023

may have been attracted by a newfangled title (Jamali & Nikzad, 2011) or a relevant author name (e.g., themselves or an acquaintance), or may have generated usage data for other random reasons (Buter & van Raan, 2011).

Consequently, there are disputes about the reliability of usage metric indicators in reflecting the impact of publications (Haustein, 2016). There are also insufficient examples of rational use of these indicators in the evaluation process. To resolve these questions of reliability and robust methodology, it is imperative to explore the motivations for using publications. This type of exploration will further explain how to measure the actual impact and usefulness of publications using usage metric indicators.

This study conducted empirical research on user motivations for accessing academic publications. Downloading is the most common and important step in the process of utilizing electronic publications; downloads have been proven to be one of the most relevant indicators in terms of citations (Naude, 2017; Xiong & Duan, 2019). Some researchers have asserted that the number of article downloads is useful in author or article evaluation (Haustein et al., 2014; Kurtz & Bollen, 2010). However, the factors that drive users to download specific publications and to what extent the number of downloads can reflect the impact of the publications remain unknown. This study selected downloading as the representative behavior to ascertain user motivation to utilize academic literature accessed electronically. The objective of this study is to investigate which motivations drive users to download publications, and the relationship between these motivations.

3. Literature review

3.1. Literature usage and downloads

Usage behaviors related to a specific publication before it is cited are browsing, viewing, downloading, and reading (Kurtz & Bollen, 2010). The publication is the carrier of knowledge and information. Users first browse through the publication; some are attracted by the titles or abstracts of specific documents and download and read them to obtain specific knowledge or information. A smaller percentage of readers will cite these specific articles in their own manuscripts or share them with others. After peer review and publication of the user's own manuscript, the knowledge and information then enter a new usage cycle (Brody, Harnad, & Carr, 2006).

In the process of this cycle, browsing is accompanied by much casual interest, reading is difficult to capture and assess, while downloading is more targeted and easier to record by servers (Wan, Hua, Rousseau, & Sun, 2010). Thus, the majority of scientometricians consider the number of downloads as useful information for authors or for output evaluation (Haustein et al., 2014; Kurtz & Bollen, 2010). Some researchers have discussed usage in terms of downloads (Brody et al., 2006; Schloegl & Gorraiz, 2010; Schloegl & Gorraiz, 2011); others treat downloads as the primary component of usage in discussing the rationality behind downloads (Bollen, de Sompel, Smith, & Luce, 2005), influencing factors on downloads (Guerrero-Bote & Moya-Anegon, 2014; Subotic & Mukherjee, 2014), and the correlation of downloads and citations (Moed, 2005; Schloegl et al., 2014; Schloegl & Gorraiz, 2010; Schloegl & Gorraiz, 2011). These studies have provided a wealth of information about the characteristics of downloads and the relationship between downloads and citations. Yet, there is limited knowledge regarding why users download specific publications, particularly the motivations for downloading.

3.2. Motivations for academic publications use

Motivation is an interdisciplinary research topic. The motivation for using computers (Davis, 1989), the Internet (Teo, Lim, & Lai, 1999), cellphone texting (Park, Lee, & Chung, 2016), blogs (Hsu & Lin, 2008; Zhang & Pentina, 2012), and YouTube (Klobas, McGill, Moghavvemi, & Paramanathan, 2018) have been studied extensively. However, very few studies have focused on the motivation for academic publications usage, and most of them have focused on sharing or mentioning that behavior on social media platforms (Jin-Cheon, 2015). For example, Kumar, Gupta, Baskaran, and Jin-Cheon (2019) built a machine learning model to classify tweet motivations into six categories: expressing opinion, interaction, promotion, sharing, and summarization. They found that tweets about medicine and environmental science directly impacted the public while tweets about chemistry showed a higher percentage of self-citation and promotion. Na and Ye (2017) conducted a content analysis of 2783 posts on Facebook and found that nearly half of the posts were simple sharing of articles without content, while 20.4% of the posts were sharing with discussion or evaluation of the articles. Only a few posts were motivated by self-citation in their study. Lee et al. (2019) investigated motivations for self-archiving on ResearchGate and detected that accessibility was the most important factor. Users archived their publications on ResearchGate because they believed it would make their research outputs more widely and more easily used.

Some scholars have investigated the motivations for bookmarking or saving documents. For example, Mohammadi et al. (2016) investigated the motivations for bookmarking publications in Mendeley and found that most respondents bookmarked publications to cite them later. Others bookmarked publications for use in professional, teaching, and other educational activities. In that way, Mendeley's readership data can be used to capture knowledge circulation in scientific activities, especially for users who read but do not author publications (Mohammadi & Thelwall, 2014).

From these existent studies, sharing motivation appears to encompass more non-academic factors, while bookmarking and saving motivation may contain more academic factors. However, researchers have limited understanding of the motivations for downloading, the most important step in the literature-use process. This study explored that key issue.

4. Research model and hypotheses development

The purpose of this study was to find out what causes users to download specific literature. Descriptive statistical analysis may only assess the correlation between the two variables and lacks the ability of causal inference. Therefore, for the analysis of the questionnaire, in addition to descriptive statistics, this research introduced structural equation modeling (SEM) to make a causal analysis of users' download motivation and download frequency. In addition, the use of SEM can make research problems more concise. For example, the use of a latent variable to reflect various similar observation variables (e.g. citations, downloads and impact factor) can enable readers to gain a clearer and deeper understanding of the phenomenon.

Downloading is the process of receiving data over the Internet (Christensson, 2014). Accordingly, this study defined downloading publications as the process of acquiring publications over the Internet. Models that focused on Internet use motivation were under preliminarily consideration as the study's potential research models. Among these potential models, the most classic and influential model is the Technology Acceptance Model (TAM) (Schöpfel & Azeroual, 2021).

TAM was developed by Davis (1989) when he investigated motivations to use computers. Now it has become one of the most influential models in studying technology acceptance and usage motivation. In this model, a user's intention to use technology is regulated by two primary factors: perceived usefulness and perceived ease of use. Perceived usefulness refers to the degree to which a person believes that his or her task performance can be improved by using the technology. Perceived ease of use refers to the extent to which a person feels free of effort to use the technology. Although some scholars consider that TAM lacks subjective norms and have extended TAM with additional factors (i.e., social influence, age, and gender), TAM still serves as a useful general framework and has demonstrated a high level of explanatory power in various

contexts.

