Evaluating the influence of social media exposure of scholarly articles: Perspectives of social media engagement and click metrics

Cui, Y.; Wang, X.; Xu, S.; Hu, Z.; Zhang, C.

STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Evaluating the influence of social media exposure of scholarly articles: Perspectives of social media engagement and click metrics[1]

Yunxue Cui*, Xianwen Wang*, Shenmeng Xu**, Zhigang Hu* and Chunbo Zhang*

*zmxiaohai@mail.dlut.edu.cn, xianwenwang@dlut.edu.cn, huzhigang@dlut.edu.cn, zwchunbo@163.com
WISE Lab, Institute of Science of Science and S&T Management, Dalian University of Technology, Dalian, 116024 (China)

**xushenmeng@gmail.com
School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3360 (USA)

Introduction

With the development of digital library, internet and social media, the access of research results is going through significant changes. Different from the traditional way of obtaining journal articles by printed copies, people can get relevant news of the latest research results through the digit way and internet, and even express their opinions on the research results conveniently. Especially, scholarly articles are increasingly shared, commented, and discussed by researchers, members of the public, organizations and journals on social media platforms like Twitter, Facebook and YouTube. We call this phenomenon social media exposure of scholarly articles. In the social media era, journals need to redefine their role in the spreading of scientific results (Ortega, 2017). More and more journals use social media as a major new channel to disseminate new studies and interact with readers (Kelly, et al., 2016).

Despite the rapid development of social media and the emergence of many social networking platforms, Twitter and Facebook are the two major social media platforms, with the vast majority of users in the social media market. As of the fourth quarter of 2017, Facebook's monthly active users has reached 2.129 billion; at the same time, Twitter had 330 million active users. Wang et al. (2016) suggests that Twitter and Facebook are the two most important social referrals that direct people to scholarly articles, accounting for over 95% of the total social referral directed visits.

The growing scholarly content on Facebook and Twitter has aroused the interest of many researchers. Some studies focus on the relationship between altmetrics and citations (Priem, & Costello, 2010; Eysenbach, 2011; Thelwall, et al., 2013; Wang, et al., 2014; Costas, et al., 2015; Winter, 2015). However, few studies have delved into the role of social media in promoting the clicks of scholarly articles shared. So far, there has not been enough evidence to support whether social media exposure can increase article visits. Besides that, with Twitter being probably the most frequently studied platform in altmetrics, only a few studies have paid attention to the role of Facebook. Hence, in this study, we aim to take a first step in evaluating the effect of social media exposure of scholarly articles in promoting the visits. Moreover, we introduce a new data source, click metrics of short links exposure in social media posts, which

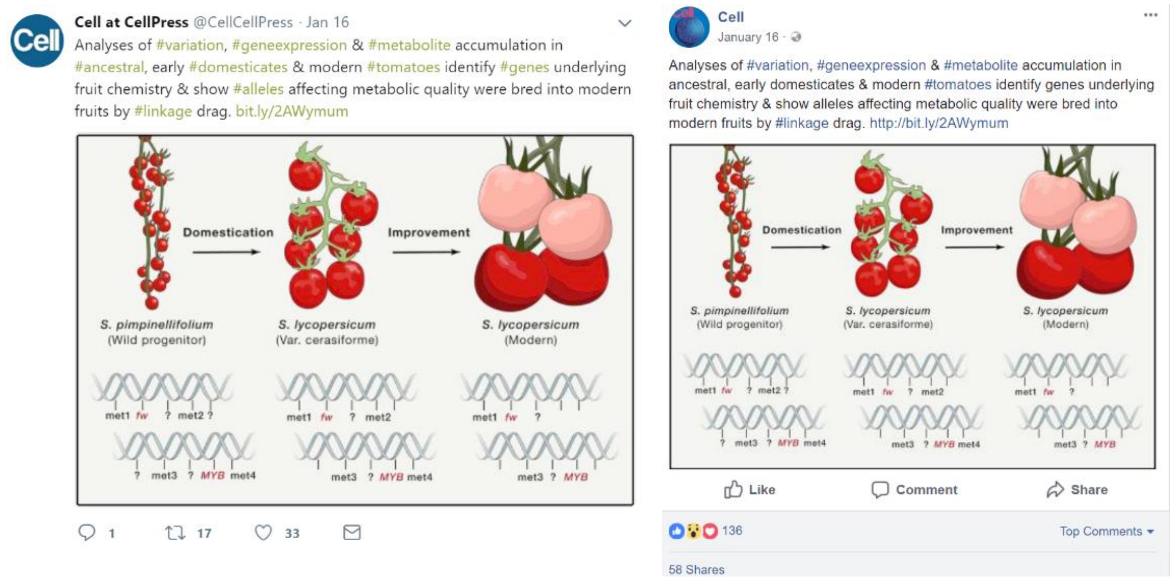
[1] This work was supported by the "Natural Science Foundation of China" (71503031, 71673038) and the "Fundamental Research Funds for the Central Universities" project (DUT18JC32).

