



# Issues in the Interpretation of “Altmetrics” Digital Traces: A Review

Shenmeng Xu\*

*School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States*

Researchers leave traces of their behavior during many stages of their research process. Parts of this process were formerly invisible. With scholarship moving online, we can now access various types of altmetrics digital traces such as reading, organizing, sharing, and discussing scientific papers, thus develop a more holistic story about researchers and their work. However, a lack of in-depth interpretation of altmetrics digital traces is observed. Therefore, this paper focuses on reviewing some of the existing altmetrics research, with a particular emphasis on the issues that need to be taken into consideration in the interpretation of altmetrics digital traces. Taking a preliminary step toward a guideline for more in-depth analysis of digital traces of scholarly acts, this review aims to bring attention to these issues to avoid misuse of altmetrics indicators.

**Keywords:** altmetrics, digital traces, social media, scholarly communication, impact indicator

## INTRODUCTION

The use of non-citation-based metrics to evaluate research is not new. In 1998, Cronin et al. (1998) identified 11 categories of invocation for a purposive sample of five highly cited researchers in library and information science. These categories include abstract, article, conference proceedings, current awareness, external home page, listserv, personal/parent organization home page, resource guide, book review, syllabus, and table of contents. Nowadays, as tools like microblogging, bookmarking, and reference management services are becoming increasingly important in scholars' workflows, we are moving from a reliance on exclusively citation-based metrics to the use of multidimensional digital traces-based metrics. The digital traces researchers leave online reflect the ways in which “academic influence is exercised and acknowledged,” thus bring possibilities to expand the “modes of influence which have historically been backgrounded in narratives of science” (Cronin et al., 1998, p. 1326).

Traditional citation-based metrics, according to Cronin et al. (1998), tells us a lot about “the formal bases of intellectual influence (p.1326).” Nevertheless, many other “modalities of influence which comprise the total impact of an individual's ideas, thinking, and general professional presence (p. 1326)” are overlooked (Cronin et al., 1998). Practitioners of webometrics have been pursuing digital tracks for some time (Almind and Ingwersen, 1997; Björneborn and Ingwersen, 2004; Thelwall et al., 2005; Thelwall, 2008; Khan and Park, 2012). From viewing, storing, discussing, recommending to citing research products, the whole process of user engagement with this research product is mirrored. Since the acts are heterogeneous in nature (Haustein et al., 2015a), they reflect various dimensions and aspects of research impact (Neylon and Wu, 2009). During this process from viewing to recommending, on the one hand, the level of interest in the research product increases, the engagement level of users increases (Lin and Fenner, 2013), and the significance of individual acts increases (Kurtz and Bollen, 2010); On the other hand, the number of counts associated with each product falls (Kurtz and Bollen, 2010), and the coverage of these acts decreases (Priem et al., 2012; Waltman and Costas, 2014; Haustein et al., 2015a; Thelwall and Wilson, 2016).

## OPEN ACCESS

### Edited by:

Xianwen Wang,  
Dalian University of Technology (DUT),  
China

### Reviewed by:

Han Woo Park,  
Yeungnam University, South Korea  
Aparna Basu,  
South Asian University, India

### \*Correspondence:

Shenmeng Xu  
xushenmeng@gmail.com

**Received:** 01 October 2017

**Accepted:** 27 August 2018

**Published:** 04 October 2018

### Citation:

Xu S (2018) Issues in the Interpretation of “Altmetrics” Digital Traces: A Review. *Front. Res. Metr. Anal.* 3:29. doi: 10.3389/frma.2018.00029

Altmetrics provide “new ways of approaching, measuring and providing evidence for impact” (Adie, 2014, p. 349). Although the name of “alternative” metrics seems to imply that they are somewhat replacing traditional citation-based indicators, the widely accepted aim of altmetrics is still to provide “different approaches to different questions” (Crotty, 2014, p. 145). Haustein et al. (2015a) propose a framework which looks at scholarly “acts” on social media through the lens of citation theories and social theories. Specifically, they categorize the scholarly acts into three categories, i.e., access, appraise, and apply, with an increasing level of engagement. Different classifications have been used to categorize social media platforms used for altmetrics. One categorization incorporating the nuanced differences of social media tools were identified by Haustein (2016): social networking platforms (e.g., Facebook, ResearchGate), social bookmarking, and reference management tools (e.g., Mendeley, Zotero), social sharing platforms allowing the sharing of datasets, code, software, presentation slides, figures, and videos, etc. (e.g., Github, Figshare), blogging platforms (e.g., ResearchBlogging, Wordpress), microblogging tools (e.g., Twitter, Weibo), wikis (e.g., Wikipedia), and social recommending, rating, and reviewing websites (e.g., Reddit, F1000Prime).

Altmetrics are widely accepted as indicators of attention and popularity (Crotty, 2014; Gruber, 2014; Sugimoto, 2015). In many cases, this is in accordance with the first criterion listed above. There is a growing body of literature discussing whether altmetrics can be proxies for scholarly impact, societal impact, or generally impact (Bornmann, 2012, 2014; Haustein, 2016). Although altmetrics are considered potential “democratizers of the scientific reward system (p. 413)” (Haustein, 2016), there is not a generally agreed upon conclusion on if and how altmetrics can be used to evaluate various types of impact. Can altmetrics indicate quality? Answering this question is an even further step in investigating what altmetrics can measure. In the discussion of using citation counts as the indicator of “quality,” it is not sufficient to “make the claim in a tautological manner (p. 114)” (Gingras, 2014). Gingras noted that “one must first test the connection between the concept (quality) and the indicator (citations) by finding a relationship between citations and an independent measure of ‘quality,’ already accepted as a valid measure” (p. 114). Sociological and bibliometrics studies since the 1970s have consistently shown that there is a correlation between how often an author is cited and how renowned he or she is, as measured by other indicators of eminence like important prizes and awards or academic nominations to scientific academies (Cole and Cole, 1974). This is probably a potential direction for future studies to observe if altmetrics can measure scholarly merit.