TAM was applied, as downloading is essentially an Internet using behavior (utilizing the Internet to find publications); downloading can thus be explained by the Internet-usage motivation model. The original TAM was extended to be more suitable for evaluating publication-usage motivation. The study's proposed model was based on four factors: perceived usefulness, perceived ease of use, perceived enjoyment, and perceived relevance. Therefore, the research questions focus on the above four factors. Specifically, the following three research questions will be answered in this study:

(1) To what extent does perceived usefulness, perceived ease of use, perceived enjoyment, perceived relevance affect users' download frequency respectively?
(2) What is the relationship between perceived usefulness, perceived ease of use, perceived enjoyment, and perceived relevance?
(3) To what extent can downloads reflect the impact of publications?

The first two of the above three questions involve the inference of causality, so some hypotheses need to be put forward and verified in advance.

4.1. Perceived usefulness

Perceived usefulness is defined as the prospective user's subjective belief that downloading a specific publication will improve his or her task performance. For example, whether the publication helps the user acquire new knowledge, an authoritative definition, or a research hotspot? Previous studies on TAM have demonstrated that perceived usefulness has a strong and consistent relationship with usage (Adams, Nelson, & Todd, 1992; Igbaria, Iivari, & Maragahh, 1995; Teo et al., 1999). In the context of literature use, this study postulated that perceived usefulness is positively related to literature usage (i.e., downloading in this study). The following hypothesis was proposed:

H1. Perceived usefulness positively affects users' download frequency.

4.2. Perceived enjoyment

People may engage in a particular activity if it yields fun or enjoyment (Teo et al., 1999). Previous research has demonstrated that affection (i.e., feelings of joy or pleasure) may affect individual behavior (Teo et al., 1999). In the context of literature use, a person is likely to download a specific document because he or she is attracted by an interesting abstract and desires more details. In this context, people perceive that they will experience fun and enjoyment from reading interesting content; this motivation is defined as perceived enjoyment. Teo et al. (1999) confirmed that perceived enjoyment had a significant effect on Internet usage. This study speculated that there is also a positive correlation between perceived enjoyment and literature download frequency.

Meanwhile, a user attracted by an interesting title may view the abstract first, and then download the full text if he or she perceives the document as useful based on the content of the abstract. This study expected that perceived enjoyment could influence download frequency indirectly via perceived usefulness, and proposed the following hypotheses:

H2. Perceived enjoyment positively affects users' download frequency.

H3. Perceived enjoyment positively affects users' perceived usefulness.

4.3. Perceived ease of use

Perceived ease of use was defined as the degree to which a person believes that downloading a specific publication is free of effort. Users may be able to download the full text needed with one click, or the download process may take a very short time. Documents that are more easily viewed would likely be downloaded more frequently because users would not need to spend extra time in retrieval. However, if the downloading process is onerous, users may give up on the process.

Perceived ease of use has been demonstrated to influence usage indirectly via perceived usefulness (Davis, 1989) and perceived enjoyment (Igbaria et al., 1995). In the context of literature use, if the document is difficult to download, it is less likely to be perceived as useful as its main content cannot be detected by users. A similar inference can be made for the effects of perceived ease of use on perceived enjoyment; a document that is difficult to download is less likely to be perceived as enjoyable. Based on the above analysis, Hypotheses 4–6 were proposed:

H4. Perceived ease of use positively affects users' download frequency.

H5. Perceived ease of use positively affects users' perceived usefulness.

H6. Perceived ease of use positively affects users' perceived enjoyment.

4.4. Perceived relevance

In previous studies of citation motivations, an interpersonal connection to cited authors was shown to be an important factor (Case & Higgins, 2000). For example, a user may cite a paper because the paper was authored, shared, or recommended by someone familiar. This study aimed to determine if this positive effect also existed in literature download behavior. Perceived relevance was introduced as the degree to which a person believes the publication is relevant to himself or herself. It should be noted that the original concept of perceived relevance refers to "the belief that information is related to one's need or perceived need for it" (Poole, 1985). While in this study, in order to investigate the impact of interpersonal connection rather than the content of literature on downloading, the perceived relevance refers to the perceived identity relevance. Generally, a person will download publications that he or she feels are relevant to him or her. If the publication is authored, shared, or recommended by someone the user is familiar with (i.e., a tutor, colleague, classmate, expert), the user would consider the publication more useful and valuable. Thus, the following hypotheses were proposed:

H7. Perceived relevance positively affects users' download frequency.

H8. Perceived relevance positively affects users' perceived usefulness.

Based on these hypotheses and TAM, a research model for the study was constructed, as shown in Fig. 1. The model was comprised of four latent variables and highlighted the eight hypotheses proposed to meet the objective of this study.

5. Methodology

5.1. Semi-structured interview

Before conducting a formal questionnaire survey, a semi-structured interview was conducted with five faculty members and five students of different majors in the authors' institution to identify the primary factors related to download motivation. First, five major disciplines (Humanities and Social Sciences, Natural Science, Engineering, Agricultural Science, and Medical Science) were determined, and then one student and one faculty were selected respectively from the above five disciplines as the potential interviewees. Student interviewees were selected from the graduate students who had taken library courses, while faculty interviewees were selected from the regular users of

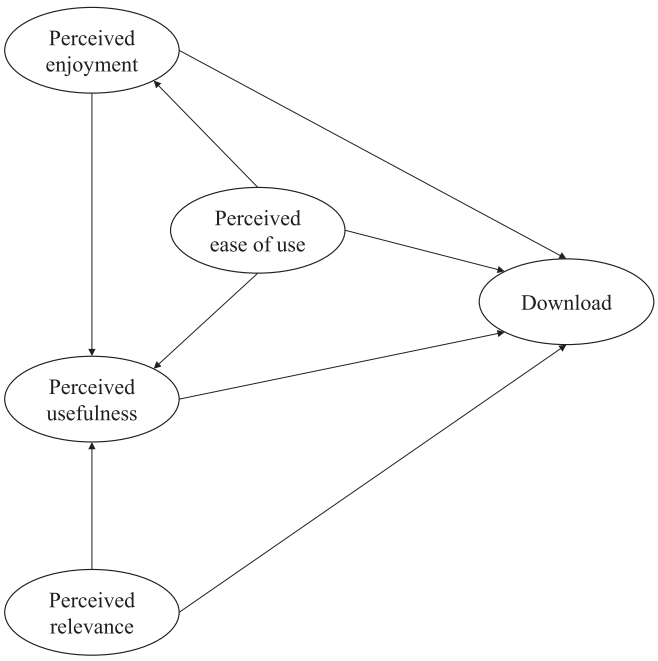

Fig. 1. Research model.

library services, to ensure all potential interviewees were users who often download literature and could be contacted. If the potential interviewees refuse, re-selection was conducted until the selection of five faculty members and five students was completed. The user information of the finalized semi-structured interview is shown in Table 1.