make it possible to track and measure the effect of social media exposure directly. In this study, our research questions are as follows:
1. Does the sharing of the same papers on Twitter and Facebook have different effects on their clicks?
2. What is the relationship between the number of clicks of a paper and the social media activities around it?

Data and method

In this study, we collect data from three sources: Twitter, Facebook, and Bitly.com. Bitly is a link management platform that provides URL shortening service. When scholarly publishers' accounts post content on Twitter and Facebook, a significant proportion of those content has links to articles they publish. The links are usually shortened by URL shortening services like Bitly, ow.ly, goo.gl, and so on. Among these services, Bitly is one of the most used. More importantly, it is also a platform to open back-end data to ordinary users, which provides us with the possibility to conduct this research. Through the Bitly info plus page, which can be accessed by adding a "+" to any Bitlink, we can track the Bitlink's performance, including clicks, referrers, and locations of the clickers. In particular, it provides the information of clicks so we can know how many clicks come from Twitter, Facebook or any other referrals.
Since our data comes from three different platforms (Twitter, Facebook and Bitly), and the data open policies of the three platforms are different, the data availability for research is limited by these policies. Consequently, papers published in the journal *Cell* from October 2016 to May 2018 are studied. *Cell* is a prestigious journal in life science that has a large audience of readers and followers, both online and offline. *Cell* has official social media accounts on both Twitter and Facebook. In addition, *Cell* usually picks up some of their recently published articles and post them on Twitter and Facebook simultaneously. Almost every post attach a shortened URLs which link to the URL of the article, while Bitly is the most used shortening service by *Cell*. So, employing the data of *Cell* posts with the same Bitlinks on Facebook and Twitter, we can compare and analyze the different performances of the two platforms. We collected the posts of the papers published during the study period (October 2016 – May 2018) from the two official social media accounts, which are @CellCellPress on Twitter and @CellCellPress on Facebook. Specifically, we collect the social media engagement data which is generated by the interaction of users on social media. i.e., reactions, shares, and comments on Facebook; and likes, retweets, and comments on Twitter. We only include posts of papers that are shared on both Twitter and Facebook with the same short URLs (Bitlinks) attached. Finally, we collected a total of 324 posts of 162 papers with 162 short links. In the following step, we collected the clicks data from bitly.com for each short URL.

Figure 1: Social media exposure of the same Cell paper on Twitter and Facebook with a short link (left panel: Twitter; right panel: Facebook)



As shown in Figure 1, Cell posts exactly the same text and figure on Twitter and Facebook to introduce a recently published paper. In these posts, a short link is attached to direct visitors to the journal website. By adding a "+" to the Bitly short link in the posts, i.e., bitly.com/2AWymum+, we can get the click metrics data. It shows the engagement metrics data in the bottom of the posts. For Twitter, it provides the number of comments, retweets and likes of this post. Correspondingly, Facebook provides the number of comments, shares and reactions of this post. The reactions on Facebook are different with the likes on Twitter, it is a total that contains 6 kinds of emotions (likes, love, haha, wow, sad and angry).

