The relationship between tweets, citations, and article views for PLOS ONE articles

de Winter, JCF

DOI
10.1007/s11192-014-1445-x

Publication date
2015

Document Version
Final published version

Published in
Scientometrics: an international journal for all quantitative aspects of the science of science, communication in science and science policy

Citation (APA)
de Winter, JCF. (2015). The relationship between tweets, citations, and article views for PLOS ONE articles. *Scientometrics: an international journal for all quantitative aspects of the science of science, communication in science and science policy*, *102*(2), 1773-1779. https://doi.org/10.1007/s11192-014-1445-x

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Scientometrics (2015) 102:1773–1779
DOI 10.1007/s11192-014-1445-x

The relationship between tweets, citations, and article views for PLOS ONE articles

J. C. F. de Winter

Received: 4 June 2014 / Published online: 26 September 2014

Abstract An analysis of article-level metrics of 27,856 PLOS ONE articles reveals that the number of tweets was weakly associated with the number of citations ($\beta = 0.10$), and weakly negatively associated with citations when the number of article views was held constant ($\beta = -0.06$). The number of tweets was predictive of other social media activity ($\beta = 0.34$ for Mendeley and $\beta = 0.41$ for Facebook), but not of the number of article views on PubMed Central ($\beta = 0.01$). It is concluded that the scientific citation process acts relatively independently of the social dynamics on Twitter.

Keywords Altmetrics · PLOS ONE · Social media · Twitter

Some articles are highly tweeted, usually in the first few days after publication, while many others receive zero 'tweetations'. How is the number of tweets related to the number of citations an article receives?

Some scholars have expressed a positive opinion about the association between tweets and citations (Eysenbach 2011; Shuai et al. 2012). Eysenbach (2011), for example, found fairly strong correlations (0.40–0.70) between tweets and citations for 55 articles published in the Journal of Medical Internet Research, and concluded that "social media activity either increases citations or reflects the underlying qualities of the article that also predict citations, but the true use of these metrics is to measure the distinct concept of social impact." Others have expressed more reserved opinions. Thelwall et al. (2013) evaluated over 100,000 PubMed articles and concluded that tweets do predict citations, but also

Electronic supplementary material The online version of this article (doi:10.1007/s11192-014-1445-x) contains supplementary material, which is available to authorized users.

J. C. F. de Winter (✉)
Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
e-mail: j.c.f.dewinter@tudelft.nl

observed that the correlations were quite weak, and that tweets do not necessarily reflect academic impact (see also Costas et al. 2014; Haustein et al. 2014; Priem et al. 2012; Zahedi et al. 2014a). Yet others are outright negative. For example, Colquhoun and Plested (2014) and Beall (2013) argued that Twitter corrupts science.

There are several mechanisms through which a strong positive correlation between the number of tweets and the number of citations could occur. For example, it is possible that after a scientifically important article is published, this is noticed by scientists and/or the general public, who will then start tweeting. Obviously—through mechanisms unrelated to Twitter—high-impact articles will be well cited several years later, and a strong positive correlation arises between tweets and citations. It is also possible that the Twitter community reinforces citation rates by making an article known to a large scientific audience.

However, several mechanisms could cause the correlation between tweets and citations to be close to zero, or even negative. According to Haustein et al. (2014), it is possible that "the general public is not interested in the same topics as scientists cite but is interested in articles that have low scientific impact". Their qualitative analysis of the 15 most highly tweeted articles in the biomedical literature showed that these articles are either curious or funny, have potential health applications, refer to a catastrophe, or are topics that concern scientific publishing or scientists' behavior. In a related vein, it has been suggested that highly tweeted articles are about "catchy" topics (Van Noorden 2012), about "offbeat topics, current events, and general curiosities" (Liu 2014), and about "climate change, human health and diet, and online information and privacy" (Taylor and Plume 2014). Yet another possibility is that Twitter users hardly read the articles they tweet about. A recent revelation by Tony Haile, the chief executive officer of Chartbeat (a company that provides real-time analytics to websites and blogs) is consistent with this possibility: he stated that they "found effectively no correlation between social shares and people actually reading" (Haile 2014). Furthermore, Twitter could be used for "pathological" self-promotion (Buela-Casal 2014; Fenner 2014; Lin 2012), a phenomenon which could undermine the predictive validity of tweetations.