The interview was conducted face to face and lasted for 5–10 minutes. Interview questions focused on the query: "Why do you download the full text of specific publications?" The responses of some interviewees were as follows:

> *Prof. Y: I am curious. When the relevant knowledge or trends of a concept or field are not clear, I will do a retrieval. If the (publication's) title fits my needs very well and the source is authoritative, I will download it. If the [publication's] source is a common journal, I will click to see the summary, and then decide to download the full text or not.*

> *Prof. G: I want to get details of the specific paper. For example, if the abstract was very relevant to the topic I want to study, or it is a preliminary study, I will download the full text.*

> *Prof. W: After reading the title or abstract, if I feel that there are useful things worth reading in-depth.*

> *Student Y: Related to the topic I want to study; Instructive to my research questions; The research question is interesting; Can be used as a reference for paper-writing.*

> *Student W: Generally, if I'm interested in the research methods of the paper, I will download it. Secondly, I will download papers that have related research topics.*

Keywords were extracted by refining respondents' replies and appear in Table 2

The interviews revealed that external attributes of the publications combined with users' perceptions of the publications were the primary factors that promoted publication usage. Based on this interview data

the questionnaire was designed to focus primarily on the two dimensions

5.2. Survey questionnaire

An online survey was designed to answer the research questions. A survey questionnaire was developed using Wenjuanxing (https://www.wjx.cn/). It consisted of four parts: (I) participant background information, (II) download ways, (III) download purposes, and (IV) publication attributes and user perceptions that led users to download literature.

The background questions asked about a user's job position, age, publishing experience, and study discipline. Download frequency was also queried, and only users who had downloaded publications were asked to answer questions about download motivations. The user's probability of full text reading after downloading was also queried to understand respondents' effective downloads, downloads that can be turned into reading.

Questions about download methods were how often respondents downloaded publications in various ways, including hitting the title directly in the full-text database, searching certain fields (excluded title) in the full-text database and downloading selectively, searching through a search engine (e.g., Google), searching through an academic social networking website, shared by other users, and other ways. These questions reflected the diversity of users' download behavior, but also increased understanding about the extent to which the download counts recorded by the database reflect users' full-text acquisition behavior.

To examine factors that motivate users to download, respondents were asked to rate 14 statements describing publication characteristics and user perceptions, with responses ranging from 1 (strongly disagree) to 5 (strongly agree). These statements were categorized into four areas: perceived usefulness, perceived enjoyment, perceived ease of use, and perceived relevance, the four latent variables of the study's model. The full questionnaire is available in Appendix A.

The formal survey questionnaire was disseminated via social media platforms (e. g. WeChat or QQ) to 1025 members of Chinese universities and research institutions on March 2, 2021, with a deadline of March 15, 2021. Participants were required to recall one of the most impressive articles they had recently downloaded and answer the questions based on that article. Two reminders were sent on March 7 and on March 14, 2021. A total of 525 users completed the survey questions by the

Table 1
User information of semi-structured interviewees.

ID	Job position	Research area	Discipline	Age
Prof. G	Supporting staff	Library and information science	Humanities and social sciences	36
Prof. L	Tenure track faculty	Aquaculture	Agricultural science	41
Prof. W	Tenure track faculty	Biomedical science	Medical science	28
Prof. X	Tenure track faculty	Software engineering	Engineering	37
Prof. Y	Tenure track faculty	Statistics	Natural science	34
Student F	Undergraduate student	Software engineering	Engineering	19
Student L	Graduate Student	Marketing management	Humanities and social sciences	26
Student S	Graduate Student	Mathematics	Natural science	24
Student W	Graduate Student	Biomedical science	Medical science	30
Student Y	Graduate Student	Aquaculture	Agricultural science	28

Table 2
Keywords from semi-structured interview content.

Category	Keywords
Publications attributes	title, abstract, source, topic
User perception	curiosity, inspiration, interesting, useful, related

deadline and thus the survey response rate was about 51.22%. 45 responses were considered invalid due to the inordinately short response times (less than 5 minutes) or identical responses to more than 10 questions. As such, there were 480 valid responses for an overall valid response rate of 46.83%. Appendix B presents the demographic information of all valid participants ($N = 480$).

5.3. Data processing and analysis

Partial Least Squares Structural Equation Modeling (PLS-SEM) was adopted to assess the proposed research model and the hypothesized relationships. PLS-SEM is an iterative estimation combining principal component analysis with multiple regression, and a method of causal modeling. In PLS-SEM, models are first evaluated for fit. Upon satisfying fit, individual paths may be evaluated. Compared to covariance-based SEM, PLS-SEM requires a relatively small sample size, does not require variables to exhibit a normal distribution, and is more appropriate for exploratory analysis and for handling formative constructs (Shiau, Sarstedt, & Hair, 2019). Internal consistency, convergent validity, and discriminant validity were examined to assess the quality of the measurement model (Fornell & Larcker, 1981). Generally, Cronbach's α and Composite Reliability (CR) values greater than 0.700 are considered to indicate acceptable reliability of the measurement items. Convergent validity and discriminant validity assessment required that the square root of each AVE all exceeded the inter-construct correlations. The path coefficients were determined by applying the resampling technique of bootstrapping in PLS-SEM to investigate the relationships among factors. All the model analysis processes were conducted using SmartPLS 3.3.3. Hair, Hult, Ringle, and Sarstedt (2022), Henseler, Ringle, and Sarstedt (2012), and Lohmöller (1989) provide detailed explanations on how the PLS-SEM algorithm operates as it is implemented in SmartPLS.