Good metrics should be able to capture the appropriate features of science (Lane et al., 2014). This criterion is applicable to both citation-based metrics and the digital trace-based metrics, which can provide a more comprehensive picture of impact. According to Moed and Plume (2011), there are three types of bibliometric indicators: basic indicators, relative, or normalized indicators, and indicators based on advanced network analysis. Digital trace-based metrics, similarly, can fit into this framework.

Although to date, altmetrics still lacks a cohesive body of theory, the history of bibliometrics or scientometrics has demonstrated that the lack of theory does not keep metrics from being useful in practice. In this review, our scope focuses on the first type—basic indicators based on digital traces.

Two comprehensive relevant review articles have been published by Bornmann (2014) and Haustein (2016). Bornmann (2014) discussed the disadvantages of altmetrics: commercialization, data quality, missing evidence, and manipulation. Haustein (2016) conducted a comprehensive review of the challenges in altmetrics, i.e., heterogeneity, data quality, and dependencies, and discussed them with a particular emphasis on past developments in bibliometrics.

This paper extracts and integrates the issues in the interpretation of altmetrics digital traces. The purpose of this review is not to question or critique the existence of a large number of indicators and empirical studies. Instead, this review sets out with the aim of pointing to the lack of robust understanding we have about altmetrics. Four major issues are discussed in four main sections of this review, including influencing factors, necessity and sufficiency, academic vs. societal impact, as well as reliability and validity. Understanding these issues is crucial to the assessment of the applicability and limitations of empirical altmetrics studies.

## INFLUENCING FACTORS

Bornmann (2014) discussed factors that might affect the probability of an article being cited in his comprehensive review: time-dependent factors, field-dependent factors, journal-dependent factors, article-dependent factors, author-reader dependent factors, availability of publications, and technical problems. Similarly, the results of altmetrics research should be interpreted in light of the following aspects in order to better understand and apply the results.

### Time

Considering the exponential increase in both scientific output and social media tools usage, an increasing amount of occurrences of altmetrics digital trace can be expected from year to year. Generally, Altmetrics and citation-based metrics show different obsolescence functions (Small, 2016). For instance, Moed (2005) and Schloegl and Gorraiz (2011) examined the differences between cited and usage half-lives. The medium correlations confirm that downloads measure a different impact than citations. Wan et al. (2010) investigated 6000 Chinese academic articles indexed in the Chinese full-text database CNKI and found that correlations were strongest when they were computed through normalized cross-covariance between citation and download curves. In other words, when the citation curve is shifted backward by 2 or 3 years, the correlations are the highest.

Due to the heterogeneous nature of the altmetrics traces, they also display different levels of immediacy. For instance, Yu et al. (2017) compared Weibo altmetrics with Twitter altmetrics. Their findings indicate that Weibo altmetrics is more immediate compared with the general altmetrics (all source that Altmetric.com aggregates). Specifically, 60% of articles with

Weibo attention were captured within 180 days, while 46% of general altmetrics happen more than 360 days after the publication date.

Although there are plenty of studies on the correlation between citations and altmetrics, as well as between various altmetrics, the impact flow is not well-understood. In other words, despite some empirical studies (e.g., Eysenbach, 2011), which showed that highly-tweeted articles were 11 times more likely to become highly-cited in the future), the evidence is still lacking about how exactly citations and research-related social media traces affect each other. It has already been shown that the more frequently a publication is cited, the more frequently it will be cited later (Garfield, 1981; Cano and Lind, 1991; Burrell, 2003). Taking this "success-breeds-success" phenomenon (Cozzens, 1985), or Matthew Effect, into consideration, the effect of time becomes an even more complex phenomenon that needs to be explored in more depth.

## Disciplines

The practice and norms of research differ in different disciplines. For instance, according to Thelwall and Wilson (2016), Medical research "is heavily funded by governments, charities, and private companies, presumably because it can lead to improvements in lifespan and quality of life and because some medical discoveries, such as new drugs, equipment, and treatments, can be highly profitable (p. 1962)." A large number of altmetrics studies focus on medical-related disciplines, partly due to the extensive citation and altmetrics data existing. For instance, Haustein et al. (2014) found a correlation of 0.39 between the number of social bookmarks and citations for PubMed papers; Excluding papers without Mendeley readers at the time of their research, the correlation was found to be 0.46. Thelwall and Wilson (2016) analyzed all medical articles in Web of Science (WoS) and reported a more detailed subject breakdown (according to the 45 Scopus Subject Categories).

Just as citation practices which vary between different disciplines in science, social sciences, and humanities (Hurt, 1987; Ziman, 2000), the usage patterns of altmetrics tools by different disciplines are quite different. Brody et al. (2006) found a correlation of 0.46 between downloads and citations from almost 15,000 physics preprints on arXiv and Citebase citations. Bar-Ilan (2012) reported a correlation of 0.46 between Mendeley readership counts and citations for articles published in the Journal of the Association for Information Science and Technology. Meanwhile, Mohammadi and Thelwall (2014) found a correlation of 0.52 for publications in the social sciences and a correlation of 0.43 for publications in the humanities. Haustein et al. (2015b) explored bibliographic and citation data from 1.3 million papers indexed in WoS as well as altmetrics data from Altmetric.com. They found that on the contrary to what is observed for citations, articles in Social Sciences and Humanities were the most often found on social media platforms. Na and Ye (2017) found a predominance of public engagement in discussions of psychological academic articles on Facebook, and thus concluded that Facebook metric better reflected the attitudes or perceptions of the general public instead of academia at least in the discipline of psychology.