PLOS ONE is a large open access journal that facilitates various altmetrics. I focused on the 27,856 articles published between 1 July 2012 and 31 June 2013. Citations were counted using Crossref. The mean number of tweets per article was 2.66 (SD = 20.0). 59.6 % of the articles received zero tweets.

I entered the publication date and the number of tweets in a linear regression analysis in an attempt to predict the number of citations (variables in all analyses were rank transformed to account for non-normality). The obtained standardized β coefficients were -0.27 for the publication date (this number is negative since citations accumulate over time) and 0.10 for the number of tweets.

Next, I repeated the regression analysis, using the publication date, the number of tweets, as well as the number of article views as predictors. The obtained β coefficients were -0.19 for publication date, -0.06 for tweets, and 0.41 for article views on the PLOS ONE website. So, for a given number of article views, tweets *negatively* predict citations.

Figures 1, 2 and 3 provide further insight into the underlying mechanisms between tweeting, article views, and citations. Figure 1 shows that a clear association exists between the number of tweets and the number of PLOS ONE article views ($\beta = 0.38$ when controlling for publication date). PLOS ONE articles can be viewed not only on the PLOS ONE website, but also on PubMed Central (PMC), a website without altmetrics services. The association between tweets and article views on PMC is close to zero ($\beta = 0.01$). These results suggests that tweets attract article views on the PLOS ONE website because people click on the hyperlink embedded in the tweet, and that the number of tweets is

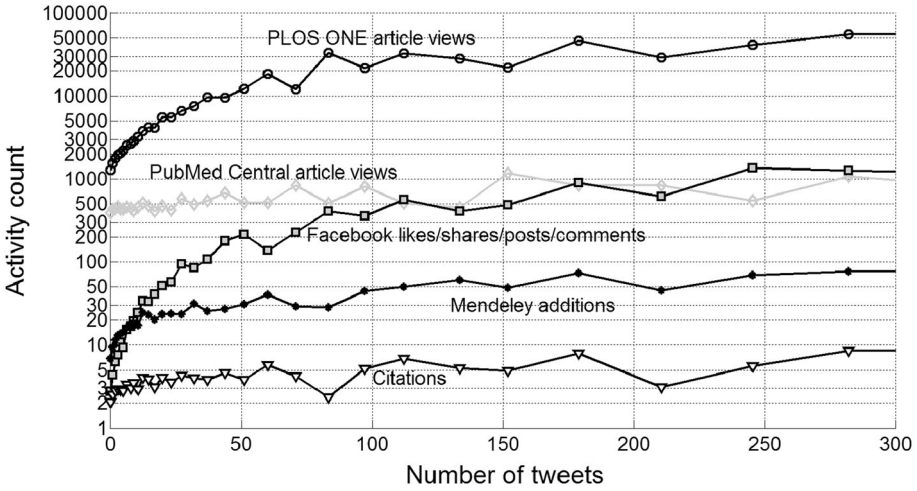

Fig. 1 Number of PLOS ONE article views, PubMed Central article views, Facebook likes/shares/posts/comments, Mendeley additions, and citations as a function of the number of tweets, for PLOS ONE articles published between 1 July 2012 and 31 June 2013. The number of tweets were counted in logarithmically spaced bins. Data were measured on 10 June 2014