6. Results

6.1. Descriptive statistics analysis

6.1.1. Participant background information

Of the valid participants, 34.59% ($n = 166$) were graduate students and 16.76% ($n = 80$) were tenure-track faculty members. These two user categories were likely to have clear research topics. The remaining participants comprised undergraduate students (32.08%, $n = 154$), supporting staff (6.04%, $n = 29$), and users in other positions (10.63%, $n = 51$). Most participants (66.88%, $n = 321$) were younger than 30 years; no user older than 60 years participated in the survey.

In terms of publishing experience, 42.71% ($n = 205$) of the participants reported being published as the first or corresponding author, followed by 39.17% (n = 154) having no writing experience; 12.92% ($n = 62$) of the sample were unpublished but with writing experience, and 5.21% ($n = 25$) were published but not as the first or corresponding author.

In terms of academic discipline, close to half of the participants (47.08%, $n = 226$) were from the field of natural science, followed by the humanities and social sciences (38.33%, $n = 184$), agricultural science (6.25%, $n = 30$), engineering (4.58%, n = 22), medical science (0.63%, n = 3), and other disciplines (3.13%, n = 15).

Most participants stated that their download frequency was at least once a week (34.17%, $n = 164$), followed by at least once a month (23.54%, $n = 113$), every day (17.29%, $n = 83$), at least once every six months (14.38%, $n = 69$), and at least once a year (7.71%, $n = 37$). All faculty members and graduate students reported literature download experience; only twelve undergraduates and two respondents with other identities (2.92%, $n = 14$) had never downloaded documents. The participants without download experience were excluded from follow-up analysis.

6.1.2. Download methods

As shown in Fig. 2, most participants reported that the primary way they downloaded full text was through academic databases; 68.94% of respondents downloaded selectively after retrieval in the database and 56.38% of respondents downloaded specific documents after hitting titles in the database; 7.02% of respondents said that they obtained full text by seeking help from others. In addition, 21.28% of respondents reported that they were likely to obtain full texts through academic social networking sites such as ResearchGate. Most of the publications used by participants came from academic databases directly. These responses indicated that database download count was a feasible method of reflecting publication usage.

6.1.3. Probability of reading after download

As shown in Fig. 3, 38.91% of respondents had read more than 80% of the publications they downloaded, 25.96% of respondents had read more than 60% and less than 80% of the publications. Only 4.47% of respondents reported that their probability of reading downloaded publications was less than 20%. Thus, 87% of respondents had read at least 40% of the publications they downloaded. This proportion is very similar to the probability of reading after bookmarking (Mohammadi et al., 2016).

6.1.4. Publication characteristics

Of publication characteristics that motivated users to download academic literature (see Fig. 4), most users considered consistency between their research field and the topic of the publication to be the most important characteristic. Articles published in high-quality journals, highly cited, and recommended or shared by someone who could be trusted were three additional characteristics that led users to download.

Earlier publication (older academic literature) was the least attractive characteristic to participants intending to download. It may be because users consider the principle of literature obsolescence (Avramescu, 1979), believing that the older a document, the less useful it is. Avramescu (1979) also considered that this phenomenon confirmed the reliability of the scientific information diffusion model.

6.2. Technology acceptance model analysis

6.2.1. Measurement model

In this study, all Cronbach's α values were above 0.800 and all CR values were greater than 0.900, thus suggesting that each factor employed in the study was acceptably reliable. All AVE values were more than the recommended values of 0.500; CR values also exceed 0.700 (Fornell & Larcker, 1981). Therefore, both CR and AVE values demonstrated acceptable validity for assessing the structural model. Cronbach's α, CR, and AVE values are shown in Table 3.

The square roots of AVEs and the correlation matrix are presented in Table 4. The square roots of each factor's AVEs (the bold texts in Table 4) were greater than the correlations between factors (the regular-face texts in Table 4). These results indicated that the study had convergent and discriminant factor validity. The assessment of the measurement model demonstrated that the factors in this study were reliable and valid in estimating the structural model in the next step.

6.2.2. Structural model

The results of the structural model assessment are shown in Fig. 5.

The total variance explained by the research model was 22.1% ($R^2 = 0.221$, $p < 0.01$), which indicated a closely moderate exploratory power (Chin, Peterson, & Brown, 2008). Perceived usefulness ($\beta = 0.277$, $p < 0.01$) and perceived relevance ($\beta = 0.150$, $p < 0.05$) both had positive statistically significant relationships with user download frequency; perceived enjoyment and perceived ease of use were found to have non-significant relationships with user download frequency.

In terms of the relationships among the four factors, perceived ease of use was found to have a significant positive influence on both

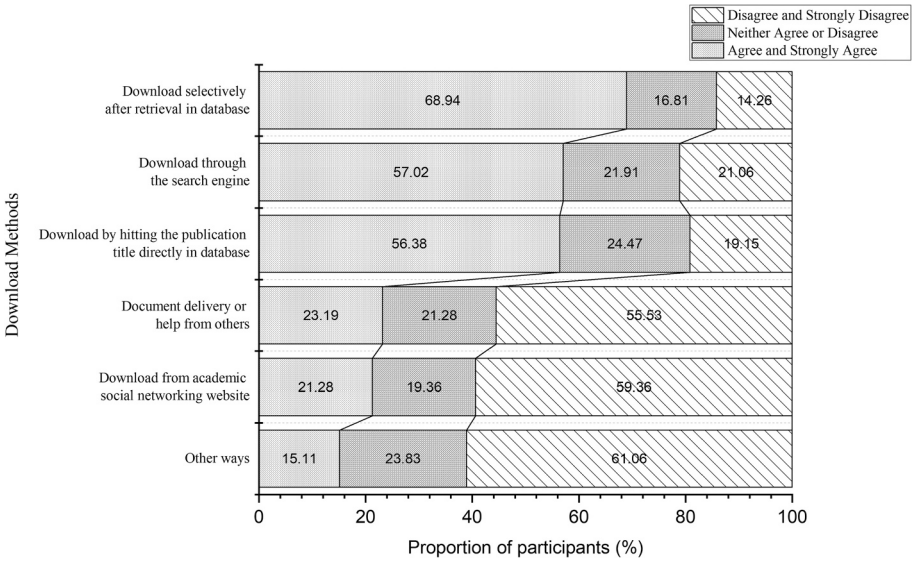

Fig. 2. Download method proportions (n = 480).