In addition, disciplines covering more general topics reach more readers than those focusing on a more specific area of research. Just as the chance of being cited by others is related to the number of publications in the field (Moed et al., 1985), smaller fields tend to attract fewer altmetrics activities than more general fields.

## Platforms and Acts

One of the challenges in the meaningful use of altmetrics is the heterogeneity of the underlying acts (Bornmann, 2014; Haustein, 2016). As altmetrics derived from different social media platforms are shaped by significantly different premises, the interpretation of altmetrics is a difficult endeavor. Different actions on different platforms are in many cases fundamentally different regarding the user's intention and degree of involvement. For example, the motivation of bookmarking a publication in Mendeley can be different from blogging the same publication; the level of engagement is also significantly different in that bookmarking takes one click while blogging takes much more effort. Empirical studies have found moderate to high correlations between citations and Mendeley readership, as well as between citations and F1000 Prime recommendations (Li and Thelwall, 2012; Bornmann et al., 2013; Thelwall and Wilson, 2016); In contrast, correlations between citations and tweets were found to be weak (Costas et al., 2015; Haustein et al., 2015a). Several studies correlating the number of downloads with the number of citations have found positive relationships, but the correlations are too weak to conclude that downloads and citations measure the same thing (Li et al., 2011). Similarly, Thelwall et al. (2013) found statistically significant but low correlations between citation counts and tweet counts for PubMed articles. Other than Mendeley and Twitter, other studies have also examined citations from Wikipedia articles (e.g., Nielsen, 2007) and blogs (e.g., Groth and Gurney, 2010; Shema et al., 2012).

Mendeley readership appears to be the most common altmetrics compared to others. Li and Thelwall (2012) found that 1,389 of the sampled 1,397 F1000 Genomics and Genetics papers were covered in Mendeley. Priem et al. (2012) reported a Mendeley coverage of 80% for articles published in the Public Library of Science (PLoS) journals. In addition, Mohammadi and Thelwall (2014) investigated Mendeley coverage of articles in the social sciences and the humanities indexed by WoS published in 2008 and reported coverages of 58 and 28% for the social sciences and humanities, respectively. Zahedi et al. (2014) sampled 20,000 publications indexed by WoS and found that 37% of them were covered in Mendeley, which was the highest among all altmetrics sources. Additionally, Haustein et al. (2014) reported a coverage of 66% for the 1.4 million papers published between 2010 and 2012 and indexed by PubMed. In addition, Mendeley readership of articles appears to have the highest correlation with citation counts than do other altmetrics (Li and Thelwall, 2012; Zahedi et al., 2014; Mohammadi et al., 2015; Maflahi and Thelwall, 2016).

The scholarly acts are different in nature due to the functions and affordances of different platforms (Haustein et al., 2015a). The absence of contextual information can oftentimes lead to

an incomplete understanding of the users' acts. According to Kurtz and Bollen (2010), "usage statistics" lack the "individual event information (p. 8)" (e.g., user information and session information) and is thus different from "usage data." Gunn (2013) also notes the difference between content-rich altmetrics (e.g., blog posts or Wikipedia articles) and others as content-poor altmetrics (e.g., tweets or Facebook's "likes"). Special caution needs to be used when analyzing the latter form of altmetrics.

## Platforms and Users

Haustein et al. (2014) found that while some papers received attention on Twitter because of their health implications or topicality, others seemed to be discussed on Twitter due to humorous or curious contents. This suggests that tweeters of academic articles do not necessarily engage in intellectual discussions, and the tweets do not necessarily reflect intellectual impact.

Bollen and Van de Sompel (2008) calculated the Usage Impact Factors using download statistics from nine California State University institutions. At the discipline level, education was the only one out of 17 disciplines studied that exhibited a significant positive correlation. They also found that journals with low IFs tend to be more useful for undergraduate teaching than those with high IFs because disciplines with relatively large graduate populations displayed positive correlations while those with relatively large undergraduate populations displayed more negative correlations (Bollen and Van de Sompel, 2008). Mohammadi et al. (2015) analyzed the professions of Mendeley users and calculated the readership counts and citations for professors, postdoctoral researchers, Ph.D. students, postgraduate students, and undergraduate students. Their results indicate that except for undergraduate students, the other professions all display high and significant correlations between Mendeley readership and citations. In the above-mentioned study of medical articles (in "Disciplines"), Thelwall and Wilson (2016) removed student readers from the Mendeley data and found a slightly decreased correlation between Mendeley readership and citation counts.

Lemke et al. (2017) conducted a survey of 3,400 researchers. Their results exhibited statistically significant differences in the frequency of usage of certain kinds of social media-related acts between early stage researchers (Ph.D. students and research assistants) and professors: while early stage researchers make more frequent use of download functionalities on various platforms, professors more often engage in publication-related interactions of diverse kinds on Facebook, Twitter, and LinkedIn (including writing posts/tweets about academic research, commenting on posts/tweets about academic research or liking/favoring such posts/tweets). As a result, Lemke et al. (2017) suggested using findings like this to specify more precise applications of altmetrics. Specifically, download counts could be used to express a publication's scientific impact in a way that emphasizes its relevance among early-stage researchers; The number of tweets about an article could be considered as a metric which better reflects that article's impact among relatively more established researchers.