For all the data, firstly we perform a descriptive statistical analysis to observe the altmetrics performance of content posted to the two social media platforms. Then we compare the data of the two platforms to find out the difference of social media engagement, i.e., reactions on Facebook versus likes on Twitter, and shares on Facebook versus retweets on Twitter. Finally, we use clicks data to measure the effectiveness of social media exposure on different platforms. In order to explore which altmetrics indicators are relevant to the effectiveness of the social media exposure, we conduct a regression analysis using the clicks data and social media engagement data on these two platforms. Considering the attributes of the data (count data that does not conform to a normal distribution and the mean and variance are unstable), we performed a log transformation on the data and found that the transformed data approximated a normal distribution. As suggested by Thelwall (Thelwall & Wilson, 2014), we finally employ an OLS model in the regression analysis using log-transforming data.

Results

Statistical analysis

Table 1 and Table 2 show the descriptive statistical results of 162 short links of 324posts' social media engagement performance on Facebook and Twitter.

Table 1. Social media engagement performance on Facebook

	Reactions	Shares	Comments	Clicks
Median	106	44	2	159
Max	793	398	36	2113
Min	13	3	0	18
Mode	113	50	0	156
Standard deviation	115	57	4	249

Table 2. Social media engagement performance on Twitter

	Likes	Retweets	Comments	Clicks
Median	26	17.5	0	74
Max	458	294	3	1251
Min	2	2	0	0
Mode	20	10	0	55
Standard deviation	47	30	0.58	132

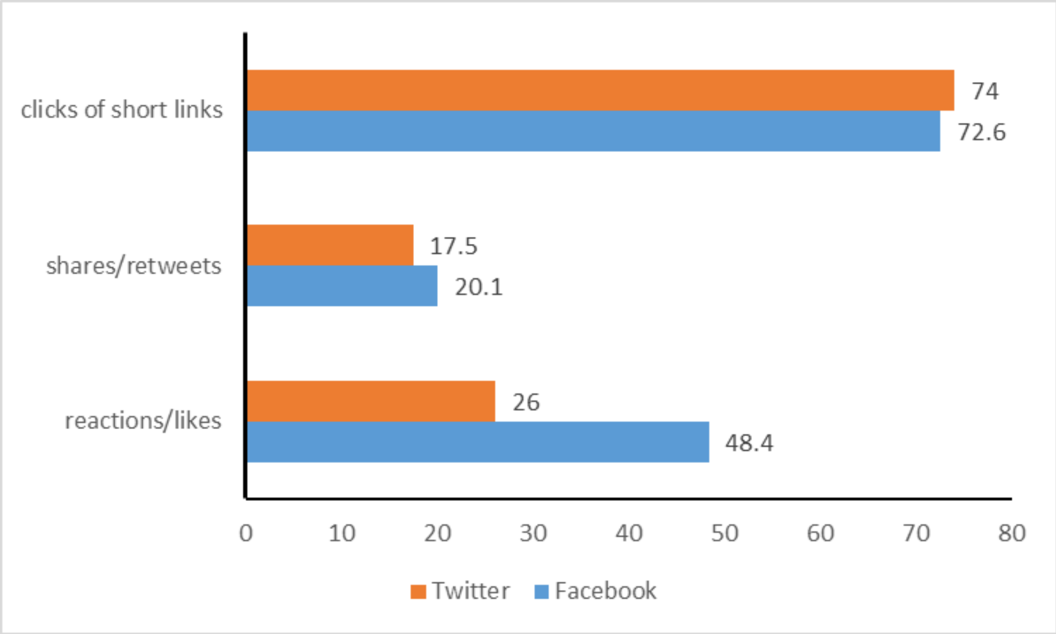
According to Table 1, each Facebook post gets 106 reactions, 44 shares, and 2 comments on median. The post with the best social media engagement performance has 793 reactions, 398 shares, and 36 comments; the post with the worst performance has only 13 reactions, 3 shares and no comments. Since the standard deviation of each group of data is relatively large, the average does not reflect the general level of the data well. We use the mode, the highest frequency in all data, to represent the general level of interest in the post. Most commonly, a post on Facebook can get about 113 reactions, 50 shares, but few comments. In addition, on median, the short links embedded in each post get 159 clicks. For the social media engagement performance on Twitter, each post gets 26 likes, 17.5 retweets, and zero comments on median. The post with the best performance gets 458 likes, 294 retweets, and 3 comments; the post with the worst performance hardly gets any social media attention. Just like the statistics of Facebook, the likes and retweets data have great variance. As the mode values indicate, most commonly, a post on Twitter can get about 20 likes, 10 retweets and 0 comments. Meantime, each short link in the post has 74 clicks on median.