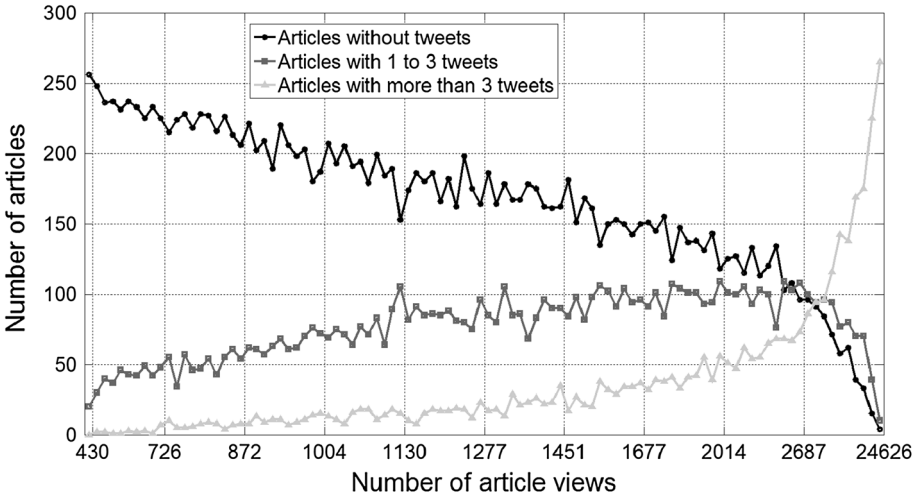

Fig. 2 Number of articles as a function of number of tweets and number of PLOS ONE article views, for PLOS ONE articles published between 1 July 2012 and 31 June 2013. Articles were divided into 100 groups based on the percentile rank of the number of views. Article views and tweets were measured on 10 June 2014

unrelated to the interests of non-Twitter (PMC) users. The number of tweets was associated with activity on other social media ($\beta = 0.34$ for the number of Mendeley posts and $\beta = 0.41$ for Facebook activity; Fig. 1). Figure 2 illustrates that articles without tweets are unlikely to have many article views. This suggests that Twitter is the primary medium for becoming a viewing hit. Figure 3 shows that highly-viewed articles are better cited when

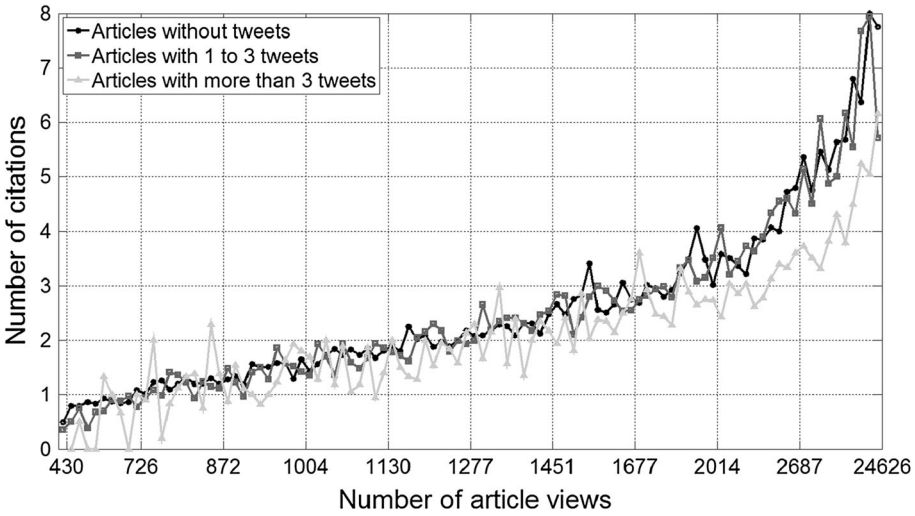

Fig. 3 Number of citations as a function of number of tweets and number of PLOS ONE article views, for PLOS ONE articles published between 1 July 2012 and 31 June 2013. Articles were divided into 100 groups based on the percentile rank of the number of views. Citations, article views, and tweets were measured on 10 June 2014

they are ignored on Twitter. This suggests that articles which have acquired their views from non-Twitter sources are more highly cited than articles which have acquired their views via Twitter.

The top 15 of the most highly tweeted PLOS ONE articles are listed in Table 1. Consistent with the above literature, highly-tweeted articles were about 'curious' topics or about human health and well-being. Specifically, several highly-tweeted articles were about psychological/neuroscience research related to gender, political orientation, and cognitive performance (articles 2, 4, 6, 11, 13, 14). These highly tweeted psychological articles often contained unique or spectacular results, but it is worth mentioning that null findings were highly tweeted too (article 4). Other popular articles were about health and weight loss (articles 7, 8, 9), consequences of the Fukushima accident (articles 3 and 15), the discovery of a new animal species (article 10), an analysis of Twitter data (article 1), and research about the media coverage of scientific findings and the success of research grants (articles 5 and 12).