## Other Factors

Research products of different types were reported to have different altmetrics patterns (Haustein et al., 2015b; Xu and Hemminger, 2015). Specifically, Xu and Hemminger (2015) investigated eight types of publications: research articles, review articles, opinion articles, educational articles, community pages, editorials and letters, synopses, and journal documents. Their results showed that compared to other types, review articles had the highest median and mean of views, saves, shares, as well as citations. Moreover, educational articles were highly saved but not as highly cited; Opinion articles are highly viewed but not as highly saved. The correlations among them also displayed different patterns—they thus suggested taking article types into consideration to assist in the interpretation of scholarly impact. In the abovementioned study (in "Disciplines"), Haustein et al. (2015b) found that while editorials and news items were seldom cited, they had the highest popularity on Twitter.

The type and topic of research can also have an influence on the altmetrics. For instance, Liu et al. (2011) investigated the downloads and citations of Chinese ophthalmology journal publications and found that those with high usage but low citation rates had an application orientation or else contained news and summaries about important conferences. Vainio and Holmberg (2017) found that scientific articles were tweeted to "promote ideological views especially in instances where the article represented a topic that divides general opinion (p. 345)."

Moreover, Haustein et al. (2015b) analyzed the main patterns of five social media metrics as a function of document characteristics including number of pages and references, title length, and number of authors, institutions and countries, as well as collaborative practices. Their findings indicate that just like citations, social media metrics increase with the extent of collaboration and the length of the references list. Meanwhile, longer papers typically attract less social media attention although an opposite trend is observed in citations (Haustein et al., 2015b).

## NECESSITY AND SUFFICIENCY

In the research process, acts from viewing, storing, discussing, recommending to citing research products are associated with an increasing significance of individual occurrences (Kurtz and Bollen, 2010). In other words, the latter acts might indicate a higher level of impact on the user than the former ones. When interpreting impact indicators, many previous studies use an underlying assumption that the acts represent the corresponding level of impact: download and click rates represent impact in the form of readership, social mentions and discussion represent a higher level of interest in the research output, and citations represent an even higher level of impact.

Nevertheless, none of the acts mentioned above can be used as a perfect necessary condition for the occurrence of impact. When discussing the possible explanations of differences between download and citation distributions, Small (2016) mentioned 10 factors, one of which being that downloading a document does not equal to reading the document. The social

constructivist citation theory believes that scientific knowledge is socially constructed through the manipulation of political and financial resources (Knorr-Cetina, 1991). Particularly, there are different motivations for citing, and citations can be perfunctory (Murugesan and Moravcsik, 1978). Similarly, scholarly acts other than citing can also be nonessential and meaningless. In other words, download and click rates estimate readership; they do not measure it (Thelwall, 2012). An empirical study example would be an extensive deep log analysis conducted by Nicholas et al. (2008), which revealed that two-thirds of all article views actually lasted less than three minutes. Their results indicate that a considerable amount of full-text access is cursory and cannot be used unconditionally to represent readership.

In addition, the former acts are not always the necessary condition of all the latter acts. For instance, one can recommend an article on Twitter without having to save it to a reference management tool or personal computer; One can cite an article without having to read it.

When talking about citation-based metrics and impact, Cronin and Sugimoto (2014) point out that although citation correlates positively with impact, it is still only an approximation of impact. This reasoning also fits in altmetrics studies. Considering the logic of necessity and sufficiency, there is no guarantee that a single indicator can tell whether the user was in fact influenced by the article. The ostensible meanings of digital traces can sometimes be deceptive and need to be interpreted with caution.

## ACADEMIC VS. SOCIETAL IMPACT

Before the term "altmetrics" was coined, the Public Library of Science (PLOS) began to offer Article-Level-Metrics (ALM) in 2009 to provide the research community with a view into the reach of their publications. Lin and Fenner (2013) grouped the types of engagement captured by the PLOS ALM data sources into five groups: viewed, saved, discussed, recommended, and cited. In their categorization, all the data sources are grouped together without distinguishing between the impact on scholars and the public. ImpactStory, co-founded by Jason Priem who coined the term "altmetrics," provides altmetrics data to help researchers measure and share the impacts of all their research outputs. ImpactStory used a similar categorization to PLOS, grouping data sources into the same five groups. In addition, they divide the data sources in each group into two subgroups—impact on scholars and impact on the public. For instance, "PDF downloads" is categorized as "impact on scholars," and "HTML downloads" is categorized as "impact on the public." This distinction is considered somewhat artificial (Bornmann, 2014) because PDF documents are not only downloaded by academics nor are HTML versions only read by the public. Another example is that Wikipedia is included under "cited" "by the public" category by ImpactStory, while PLOS has it in the "discussed" category. Since not all Wikipedia editors are the general public, the interpretation of such data still needs further research.

The societal impact of research is concerned with "the assessment of social, cultural, environmental, and economic returns (impact and effects) from results (research output) or products (research outcome) of publicly funded research" (Bornmann, 2012, p. 673). Governments and funding agencies increasingly expect scholars to expound and demonstrate societal impact and relevance of their work (Higher Education Funding Council for England, 2011; Bornmann, 2012; Sugimoto et al., 2017). Bornmann (2012) provided a comprehensive review on how societal impact of research is assessed and how diverse names are used when describing societal impact: third stream activities, societal benefits, societal quality, usefulness, public values, knowledge transfer, and societal relevance.