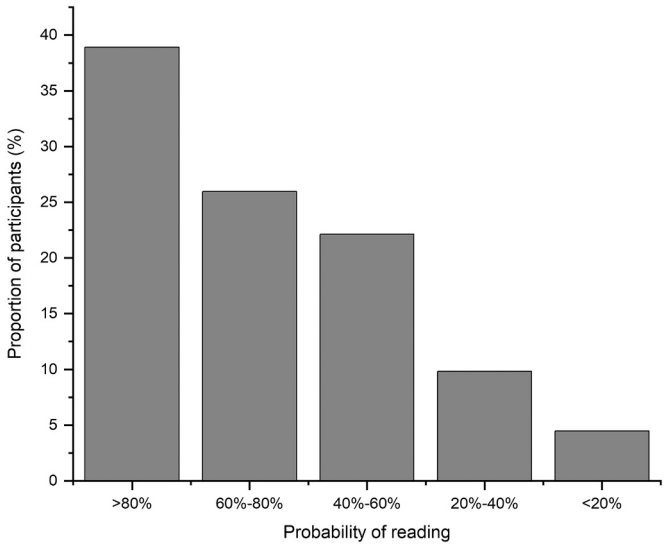

Fig. 3. Probability of reading after download (n = 480).

perceived enjoyment ($\beta = 0.728$, $p < 0.01$) and perceived usefulness ($\beta = 0.405$, $p < 0.01$). Perceived relevance and perceived enjoyment both had a significant positive influence on perceived usefulness ($\beta = 0.321$, $\beta = 0.252$, $p < 0.01$). Therefore, perceived ease of use, perceived enjoyment, and perceived relevance had indirect effects on user download frequency through perceived usefulness; the indirect effect values were 0.112 ($p < 0.01$), 0.070 ($p < 0.01$), and 0.089 ($p < 0.01$), respectively. Perceived usefulness can be explained by perceived ease of use, perceived enjoyment, and perceived relevance. The total variance explained by those three factors is 80.1% ($R^2 = 0.801$, $p < 0.01$). Perceived enjoyment explained 52.9% of the total variance of perceived ease of use ($R^2 = 0.529$, $p < 0.01$).

7. Discussion

7.1. Factors drive users to download literature

Descriptive statistical analysis results demonstrated that topic consistency, publication in high-quality journals, and being highly cited were three key characteristics that led users to download academic literature. They were also shown to be indicators of perceived usefulness. The inference is that perceived usefulness was the key factor driving users to download specific publications; results of TAM analysis confirmed this inference. Although existing research had previously demonstrated that perceived usefulness was the most important motivation in Internet usage (Teo et al., 1999), this study was the first to prove that the conclusion also applied to literature download motivation.

In the context of the literature-use process, the conclusion is easily explained. Aside from knowledge of the research field itself, a user's perception of the usefulness of a document before reading its full text could only be through an external characteristic, such as the citation count and the impact factor of the journal. After downloading, some users choose to cite certain documents that they consider novel, well-known, or concept markers (Case & Higgins, 2000), which are useful to the user. Therefore, the download and citation counts show a moderate correlation (Schloegl et al., 2014) due to their consistency in reflecting the usefulness and perceived usefulness of documents but not the causal relationship between downloads and citations. Since the number of citations is considered as an important indicator to evaluate the impact of publications, this consistency in reflecting the usefulness between downloads and citations also adds the feasibility of downloads in evaluating the impact of publications.

Perceived relevance was also confirmed as an important factor affecting download behavior. This factor can be considered the most fundamental motivation, as, no matter how high the citation rate, downloading will never occur if users perceive that the document is irrelevant to them. As White (2001) described in a study on citation motivation, the consistent motivation underlying all citing is perceived relevance. White (2001) categorized perceived relevance into perceived topical relevance, perceived analogical relevance (parallels are drawn between concepts), perceived methodological relevance, perceived evidentiary relevance, and so on. These types of perceived relevance are related to the content of literature. Users consider that the concepts or methods of the literature are useful to them, so they can also be labeled as offering perceived usefulness. Here, this study proposed a new type of perceived relevance, independent of content: perceived identity relevance. Users may download literature authored, shared, or recommended by someone familiar (i.e., tutor, colleague, classmate, expert). Perceived identity relevance is not the reflection of the usefulness of the document content, but the reflection of the author's social network relationship. Downloads attracted by perceived relevance are easier to convert into citations, as people tend to trust someone they are familiar