In conclusion, yes: tweets do predict citations, but the effect is very weak. Article views obtained from Twitter hyperlinks have less academic substance than article views obtained from other referral sources. It seems that the scientific citation process acts relatively independently of the social dynamics on Twitter.

The virtual lack of correlation between tweets and citations does not imply that Twitter is inconsequential for scientists. It is of course possible that the number of tweets represents something else than academic impact, for example 'hidden impact' (i.e., academic impact that is not detected using citation counts), 'social impact', or relevance for practitioners (e.g., Darling et al. 2013; Taylor 2013). Furthermore, it is possible that tweets influence science in indirect ways, for example by steering the popularity of research topics, by faming and defaming individual scientists, or by facilitating open peer review (Liu 2014; Mandavilli 2011). The exact causal mechanisms between tweets and citations

Table 1 Statistics of the most highly tweeted PLOS ONE articles published between 1 July 2012 and 31 June 2013

Rank	Article title	Number of tweets	Number of article views	Number of citations
1	The Twitter of Babel: Mapping World Languages through Microblogging Platforms	1,870	12,846	4
2	The Power of Kawaii: Viewing Cute Images Promotes a Careful Behavior and Narrows Attentional Focus	1,163	152,460	5
3	Concentration of Strontium-90 at Selected Hot Spots in Japan	889	15,963	5
4	The Eyes Don't Have It: Lie Detection and Neuro-Linguistic Programming	792	71,663	2
5	Big Science vs. Little Science: How Scientific Impact Scales with Funding	754	47,567	5
6	Self-Affirmation Improves Problem-Solving under Stress	721	19,353	2
7	Minimal Intensity Physical Activity (Standing and Walking) of Longer Duration Improves Insulin Action and Plasma Lipids More than Shorter Periods of Moderate to Vigorous Exercise (Cycling) in Sedentary Subjects When Energy Expenditure Is Comparable	396	34,690	18
8	The Relationship of Sugar to Population-Level Diabetes Prevalence: An Econometric Analysis of Repeated Cross-Sectional Data	347	81,445	18
9	BCG Vaccination Reduces Risk of Tuberculosis Infection in Vaccinated Badgers and Unvaccinated Badger Cubs	344	9,241	6
10	Lesula: A New Species of Cercopithecus Monkey Endemic to the Democratic Republic of Congo and Implications for Conservation of Congo's Central Basin	331	131,750	9
11	Creativity in the Wild: Improving Creative Reasoning through Immersion in Natural Settings	331	33,282	8
12	Why Most Biomedical Findings Echoed by Newspapers Turn Out to be False: The Case of Attention Deficit Hyperactivity Disorder	318	22,731	5
13	Red Brain, Blue Brain: Evaluative Processes Differ in Democrats and Republicans	302	34,112	4
14	Why Don't Men Understand Women? Altered Neural Networks for Reading the Language of Male and Female Eyes	285	50,579	0
15	Distribution of Artificial Radionuclides in Abandoned Cattle in the Evacuation Zone of the Fukushima Daiichi Nuclear Power Plant	273	10,136	3
Mean, top 15 articles		607.73	48,521	6.27
Mean, all 27,856 articles		2.66	1,804	2.39

The number of article views represents the number of article views on the PLOS ONE website. The number of citations was measured with Crossref. Data were measured on 10 June 2014

require further study, and this may require longitudinal rather than cross-sectional research designs.

One limitation of the present analysis is that it focused on PLOS ONE only, a large open-access journal covering a variety of research topics. The obtained β coefficient between tweets and citations of 0.10 is in line with Haustein et al. (2014) who reported zero-order Spearman correlations between tweets and citations in the range of 0.10–0.20 for 134,929 biomedical articles that were mentioned on Twitter at least once. However, there exist large differences in Twitter use between scientific disciplines, with tweets being relatively uncommon in for example economics, sociology, and history of science (Holmberg and Thelwall 2014). Furthermore, some journals have an official tweeting account, which increases the probability that their articles receive tweetations (Haustein et al. 2014). The PLOS and PLOS ONE Twitter accounts have posted 6,398 and 7,594 tweets, respectively, as of 16 August 2014. Altmetrics services are not perfectly reliable. For example, it has been found that the PLOS application programming interface (API) does not recover the same tweets as Altmetric.com (Zahedi et al. 2014b). Another limitation is that the use of Twitter for the dissemination of scholarly information is still in its infancy. The social media landscape might change dramatically in the years to come (Van Noorden 2014).