Comparative analysis

The purpose of this analysis is to explore whether there are significant differences between the two platforms in terms of exposure effect. Although social media platforms have different content forms and user groups, their essential patterns are the same. That is, users generate content for sharing and discussion, and the content get widespread exposure and attention. Therefore, social media has many common natures and metrics that can be compared, especially on Facebook and Twitter. For example, Likes on Twitter is comparable with reactions on Facebook; so does the retweets on Twitter and shares on Facebook. Of course, Facebook's user volume and structure are different from Twitter. The intuitive difference in this study is that Cell’s Facebook account has many more followers than its Twitter account. As of August 8, 2018, Cell had 161,310 followers on Facebook, while on Twitter it was 73,590. That means the number of followers on Facebook is about 2.19 times that of Twitter. In order to control the impact of this difference on this study, we processed the data from Facebook, that is, the metrics of Facebook is divided by the weight of 2.19 and then compared with the metrics of Twitter.

Figure 2 shows the comparison of the median values from these two platforms. After dividing by the weight, all the engagement indicators of Facebook are close to Twitter except reactions: the number of reactions on Facebook is about twice as many as that of likes on Twitter. However, the likes on Twitter can only express positive emotions, while the reactions on Facebook contain 6 emotions, and two of them are negative emotions (sad, angry). If we take this into account and multiply the data of the reactions by 2/3, the number of reactions will be 31.9, just a little more than the likes on Twitter. Comments are not considered in the comparison because the median value of Twitter is too small; besides, the mode value indicates that there are few comments in general both on Facebook and Twitter.

Figure 2: Comparison of Twitter and Facebook



The result of the comparison suggests that if we do not consider the impact of platform differences, posts on Facebook attract greater user engagement than tweets. In addition, short links from posts on Facebook receive much more clicks than Twitter. But in fact, the social media engagement metrics of the two platforms are almost the same when we eliminated the difference in platform user volume and function.

Regression analysis

Clicks of the short links attached in the posts are considered as the direct indicator to measure the performance of social media exposure. Therefore, in the regression analysis, clicks are considered as the response variable. The social media engagement indicators, including likes/moods, retweets/shares, are regarded as predictor variables. Considering that both Facebook and Twitter's comments data are scarce, comments are not taken into consideration in our analysis. The regression model is as follows:

$$\log(Y) = \beta_0 + \sum_{j=1}^{p} \beta_j \log(X_j)$$

where Y is the response variable, X_j is the predictor variables, and β_j is the regression coefficients.

Because all the data are approximately subject to a normal distribution after log transformation, we choose the OLS regression model to conduct the regression analysis. Moreover, we conduct two regression analyses separately for Facebook and Twitter. The results are shown in Table 3 and Table 4.

Table 3. Regression analysis of Facebook

	Coefficients	**Exp**	**Std. Error**	**t value**	**P (>|t|)**
Intercept	1.482***	4.402	0.296	5.000	1.5e-06
Log(Reactions)	0.498***	1.646	0.138	3.613	0.00049
Log(Shares)	0.346***	1.414	0.115	3.024	0.003

** p<0.05, ***p<0.01

Table 4. Regression analysis of Twitter

	Coefficients	Exp	Std. Error	t value	P (>|t|)
Intercept	2.156***	8.634	0.172	12.538	<2e-16
Log(Likes)	0.265**	1.303	0.108	2.450	0.015
Log(Retweets)	0.448***	1.566	0.109	4.118	6.17e-05

** $p<0.05$, ***$p<0.01$

sAs shown in Table 3, the coefficients of Reactions and Shares for Facebook are 0.498 and 0.346. The corresponding exponential value is 1.646 and 1.414. In other words, if the reactions/shares increase by one unit, the difference in the number of clicks would be expected to increase by 1.646/1.414 unit, while holding the other variables in the model constant.
The results of Twitter are similar. The corresponding exponential coefficient of likes and retweets are 1.303 and 1.566, which means that if the likes/retweets increase by one unit, the difference in the number of clicks would be expected to increase by 1.303/1.566 unit, while holding the other variables in the model constant.
The regression results indicate the similar characteristics of the two platforms. Both reactions/shares on Facebook and likes/retweets on Twitter have significant effects on directing people to visit Cell publications. Additionally, comparing these two platforms, retweets are more efficient in directing visitors to Cell article through clicks than likes on Twitter while that is just the opposite on Facebook.