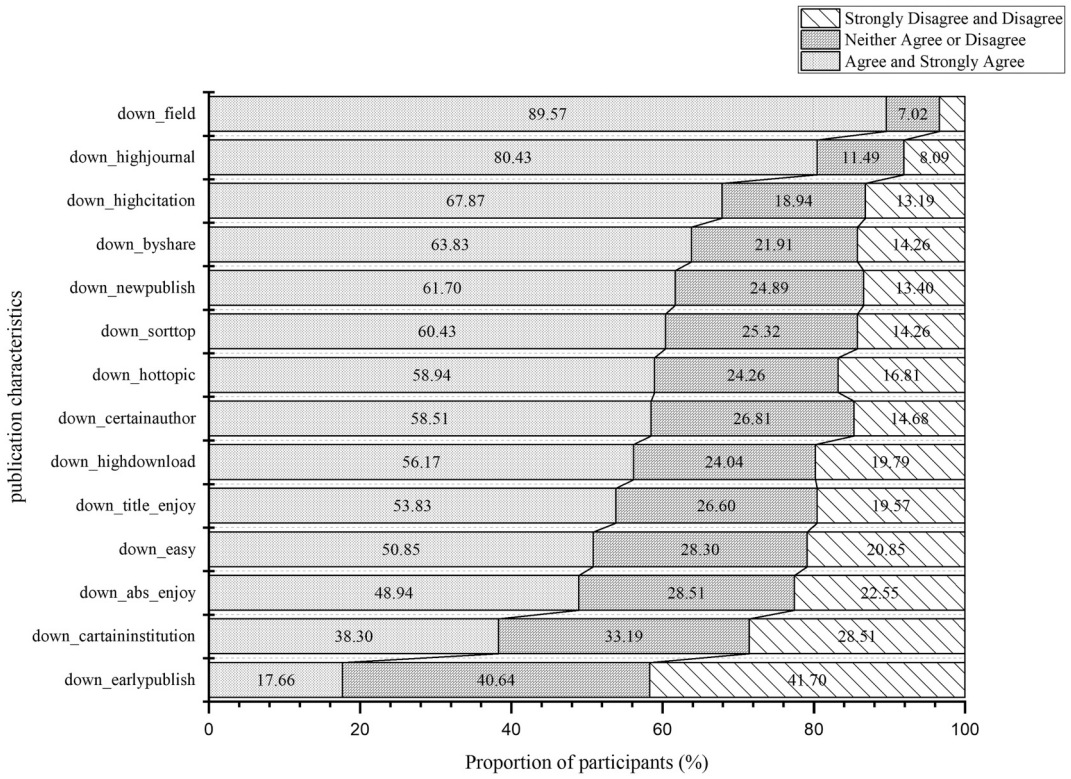

Fig. 4. Publication characteristics that lead users to download (n = 480).
Note: See Appendix C for abbreviations of publication characteristics and corresponding descriptions.

Table 3
Reliability and validity values.

Factors	Cronbach's α	CR	AVE
Perceived enjoyment	0.890	0.932	0.820
Perceived ease of use	0.841	0.904	0.759
Perceived usefulness	0.924	0.943	0.770
Perceived relevance	0.888	0.930	0.817
Download frequency	1.000	1.000	1.000

with, and they are more inclined to cite documents authored, shared, or recommended by familiar people in the face of two similar documents.

Different from the results of Internet use motivation, perceived enjoyment was found to have non-significant relationships with download frequency, perhaps because Internet use is mostly for entertainment purposes, while literature use is mostly for academic purposes. Previous studies hold that humorous titles communicate a nonserious subject matter and are even harmful to the credibility of the paper concerned (Bryant, Brown, Silberberg, & Elliott, 1981; Klein, Bryant, & Zillmann, 1982). Correlation analysis also found that amusing titles reduced the citation of the paper (Sagi & Yechiam, 2008). An interesting title may attract researchers and students to view an abstract, but if they cannot obtain information that is relevant to them or useful for their research, teaching, and other scientific activities, they will not download the full text. In this way, perceived enjoyment indirectly affects users' download behavior through perceived usefulness.

As with perceived enjoyment, users will not download a paper merely because it is perceived as easy to download. However, the easier it is to download a paper, the easier its content can be read by users, and the easier its usefulness can be perceived. For example, open access journal papers receive more citations and downloads (Davis, Lewenstein, Simon, Booth, & Connolly, 2008; Wang, Liu, Mao, & Fang, 2015) not solely because they can be obtained for free but because their usefulness is easier to perceive. As the results of this study reveal, perceived ease of use indirectly affects users' download behavior through perceived usefulness.

In addition to open access, sharing, email pushing, article recommendations, and other promotional activities can improve the perceived ease of use of literature, as they can all save retrieval time, if the literature is useful to users. These efforts at accessibility are effective to some extent in improving the impact of the literature.

In response to the third research question, 64.87% of respondents had read at least 60% of the publications they downloaded. A similar result from the research of Mohammadi et al. (2016) on Mendeley bookmarks showed that 55% of users with a Mendeley library had read or intended to read at least 50% of their bookmarked publications. Thus, total downloads generally reflect the real usage of users. Although readers can obtain required resources via other users' sharing or

Table 4
Square roots of AVEs and correlation matrix.

	Perceived enjoyment	Perceived ease of use	Perceived usefulness	Perceived relevance	Download frequency
Perceived enjoyment	**0.906**				
Perceived ease of use	0.728	**0.871**			
Perceived usefulness	0.778	0.842	**0.877**		
Perceived relevance	0.720	0.789	0.822	**0.904**	
Download frequency	0.399	0.405	0.458	0.431	**1.000**

Note: The boldfaced diagonal values are the square root of AVE. Off-diagonal values are the inter-factor correlations.

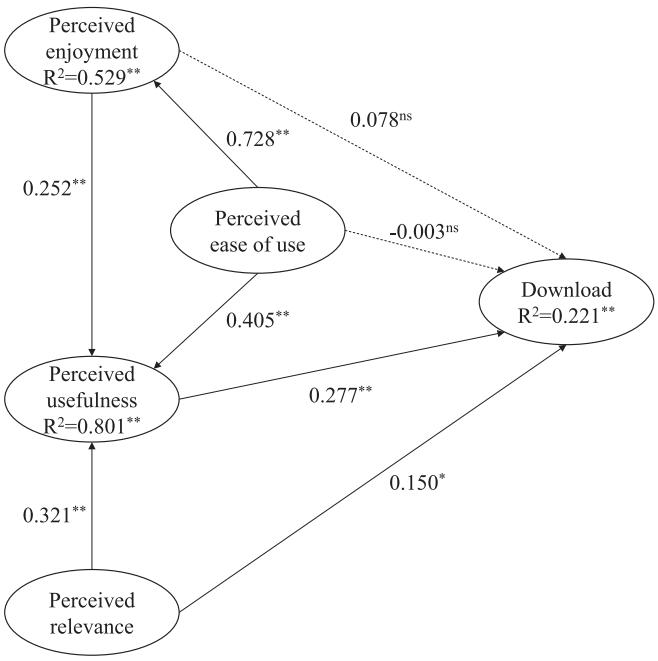

Fig. 5. Structural model results.
Note: ^{ns} = Not significant; * $p < 0.05$; ** $p < 0.01$.

document delivery services, downloading from databases is the predominant method of accessing academic resources using the Internet (see Fig. 2). Downloads recorded by databases generally represent total downloads and encompass various download pathways.

Although users download documents for different purposes, most are academic purposes. The value of literature is in its usage; the more it is used, the more useful it is perceived. The number of downloads recorded by databases has been shown, here and in complementary research, to be an effective indicator for measuring the usefulness and impact of literature.