One of the problems with altmetrics is the representativeness of data in terms of who is using the resource (Haustein, 2014). Studies have shown that tweets to scientific papers are created mostly by academics (Alperin, 2015; Tsou et al., 2015; Vainio and Holmberg, 2017), but it has also been shown that in some disciplines, non-academic users dominate in the discussion (Alperin and Haustein, 2017; Na and Ye, 2017). Alperin (2015) conducted surveys on Twitter to ask users who had recently shared academic articles if they were affiliated with a university. In the 286 responses he got, 184 (64%) were affiliated with a university. Tsou et al. (2015) investigated the tweeters who had tweeted at least one link to an article in four leading journals (Nature, Science, PNAS, and PLoS One). They identified 34.4% of the tweeters as Ph.D. degree owners. In addition, they found that the tweeters were more male-biased (70%), despite the fact that women are overall slightly more likely to use social networking sites than men (Kimbrough et al., 2013; Pew Research Center, 2014). Based on these two findings, they concluded that the demographics of the tweeters studied did not reflect the general population of Twitter users—instead, they included more academics. Vainio and Holmberg (2017) examined tweeters of Finnish produced articles (or collaborated articles) in four areas of science (Agricultural, Engineering, and Technological Science; Medical and Health Sciences; Natural Sciences; Social Sciences and Humanities). They used keyword categorization, co-word analysis, and content analysis to study the user profile descriptions, and found that researchers were strongly represented among Twitter users responsible for tweeting scientific articles. Different from the findings above, Na and Ye (2017) found a predominance of public engagement in discussions of psychological academic articles on Facebook. They conducted a content analysis on 1,711 Facebook users and found 71.4% them to be non-academic users. The motivation of discussing psychological articles was also investigated: discussion and evaluation toward articles (20.4%), application to real life practices (16.5%), self-promotion (6.4%), and data source exchange (6.0%), and most significantly, perfunctory sharing without additional user comments (50.1%). Alperin and Haustein (2017) used social network analysis to analyze Twitter diffusion patterns. Specifically, they studied tweets tweeting seven highly tweeted articles published in the open access journal BMC Biology. Their results confirmed that research on Twitter is shared primarily among academic communities of users

who were already well-connected outside of Twitter. However, they also found that certain publications were able to gain the attention of more diverse communities and disconnected users.

The paragraph above describes some research on users on general social media like Facebook and Twitter. Besides these tools, there are social networking tools and platforms targeting at scholars—like Mendeley, F1000, ResearchGate, and Academia.edu—where interactions with scientific publications are committed by researchers (Sugimoto et al., 2017). Bornmann (2014) conducted an empirical study using data from F1000, Altmetric, and Web of Science. The findings indicated that papers tailored for readers outside the specific research area led to societal impact because papers with the tag “good for teaching” by F1000 experts received higher altmetric counts. In contrast, papers with the tag “new finding,” which was relatively more scientifically oriented, tended to have higher citation counts.

## RELIABILITY AND VALIDITY

Data quality issues have been thoroughly discussed in some previous review articles (Bornmann, 2014; Haustein, 2016). Haustein (2016) has discussed the dynamic nature of social media events and how they can affect the accuracy, consistency, and replicability of various altmetrics.

Altmetrics depend on the availability of Digital Object Identifiers (DOI) and “are shaped by technical possibilities (p. 413)” (Haustein, 2016). The technical affordances of the various underlying platforms and the different ways the data providers and aggregators work determine that the retrieval of data from altmetrics platforms normally requires a certain amount of data cleaning. For instance, incompleteness and errors have been found in the metadata of bibliographic entries in online reference managers. This can potentially cause a publication bookmarked by more than one user not to be recognized as the same one. Haustein and Siebenlist (2011) showed that it was better to apply a search strategy based on different metadata fields to retrieve bookmarks on CiteULike, Connotea, and BibSonomy. Similarly, Bar-Ilan (2012) showed that the 33% of the records retrieved from Mendeley Application Program Interface (API) did not contain a document object identifier (DOI) and would be missed by the API altogether, and that the use of the Mendeley could result in the loss of a significant number of data. The retrieval of data from Twitter, Facebook, and other social media platforms can be even more complex and problematic.

Gaming is one of the important issues that can undermine the reliability and validity of altmetrics as indicators of research impact Haustein (2016). Borrego and Fry (2012) found that the majority (78%) of bookmarks in BibSonomy were created by only 14 users; it was suspected that these users were managers of digital libraries keen to enhance usage, given that the entries were created within a few days.

Automated agents (also called robots or bots) has been listed as a major concern regarding the validity of altmetrics (Darling et al., 2013). Haustein et al. (2014) analyzed how arXiv and

journal versions of scientific papers were tweeted. Their findings revealed a series of automatic Twitter accounts such as @hep\_th, @hep\_ph, @hep\_ex, and @hep\_lat, which automatically tweet new submissions to arXiv. Similarly, Xu et al. (2018) examined how video articles in the Journal of Visualized Experiments were tweeted and found at least seven out of the top 10 tweeters of JoVE articles to be bots. As bots tweet articles without human selection, they undermine the function of tweet counts as a filter or indicator of impact as suggested in the altmetrics manifesto (Priem et al., 2010). While bots contribute positively to Twitter by creating “a large volume of benign tweets, like news and blog updates” (Chu et al., 2012, p. 812), they can potentially have a big effect on altmetrics calculations if not properly recognized and discounted.

## DISCUSSION

No single metric can provide the whole picture. Simplifying a complex system by applying metrics and indicators to it can be a promising method to know better about what we are measuring. However, oversimplifying this system can be dangerous. Two Albert Einstein quotes might be used to express the tradeoff: “Not everything that can be counted counts and not everything that counts can be counted”; “Make everything as simple as possible, but not simpler” (Shapiro, 2006, p. 231; Kurtz and Bollen, 2010). To ameliorate the lack of a clear conceptual or theoretical framework for altmetrics, more enriched, complex, but more accurate and reliable measures are needed.