Conclusions and discussion

Social media exposure of scholarly content by journals have positive effects on promoting the clicks of the publications. The effects manifest in the social media engagement and would transform to real clicks of articles. In absolute terms, for the Cell journal, the effects of social media exposure on Twitter and Facebook are different. Social media exposure on Facebook has greater influence than that on Twitter. The social media engagement of scholarly content on Facebook is 2.5~4 times as much as on Twitter. The number of clicks of short links in Facebook posts are twice as many as that in tweets, which means that posts on Facebook direct more readers to the papers. However, the true exposure effect of the two platforms is similar when the total followers difference between the two platforms are eliminated. The regression analysis shows that social media engagement is related to the clicks of articles. In other words, when the level of social media engagement is high around a paper, the number of clicks of that paper also tends to be high. For Cell journal, whether reactions and shares on Facebook or likes and retweets on Twitter, it has positive effects on the clicks of short links. Based on the sample that we investigated, our study confirms the importance of social media exposure, especially exposure on Facebook, which is as an important disseminating channel for Cell publications as Twitter. Even in absolute terms, Facebook has brought far more paper clicks than Twitter, benefiting from its huge number of users. Although sometimes altmetrics studies paid more attention to Twitter, the role of Facebook in altmetrics cannot be ignored.
There are also some limitations in this research. We only use one journal as the research object. Research results may be biased when extended to multiple journals. In addition, clicks and social media performance are still affected by many other factors (i.e. the temporality of posts, paid vs unpaid posts, bot traffic on Facebook vs Twitter, Facebook may have less scholars, but more people on the platform than Twitter), and this article is limited by research conditions that cannot control these factors. Subsequent in-depth research should expand the scope of the journal and take other variables into account. Our research is based solely on a rough estimate and analysis of altmetrics data on social media, but given the impact of Cell journals, we believe that the results of this case study are still somewhat informative. This study provides a case for altmetrics researchers and academic journals. We notice that many journals prefer to post their

newly published papers on Twitter, rather than Facebook. Moreover, many journals only have official social media accounts on Twitter. Given the results of this study, Facebook can also be an effective promotion channel for scholarly journals and their publications.

References

Costas, R., Zahedi, Z., & Wouters, P. (2015). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003-2019.

Desai T, Shariff A, Shariff A, Kats M, Fang X, et al. (2012) Tweeting the Meeting: An In-Depth Analysis of Twitter Activity at Kidney Week 2011. *PLoS ONE*, 7: e40253.

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(4).

Kelly, B. S., Redmond, C. E., Nason, G. J., Healy, G. M., Horgan, N. A., & Heffernan, E. J. (2016). The use of Twitter by radiology journals: an analysis of Twitter activity and impact factor. *Journal of the American College of Radiology*, *13*(11), 1391-1396.

Ortega, J. L. (2017). The presence of academic journals on Twitter and its relationship with dissemination (tweets) and research impact (citations). *Aslib Journal of Information Management*, *69*(6), 674-687.

Priem, J., & Costello, K. L. (2010b). How and why scholars cite on twitter. Proceedings of the Association for Information Science & Technology, 47(1), 1-4.

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? twitter and ten other social web services. Plos One, 8(5), e64841.

Thelwall, M., & Wilson, P. (2014). Regression for citation data: an evaluation of different methods. Journal of Informetrics, 8(4), 963-971.

Wang, X., Liu, C., Fang, Z., & Mao, W. (2014). From Attention to Citation, What and How Does Altmetrics Work?. *arXiv preprint arXiv:1409.4269*.

Wang, X. et al. (2016). Tracking the digital footprints to scholarly articles from social media. Scientometrics, 109(2), 1365-1376.

Winter, J. C. F. D. (2015). The relationship between tweets, citations, and article views for plos one articles. Scientometrics, 102(2), 1773-1779.