7.2. Implications and limitations

Findings of this study have certain practical implications for publishers, authors, and librarians. For example, publishers can provide users with personalized subscription function, so that users can timely and easily obtain the latest literature related to their own research; and if the original language of the literature is not English (e.g. Chinese), an English title and abstract can be provided so that the main information of the literature can be obtained by worldwide users more easily. Authors can make an intelligible title for their article, provide more useful information in abstract, and make their publications open access. Librarians should consider the research topics of their service objects to provide more accurate services and using more interesting titles to attract users to read further. These measures may effectively improve the frequency of literature use to a certain extent.

This study had certain limitations that must be highlighted as they would provide some reference for future studies. The first limitation was related to data collection, which was conducted in the Chinese context. Hence, future studies could compare users' motivation for downloading from multiple countries and analyze how these factors differ in different countries. Secondly, it is unclear from this study whether differences in discipline, job position, and writing experience will affect user motivation for downloading, as research samples of this study do not meet the invariance requirements in disciplines, job positions, and writing experience (Aung et al., 2019). Future studies to understand the motivation for using academic literature should include more users from multiple groups.

8. Conclusion

Different from the studies on correlation analysis of downloads and citations, the novelty of this study is proving the causal relationship between different motivations and download frequency with PLS-SEM. According to this study, perceived usefulness is the fundamental factor affecting users' download behavior. This conclusion is similar to the previous studies on Internet use (Teo et al., 1999). However, perceived enjoyment and perceived ease of use need to indirectly affect users' downloads through perceived usefulness, which is different from previous studies (Igbaria et al., 1995; Lee, Cheung, & Chen, 2005). In other words, simply attracting users to download through newfangled titles or other tricks is not feasible. Thus, downloads can be used as an effective metric to reflect the impact of publications. To our knowledge, this study is the first one to investigate download behavior using motivation model. More research on the generation mechanism of altmetric indicators (e.g. sharing, discussing and bookmarking) are expected to be conducted using psychological theories such as motivation theory to enrich the theoretical basis of altmetrics.

Declaration of Competing Interest

None.

Acknowledgments

This research is supported by the Shanghai Office of Philosophy and Social Science (Grant No.2019ETQ004 and 2021ETQ002) and open fund from Key Laboratory of Knowledge Mining and Knowledge Services in Agricultural Converging Publishing, National Press and Publication Administration (Grant No. 2019ETQ004).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lisr.2023.101239.

References

Adams, D. A., Nelson, R. R., & Todd, P. A. (1992). Perceived usefulness, ease of use, and usage of information technology - a replication. *MIS Quarterly, 16*, 227–247.
Aung, H. H., Zheng, H., Erdt, M., Aw, A. S., Sin, S. J., & Theng, Y. (2019). Investigating familiarity and usage of traditional metrics and altmetrics. *Journal of the Association for Information Science and Technology, 70*, 872–887.
Avramescu, A. (1979). Actuality and obsolescence of scientific literature. *Journal of the American Society for Information Science, 30*, 296–303.
Bollen, J., de Sompel, H. V., Smith, J. A., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing & Management, 41*, 1419–1440.
Brody, T., Harnad, S., & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *Journal of the Association for Information Science and Technology, 57*, 1060–1072.
Bryant, J., Brown, D., Silberberg, A. R., & Elliott, S. M. (1981). Effects of humorous illustrations in college textbooks. *Human Communication Research, 8*, 43–57.
Buter, R. K., & van Raan, A. F. J. (2011). Non-alphanumeric characters in titles of scientific publications: An analysis of their occurrence and correlation with citation impact. *Journal of Informetrics, 5*, 608–617.
Case, D. O., & Higgins, G. M. (2000). How can we investigate citation behavior? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science, 51*, 635–645.
Chin, W. W., Peterson, R. A., & Brown, S. P. (2008). Structural equation modeling in marketing: Some practical reminders. *Journal of Marketing Theory and Practice, 16*, 287–298.
Christensson, P. (2014, September 18). Download. TechTerms.Com. https://techterms.com/definition/download.
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*, 2003–2019.
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 319–340.
Davis, P. M., Lewenstein, B. V., Simon, D. H., Booth, J. G., & Connolly, M. J. L. (2008). Open access publishing, article downloads, and citations: Randomised controlled trial. *British Medical Journal, 337*, Article a568.