The present analyses were based on citation data which can be downloaded from the PLOS website (http://www.plosone.org/static/almInfo). MATLAB code can be found in the Supplementary Material.

References

Beall, J. (2013). *Article-level metrics: An ill-conceived and meretricious idea*. Retrieved from http://scholarlyoa.com/2013/08/01/article-level-metrics.

Buela-Casal, G. (2014). Pathological publishing: A new psychological disorder with legal consequences? *The European Journal of Psychology Applied to Legal Context, 6*, 91–97. doi:10.1016/j.ejpal.2014.06.005.

Colquhoun, D., & Plested, A. (2014). *Why you should ignore altmetrics and other bibliometric nightmares*. Retrieved from: http://www.dcscience.net/?p=6369.

Costas, R., Zahedi, Z., & Wouters, P. (2014). *Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective*. http://arxiv.org/ftp/arxiv/papers/1401/1401.4321.pdf.

Darling, E. S., Shiffman, D., Côté, I. M., & Drew, J. A. (2013). *The role of Twitter in the life cycle of a scientific publication*. PeerJ PrePrints.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research, 13*. doi:10.2196/jmir.2012.

Fenner, M. (2014). Altmetrics and other novel measures for scientific impact. In *Opening science* (pp. 179–189). Berlin: Springer. doi:10.1007/978-3-319-00026-8_12.

Haile, T. [arctictony]. (2014, February 2). *We've found effectively no correlation between social shares and people actually reading [Tweet]*. Retrieved from https://twitter.com/arctictony/statuses/430008055263408128.

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

Holmberg, K., & Thelwall, M. (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*. doi:10.1007/s11192-014-1229-3.

Lin, J. (2012). *A case study in anti-gaming mechanisms for altmetrics*: PLOS ALMs and DataTrust. Retrieved from http://altmetrics.org/altmetrics12/lin.

Liu, J. (2014). *Interactions: January High Five (2014)*. Retrieved from http://www.altmetric.com/blog/interactions-january-high-five-2014.

Mandavilli A. (2011). Peer review: Trial by Twitter. *Nature 469*, 286–287. doi:10.1038/469286a.

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). *Altmetrics in the wild: Using social media to explore scholarly impact*. http://arxiv.org/html/1203.4745v1.

Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads, twitter mentions, and citations. *PLOS ONE*, *7*, e47523. doi:10.1371/journal.pone.0047523.

Taylor, M. (2013). Towards a common model of citation: Some thoughts on merging altmetrics and bibliometrics. *Research Trends*, *35*, 19–22. Retrieved from http://www.researchtrends.com/issue-35-december-2013/towards-a-common-model-of-citation-some-thoughts-on-merging-altmetrics-and-bibliometrics/.

Taylor, M., & Plume, A. (2014). Party papers or policy discussions: An examination of highly shared papers using altmetric data. *Research Trends, 36*, 17–20. Retrieved from http://www.researchtrends.com/issue-36-march-2014/highly-shared-papers/.

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLOS ONE*, *8,* e64841. doi:10.1371/journal.pone.0064841.

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature*, *512,* 126–129. doi:10.1038/512126a.

Van Noorden, R. (December 21, 2012). What were the top papers of 2012 on social media? [blog post]. *Nature News.* Retrieved from http://blogs.nature.com/news/2012/12/what-were-the-top-papers-of-2012-on-social-media.html.

Zahedi, Z., Costas, R., & Wouters, P. (2014a). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*. doi:10.1007/s11192-014-1264-0.

Zahedi, Z., Fenner, M., & Costas, R. (2014b). *How consistent are altmetrics providers? Study of 1000 PLOS ONE publications using the PLOS ALM, Mendeley and Altmetric.com APIs*. Retrieved from http://files.figshare.com/1515858/How_consistent_are_altmetrics_providers.doc.