It is important to exclude “empty buzz” and learn from the contextual clues of digital traces when reading stories about the impact of research products (Priem et al., 2010; Bornmann, 2014). It is more meaningful to understand who has used a research product, how and why it is used, what effect it has had, rather than simply knowing how many people have viewed, downloaded, or mentioned it on Twitter. Currently, the gold standard in establishing valid meanings of trace data remains empirical research. This review would like to particularly point to the importance of qualitative research of altmetrics digital traces, which could help provide an “interpretative lens” to understand the motivations behind the social media acts.

Last but not least, it is essential to keep an open mind in the interpretation of altmetrics. For instance, digital access to documents via self-archiving and print access is largely ignored when studying readership. Limitations like this should also be part of our interpretation of what altmetrics can truly measure.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## ACKNOWLEDGMENTS

The author would like to thank Dr. Bradley M. Hemminger and Dr. Cassidy R. Sugimoto for their valuable comments and suggestions.

## REFERENCES

- Adie, E. (2014). Taking the alternative mainstream. *Prof. Inf.* 23, 349–351. doi: 10.3145/epi.2014.jul.01
- Almind, T. C., and Ingwersen, P. (1997). Informetric analyses on the world wide web: methodological approaches to “webometrics”. *J. Doc.* 53, 404–426. doi: 10.1108/EUM000000007205
- Alperin, J. P. (2015). “Moving beyond counts: a method for surveying twitter users” in *Altmetrics15 Workshop* (Amsterdam).
- Alperin, J. P., and Haustein, S. (2017). “Applying social network analysis to explore Twitter diffusion patterns,” in *The 2017 Altmetrics Workshop* (Toronto).
- Bar-Ilan, J. (2012). “JASIST@ mendeley,” in *ACM Web Science Conference 2012 Workshop* (Evanston).
- Björneborn, L., and Ingwersen, P. (2004). Toward a basic framework for webometrics. *J. Assoc. Inf. Sci. Technol.* 55, 1216–1227. doi: 10.1002/asi.20077
- Bollen, J., and Van de Sompel, H. (2008). Usage impact factor: the effects of sample characteristics on usage-based impact metrics. *J. Assoc. Inf. Sci. Technol.* 59, 136–149. doi: 10.1002/asi.20746
- Bornmann, L. (2012). Measuring the societal impact of research. *EMBO Rep.* 13, 673–676. doi: 10.1038/embor.2012.99
- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *J. Informetr.* 8, 895–903. doi: 10.1016/j.joi.2014.09.005
- Bornmann, L., Leydesdorff, L., and Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: opportunities and limits. *J. Informetr.* 7, 158–165. doi: 10.1016/j.joi.2012.10.001
- Borrego, Á., and Fry, J. (2012). Measuring researchers’ use of scholarly information through social bookmarking data: a case study of BibSonomy. *J. Inf. Sci.* 38, 297–308. doi: 10.1177/0165551512438353
- Brody, T., Harnad, S., and Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *J. Assoc. Inf. Sci. Technol.* 57, 1060–1072. doi: 10.1002/asi.20373
- Burrell, Q. L. (2003). Predicting future citation behavior. *J. Assoc. Inf. Sci. Technol.* 54, 372–378. doi: 10.1002/asi.10207
- Cano, V., and Lind, N. (1991). Citation life cycles of ten citation classics. *Scientometrics* 22, 297–312. doi: 10.1007/BF02020003
- Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. (2012). Detecting automation of twitter accounts: are you a human, bot, or cyborg? *IEEE Trans. Depend. Secure Comput.* 9, 811–824.
- Cole, J. R., and Cole, S. (1974). Social stratification in science. *Am. J. Phys.* 42, 923–924. doi: 10.1119/1.1987897
- Costas, R., Zahedi, Z., and Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *J. Assoc. Inf. Sci. Technol.* 66, 2003–2019. doi: 10.1002/asi.23309
- Cozzens, S. E. (1985). Changing order: replication and induction in scientific practice. *Science* 230, 1267–1268. doi: 10.1126/science.230.4731.1267
- Cronin, B., Snyder, H. W., Rosenbaum, H., Martinson, A., and Callahan, E. (1998). Invoked on the Web. *J. Assoc. Inf. Sci. Technol.* 49, 1319–1328. doi: 10.1002/(SICI)1097-4571(1998)49:14<1319::AID-ASI9>3.0.CO;2-W
- Cronin, B., and Sugimoto, C. R. (eds.). (2014). *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*. Cambridge: MIT Press.
- Crotty, D. (2014). Altmetrics: finding meaningful needles in the data haystack. *Serials Rev.* 40, 141–146. doi: 10.1080/00987913.2014.947839
- Darling, E. S., Shiffman, D., Côté, I. M., and Drew, J. A. (2013). The role of Twitter in the life cycle of a scientific publication. *arXiv:1305.0435* [preprint].
- Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *J. Med. Internet Res.* 13:e123. doi: 10.2196/jmir.2012
- Garfield, E. (1981). Citation-classics-4 years of the human side of science. *Curr. Contents* 22, 5–16.