Duan, Y., & Xiong, Z. (2017). Download patterns of journal papers and their influencing factors. *Scientometrics, 112*, 1761–1775.
Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research, 18*, 382–388.
Guerrero-Bote, V. P., & Moya-Anegon, F. (2014). Relationship between downloads and citations at journal and paper levels, and the influence of language. *Scientometrics, 101*, 1043–1065.
Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Thousand Oaks: Sage.
Haustein, S. (2016). Grand challenges in altmetrics: Heterogeneity, data quality and dependencies. *Scientometrics, 108*, 413–423.
Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics, 101*, 1145–1163.
Henseler, J., Ringle, C. M., & Sarstedt, M. (2012). Using partial least squares path modeling in international advertising research: Basic concepts and recent issues. In S. Okazaki (Ed.), *Handbook of research in international advertising* (pp. 252–276). Cheltenham: Edward Elgar Publishing.
Hsu, C., & Lin, J. C. (2008). Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation. *Information & Management, 45*(1), 65–74.
Htoo, T. H. H., & Na, J. (2017). Who are tweeting research articles and why? *Journal of Information Science Theory and Practice, 5*(3), 48–60.
Igbaria, M., Iivari, J., & Maragahh, H. (1995). Why do individuals use computer technology? A Finnish case study. *Information & Management, 29*, 227–238.
Jamali, H. R., & Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics, 88*, 653–661.
Jin-Cheon, N. (2015). User motivations for tweeting research articles: A content analysis approach. In R. B. Allen, J. Hunter, & M. L. Zeng (Eds.), *Vol. 9469. Digital libraries: Providing quality information. Lecture notes in computer science* (pp. 197–208). Springer.
Klein, D. M., Bryant, J., & Zillmann, D. (1982). Relationship between humor in introductory textbooks and students' evaluations of the texts' appeal and effectiveness. *Psychological Reports, 50*, 235–241.
Klobas, J. E., McGill, T. J., Moghavvemi, S., & Paramanathan, T. (2018). Compulsive YouTube usage: A comparison of use motivation and personality effects. *Computers in Human Behavior, 87*, 129–139.
Kousha, K., Thelwall, M., & Abdoli, M. (2017). Goodreads reviews to assess the wider impacts of books. *Journal of the Association for Information Science and Technology, 68*, 2004–2016.
Kumar, M. S., Gupta, S., Baskaran, S., & Jin-Cheon, N. (2019). User motivation classification and comparison of tweets mentioning research articles in the fields of medicine, chemistry and environmental science. In A. Jatowt, A. Maeda, & S. Yeon Syn (Eds.), *Digital libraries at the crossroads of digital information for the future. Lecture notes in computer science (Vol. 11853, pp. 40–53)*. Springer.
Kurtz, M. J., & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology, 44*, 3–64.
Lee, J., Oh, S., Dong, H., Wang, F., & Burnett, G. (2019). Motivations for self-archiving on an academic social networking site: A study on ResearchGate. *Journal of the Association for Information Science and Technology, 70*, 563–574.
Lee, M., Cheung, C., & Chen, Z. H. (2005). Acceptance of internet-based learning medium: The role of extrinsic and intrinsic motivation. *Information & Management, 42*, 1095–1104.
Lippi, G., & Favaloro, E. J. (2013). Article downloads and citations: Is there any relationship? *Clinica Chimica Acta, 415*, 195. https://doi.org/10.1016/j.cca.2012.10.037
Lohmöller, J. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica-Verlag.
Moed, H. F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. *Journal of the American Society for Information Science and Technology, 56*, 1088–1097.
Moed, H. F., & Halevi, G. (2016). On full text download and citation distributions in scientific-scholarly journals. *Journal of the Association for Information Science and Technology, 67*, 412–431.
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*, 1627–1638.
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*, 1198–1209.
Na, J., & Ye, Y. E. (2017). Content analysis of scholarly discussions of psychological academic articles on Facebook. *Online Information Review, 41*, 337–353.
Naude, F. (2017). Comparing downloads, Mendeley readership and Google scholar citations as indicators of article performance. *The Electronic Journal of Information Systems in Developing Countries, 78*(1), 1–25.
Park, N., Lee, S., & Chung, J. E. (2016). Uses of cellphone texting: An integration of motivations, usage patterns, and psychological outcomes. *Computers in Human Behavior, 62*, 712–719.
Poole, H. (1985). *Theories of the middle range*. Norwood, N.J: Ablex.
Sagi, I., & Yechiam, E. (2008). Amusing titles in scientific journals and article citation. *Journal of Information Science, 34*, 680–687.
Schloegl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: The case of oncology journals. *Scientometrics, 82*, 567–580.
Schloegl, C., & Gorraiz, J. (2011). Global usage versus global citation metrics: The case of pharmacology journals. *Journal of the Association for Information Science and Technology, 62*, 161–170.
Schloegl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014). Comparison of downloads, citations and readership data for two information systems journals. *Scientometrics, 101*, 1113–1128.
Schöpfel, J., & Azeroual, O. (2021). Current research information systems and institutional repositories: From data ingestion to convergence and merger. In D. Baker, & L. Ellis (Eds.), *Future directions in digital information* (pp. 19–37). Chandos Publishing.
Shiau, W., Sarstedt, M., & Hair, J. F. (2019). Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Research, 29*, 398–406.
Subotic, S., & Mukherjee, B. (2014). Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles. *Journal of Information Science, 40*(1), 115–124.
Teo, T., Lim, V., & Lai, R. (1999). Intrinsic and extrinsic motivation in internet usage. *Omega-International Journal of Management Science, 27*(1), 25–37.
Thelwall, M., Haustein, S., Larivière, V., Sugimoto, C. R., & Bornmann, L. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS One, 8*(5), Article e64841.
Torres-Salinas, D., Robinson-Garcia, N., & Gorraiz, J. (2017). Filling the citation gap: Measuring the multidimensional impact of the academic book at institutional level with PlumX. *Scientometrics, 113*, 1371–1384.
Wan, J., Hua, P., Rousseau, R., & Sun, X. (2010). The journal download immediacy index (DII): Experiences using a Chinese full-text database. *Scientometrics, 82*, 555–566.
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*, 917–926.
Wang, X., Liu, C., Mao, W., & Fang, Z. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics, 103*, 555–564.
White, H. D. (2001). Authors as citers over time. *Journal of the American Society for Information Science and Technology, 52*, 87–108.
Xiong, Z., & Duan, Y. (2019). Evaluating the impact of books in Chinese studies: A case study of books authored by UCLA scholars. *Malaysian Journal of Library & Information Science, 24*(3), 23–38.
Zhang, L., & Pentina, I. (2012). Motivations and usage patterns of Weibo. *Cyberpsychology, Behavior and Social Networking, 15*, 312–317.

Zequan Xiong attained a PhD in industrial economics from East China Normal University. He is an associate research librarian at East China Normal University library where he is responsible for information literacy education. His research interests include digital libraries, information behavior analytics, and scientometrics. His articles have been published in several journals, including *Scientometrics*, *Journal of Librarianship and Information Science*, and *Malaysian Journal of Library & Information Science.*

Xia Peng is a reference librarian at East China Normal University. Her research interests include reference service, academic outcome evaluation, and user behavior analytics.

Li Yang is a librarian at East China Normal University. Her research interests include information literacy, information behavior, and academic outcome evaluation.

Wen Lou attained a PhD in information science from Wuhan University. She is an associate professor at Department of Information Management, East China Normal University. Her research and teaching interests are focused on information visualization and organization, scholarly communication, sociology of science. Her articles have been published in several journals, including *Information Processing & Management*, *Journal of the Association for Information Science and Technology*, *Scientometrics*, and *Journal of Information Science*.

Star X. Zhao attained a PhD in information science from Zhejiang University. He is a Professor at Institute of Big Data (IBD), Fudan University. His research interests involve information science, data science, and informetrics. He has published more than 20 English and peer-review papers in journals such as *Journal of the American Society for Information Science and Technology* and *Journal of Informetrics*. He is an editorial board member for several international peer-review journals, including *Journal of Informetrics* and *Frontiers in Scholarly Metrics and Analytics*.