- Gingras, Y. (2014). “Criteria for evaluating indicators,” in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, eds B. Cronin and C. R. Sugimoto (Cambridge: MIT Press), 109–125.
- Groth, P., and Gurney, T. (2010). “Studying scientific discourse on the web using bibliometrics: A chemistry blogging case study,” in *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line* (Raleigh, NC).
- Gruber, T. (2014). Academic sell-out: how an obsession with metrics and rankings is damaging academia. *J. Market. Higher Educ.* 24, 165–177. doi: 10.1080/08841241.2014.970248
- Gunn, W. (2013). Social signals reflect academic impact: what it means when a scholar adds a paper to Mendeley. *Inf. Stand. Q.* 25, 33–39. doi: 10.3789/isqv25no2.2013.06
- Haustein, S. (2014). “Readership metrics,” in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, eds B. Cronin and C. R. Sugimoto (Cambridge: MIT Press), 327–344.
- Haustein, S. (2016). Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics* 108, 413–423. doi: 10.1007/s11192-016-1910-9
- Haustein, S., Costas, R., and Larivière, V. (2015a). “Interpreting “altmetrics”: viewing acts on social media through the lens of citation and social theories,” in *Theories of Informetrics: A Festschrift in Honor of Blaise Cronin*, ed C. R. Sugimoto (Berlin; Boston, MA: Walter de Gruyter GmbH), 372–405.
- Haustein, S., Costas, R., and Larivière, V. (2015b). Characterizing social media metrics of scholarly papers: the effect of document properties and collaboration patterns. *PLoS ONE* 10:e0120495. doi: 10.1371/journal.pone.0120495
- Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., and Larivière, V. (2014). Tweeting biomedicine: an analysis of tweets and citations in the biomedical literature. *J. Assoc. Inf. Sci. Technol.* 65, 656–669. doi: 10.1002/asi.23101
- Haustein, S., and Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *J. Informetr.* 5, 446–457. doi: 10.1016/j.joi.2011.04.002
- Higher Education Funding Council for England (2011). *Decisions on Assessing Research Impact. Research Excellent Framework (REF) 2014 (Research Excellence Framework No. 01.2011)*. Available online at: [http://www.ref.ac.uk/media/ref/content/pub/decisionsonassessingresearchimpact/01\\_11.pdf](http://www.ref.ac.uk/media/ref/content/pub/decisionsonassessingresearchimpact/01_11.pdf)
- Hurt, C. D. (1987). Conceptual citation differences in science, technology, and social sciences literature. *Inf. Process. Manage.* 23, 1–6. doi: 10.1016/0306-4573(87)90033-1
- Khan, G. F., and Park, H. W. (2012). Triple Helix and innovation in Asia using scientometrics, webometrics, and informetrics. *Scientometrics* 90, 1–7. doi: 10.1007/s11192-011-0506-7
- Kimbrough, A. M., Guadagno, R. E., Muscanell, N. L., and Dill, J. (2013). Gender differences in mediated communication: women connect more than do men. *Comput. Human Behav.* 29, 896–900. doi: 10.1016/j.chb.2012.12.005
- Knorr-Cetina, K. D. (1991). Epistemic cultures: forms of reason in science. *Hist. Polit. Econ.* 23, 105–122. doi: 10.1215/00182702-23-1-105
- Kurtz, M. J., and Bollen, J. (2010). Usage bibliometrics. *Ann. Rev. Inf. Sci. Technol.* 44, 1–64. doi: 10.1002/aris.2010.1440440108
- Lane, J., Largent, M., and Rosen, R. (2014). “Science metrics and science policy,” in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, eds B. Cronin and C. R. Sugimoto (Cambridge: MIT Press), 397–411.
- Lenke, S., Mehrazar, M., and Peters, I. (2017). “Evaluating altmetrics acts through their creators – how to advance,” in *The 2017 Altmetrics Workshop* (Toronto).
- Li, X., and Thelwall, M. (2012). “F1000, Mendeley and traditional bibliometric indicators,” in *Proceedings of the 17th International Conference on Science and Technology Indicators*, (Montréal) Vol. 2, 451–551.
- Li, X., Thelwall, M., and Giustini, D. (2011). Validating online reference managers for scholarly impact measurement. *Scientometrics* 91, 461–471. doi: 10.1007/s11192-011-0580-x
- Lin, J., and Fenner, M. (2013). Altmetrics in evolution: defining and redefining the ontology of article-level metrics. *Inf. Stand. Q.* 25, 20–26. doi: 10.3789/isqv25no2.2013.04
- Liu, X., Fang, H., and Wang, M. (2011). Correlation between download and citation and download-citation deviation phenomenon for some papers in Chinese medical journals. *Serials Rev.* 37, 157–161. doi: 10.1016/j.serrev.2011.02.001
- Mafahi, N., and Thelwall, M. (2016). When are readership counts as useful as citation counts? Scopus versus Mendeley for LIS journals. *J. Assoc. Inf. Sci. Technol.* 67, 191–199. doi: 10.1002/asi.23369
- Moed, H. F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. *J. Assoc. Inf. Sci. Technol.* 56, 1088–1097. doi: 10.1002/asi.20200
- Moed, H. F., Burger, W. J. M., Frankfort, J. G., and Van Raan, A. F. (1985). The use of bibliometric data for the measurement of university research performance. *Res. Policy* 14, 131–149. doi: 10.1016/0048-7333(85)90012-5

- Moed, H. F., and Plume, A. (2011). The multi-dimensional research assessment matrix. *Res. Trends* 23, 5–7. Available online at: <https://www.researchtrends.com/issue23-may-2011/the-multi-dimensional-research-assessment-matrix/>
- Mohammadi, E., and Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: research evaluation and knowledge flows. *J. Assoc. Inf. Sci. Technol.* 65, 1627–1638. doi: 10.1002/asi.23071
- Mohammadi, E., Thelwall, M., Haustein, S., and Larivière, V. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. *J. Assoc. Inf. Sci. Technol.* 66, 1832–1846. doi: 10.1002/asi.23286
- Murugesan, P., and Moravcsik, M. J. (1978). Variation of the nature of citation measures with journals and scientific specialties. *J. Am. Soc. Inform. Sci.* 29, 141–147.
- Na, J. C., and Ye, Y. E. (2017). Content analysis of scholarly discussions of psychological academic articles on Facebook. *Online Inf. Rev.* 41, 337–353. doi: 10.1108/OIR-02-2016-0058
- Neylon, C., and Wu, S. (2009). Level metrics and the evolution of scientific impact. *PLoS Biol.* 7:e1000242. doi: 10.1371/journal.pbio.1000242
- Nicholas, D., Huntington, P., and Jamali, H. R. (2008). User diversity: as demonstrated by deep log analysis. *Electron. Libr.* 26, 21–38. doi: 10.1108/02640470810851716
- Nielsen, F. Å. (2007). Scientific citations in Wikipedia. *arXiv:0705.2106* [preprint].
- Pew Research Center (2014). *Social Networking Fact Sheet*. Available online at: <http://www.pewinternet.org/fact-sheet/social-media/>
- Priem, J., Piwowar, H. A., and Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv:1203.4745* [preprint].
- Priem, J., Taraborelli, D., Groth, P., and Neylon, C. (2010). *Altmetrics: A Manifesto*. Available online at: <http://altmetrics.org/manifesto/>
- Schloegl, C., and Gorraiz, J. (2011). Global usage versus global citation metrics: The case of pharmacology journals. *J. Assoc. Inf. Sci. Technol.* 62, 161–170. doi: 10.1002/asi.21420
- Shapiro, F. R. (ed.). (2006). *The Yale Book of Quotations*. New Haven, CT: Yale University Press.
- Shema, H., Bar-Ilan, J., and Thelwall, M. (2012). Research blogs and the discussion of scholarly information. *PLoS ONE* 7:e35869. doi: 10.1371/journal.pone.0035869
- Small, H. (2016). "Referencing as cooperation or competition," in *Theories of Informetrics: A Festschrift in Honor of Blaise Cronin*, ed C. R. Sugimoto (Berlin; Boston, MA: Walter de Gruyter GmbH), 49–71.
- Sugimoto, C. (2015). "Attention is not Impact" and Other Challenges for Altmetrics. Wiley Exchange. Available online at: <https://hub.wiley.com/community/exchanges/discover/blog/2015/06/23/attention-is-not-impact-and-other-challenges-for-altmetrics>
- Sugimoto, C. R., Work, S., Larivière, V., and Haustein, S. (2017). Scholarly use of social media and altmetrics: a review of the literature. *J. Assoc. Inf. Sci. Technol.* 68, 2037–2062. doi: 10.1002/asi.23833
- Thelwall, M. (2008). Bibliometrics to webometrics. *J. Inf. Sci.* 34, 605–621. doi: 10.1177/0165551507087238
- Thelwall, M. (2012). Journal impact evaluation: a webometric perspective. *Scientometrics* 92, 429–441. doi: 10.1007/s11192-012-0669-x
- Thelwall, M., Haustein, S., Larivière, V., and Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS ONE* 8:e64841. doi: 10.1371/journal.pone.0064841
- Thelwall, M., Vaughan, L., and Björneborn, L. (2005). Webometrics. *ARIST* 39, 81–135. doi: 10.1002/aris.1440390110
- Thelwall, M., and Wilson, P. (2016). Mendeley readership altmetrics for medical articles: an analysis of 45 fields. *J. Assoc. Inf. Sci. Technol.* 67, 1962–1972. doi: 10.1002/asi.23501
- Tsou, A., Bowman, T. D., Ghazinejad, A., and Sugimoto, C. R. (2015). "Who Tweets About Science?," in *ISSI (Istanbul)*.
- Vainio, J., and Holmberg, K. (2017). Highly tweeted science articles: who tweets them? An analysis of Twitter user profile descriptions. *Scientometrics* 112, 345–366. doi: 10.1007/s11192-017-2368-0
- Waltman, L., and Costas, R. (2014). F1000 recommendations as a potential new data source for research evaluation: a comparison with citations. *J. Assoc. Inf. Sci. Technol.* 65, 433–445. doi: 10.1002/asi.23040
- Wan, J. K., Hua, P. H., Rousseau, R., and Sun, X. K. (2010). The journal download immediacy index (DII): experiences using a Chinese full-text database. *Scientometrics* 82, 555–566. doi: 10.1007/s11192-010-0171-2
- Xu, S., and Hemminger, B. M. (2015). What flavors are different types of scholarly articles? An investigation of PLOS publications. *Proc. Assoc. Inf. Sci. Technol.* 52, 1–4. doi: 10.1002/pr2.2015.1450520100113
- Xu, S., Yu, H., Hemminger, B. M., and Dong, X. (2018). Who, what, why? An exploration of JoVE scientific video publications in tweets. *Scientometrics* 1–12. doi: 10.1007/s11192-018-2880-x
- Yu, H., Xu, S., Xiao, T., Hemminger, B. M., and Yang, S. (2017). Global science discussed in local altmetrics: weibo and its comparison with Twitter. *J. Informetr.* 11, 466–482. doi: 10.1016/j.joi.2017.02.011
- Zahedi, Z., Costas, R., and Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of "alternative metrics" in scientific publications. *Scientometrics* 101, 1491–1513. doi: 10.1007/s11192-014-1264-0
- Ziman, J. (2000). *Beyond the Science Wars: The Missing Discourse About Science and Society*. Postacademic Science, 135.

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.