Towards advanced social media metrics: understanding the diversity and characteristics of Twitter interactions around science

Fang, Z.

CHAPTER 4

User engagement with scholarly tweets of scientific papers: A large-scale and cross-disciplinary analysis[1]

[1] **This chapter is based on**:

Fang, Z., Costas, R., & Wouters, P. User engagement with scholarly tweets of scientific papers: A large-scale and cross-disciplinary analysis. (Under review).

Author contributions:

Fang, Z. (Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data Curation, Writing - Original Draft, Writing - Review & Editing)

Costas, R. (Conceptualization, Methodology, Investigation, Supervision, Writing - Review & Editing)

Wouters, P. (Conceptualization, Supervision, Writing - Review & Editing)

Abstract

This study investigates the extent to which scholarly tweets of scientific papers are engaged with by Twitter users through four types of user engagement behaviors, i.e., liking, retweeting, quoting, and replying. Based on a sample consisting of 7 million scholarly tweets of Web of Science papers, our results show that *likes* is the most prevalent engagement metric, covering 44% of scholarly tweets, followed by *retweets* (36%), whereas *quotes* and *replies* are only present for 9% and 7% of all scholarly tweets, respectively. From a disciplinary point of view, scholarly tweets in the field of *Social Sciences and Humanities* are more likely to trigger user engagement over other subject fields. The presence of user engagement is more associated with other Twitter-based factors (e.g., number of mentioned users in tweets and number of followers of users) than with science-based factors (e.g., citations and Mendeley readers of tweeted papers). Building on these findings, this study sheds light on the possibility to apply user engagement metrics in measuring deeper levels of Twitter *reception* of scientific information.

Keywords

Altmetrics, social media metrics, Twitter engagement, scholarly communication, retweet

4.1 Introduction

With the growing enthusiasm for sharing scientific information via Twitter, tweets mentioning scientific papers (hereinafter "scholarly tweets") are widespread. In the altmetrics realm, Twitter has arguably become one of the most crucial data sources, with more than one third of recent scientific papers being tweeted (Fang, Costas, et al., 2020). Scholarly tweets, therefore, have long been seen as measureable traces possibly capturing the impact of research outputs in a broader sense (Bornmann & Haunschild, 2016; Eysenbach, 2011).

Instead of merely serving as countable information carriers bringing scientific papers to the attention of Twitter users, scholarly tweets per se are also informative in terms of the content incorporated, the characteristics of users involved, as well as the possible user engagement triggered, collectively making Twitter a valuable source of social media metrics. In other words, the creation of scholarly tweets stands not only for an outcome of Twitter reception of science by users who posted them, but also a prologue of another narrative about how other users might interact with them in the Twitter universe, being relevant to quantitative elaboration of science-social media interactions (Costas et al., 2021).

4.1.1 Scholarly tweets as the objects of study

Díaz-Faes et al. (2019) proposed the umbrella term *secondary social media metrics* to conceptualize metrics taking "social media objects" (i.e., social media users and their online activities) as the objects of study, distinguishing them from *primary social media metrics* which focus on "research objects" (e.g., publications, datasets, journals, and individual scholars), in particular "the use and visibility of publications on social media". To date, in the direction of *secondary social media metrics*, many research efforts centering on scholarly tweets have been made to characterize the mechanisms of how Twitter users process, circulate, and engage with scientific information from different perspectives.

At the tweet level, content analyses provide straightforward insights into the tweeting behavior of users who are disseminating scientific information. For example, by scrutinizing the content of scholarly tweets received by the top ten most tweeted papers in the field of dentistry, Robinson-Garcia et al. (2017) exemplified the scarce existence of original thought but more mechanical nature of the bulk of tweet content. Similarly, with a case study containing 270 tweets, Thelwall, Tsou, et al. (2013) reported that the majority of the observed scholarly tweets only echoed a paper title or presented a brief summary. Regarding the sentiment of tweet texts, scholarly tweets were found to be generally neutral, with limited share showing positive or negative sentiment expressed by users (Friedrich et al., 2015; Thelwall, Tsou, et al., 2013). Besides, the use of some tweet features in scholarly tweets, such as hashtags (word or phrase prefixed with #) and user mentions (user's handle name prefixed with @), was also of interest by some altmetric research (Haustein, Bowman, Holmberg, Peters, et al., 2014; S. Xu et al., 2018), because it represents a particular form of user interactions enhancing the description and visibility of tweets and facilitating connections amongst users (Haustein, 2019; Holmberg et al., 2014).

At the user level, the presence of scholarly tweets makes it possible to recognize and characterize users discussing science on Twitter. Scholarly tweets, therefore, were drawn upon for identifying and classifying Twitter users participating in scholarly communication (Costas et al., 2020; Díaz-Faes et al., 2019; Vainio & Holmberg, 2017; Yu et al., 2019), and for further exploring how users by type performed differently while utilizing Twitter for scholarly communication (Didegah et al., 2018; Holmberg & Thelwall, 2014; Mohammadi et al., 2018; Yu, 2017). Moreover, the aforementioned objects derived from scholarly tweets, either at the tweet or user level, were not only studied separately, but sometimes networked in different ways to map the contexts in which Twitter interactions with science happened. The network methods include but not limited to co-occurrence of hashtags (Haunschild et al., 2019), co-occurrence of users and hashtags (Hellsten & Leydesdorff, 2020), user mentions network (Said et al., 2019), and follower/friend network of users (Alperin et al., 2019; Robinson-Garcia et al., 2018), which were collectively conceptualized as *heterogeneous couplings* by Costas et al. (2021).

4.1.2 User engagement behaviors around scholarly tweets

In addition to tweet content and user characteristics, user engagement behavior[1] around scholarly tweets is also a focal point of *secondary social media metrics*. Conceptually speaking, scholarly tweets offer the possibility for a wider range of users to participate in science-focused discussions through many engagement behaviors enabled by Twitter. In the current platform version (2021), Twitter provides several engagement functionalities for users to interact with tweets on their own initiative. As illustrated with a tweet example in Figure 1, there are four main types of engagement functionalities with corresponding metrics visible at the bottom of tweets and publicly retrievable through the Twitter API, including (1) *like*, (2) *retweet*, (3) *quote tweet*, and (4) *reply*. These engagement behaviors differ in both input and output. In terms of input, *liking* and *retweeting* are relatively basic and simple engagement behaviors because they are both devoid of extra original content added, whereas *quoting* and *replying* are comparatively more informative and conversational because they enable users to express original thought and content. As to output, except liking, the other

[1] We referred to the definition of "engagements" metrics interpreted by Twitter as "total number of times a user interacted with a tweet. Clicks anywhere on the tweet, including retweets, replies, follows, likes, links, cards, hashtags, embedded media, username, profile photo, or tweet expansion" (https://help.twitter.com/en/managing-your-account/using-the-tweet-activity-dashboard) (Accessed April 28, 2021). Therefore, in this study user engagement behavior refers to any interaction behavior performed by Twitter users on existing tweets.

three types of engagement behaviors can produce new tweets (i.e., retweets, quote tweets, and replies) which are accounted for users' total number of tweets posted.

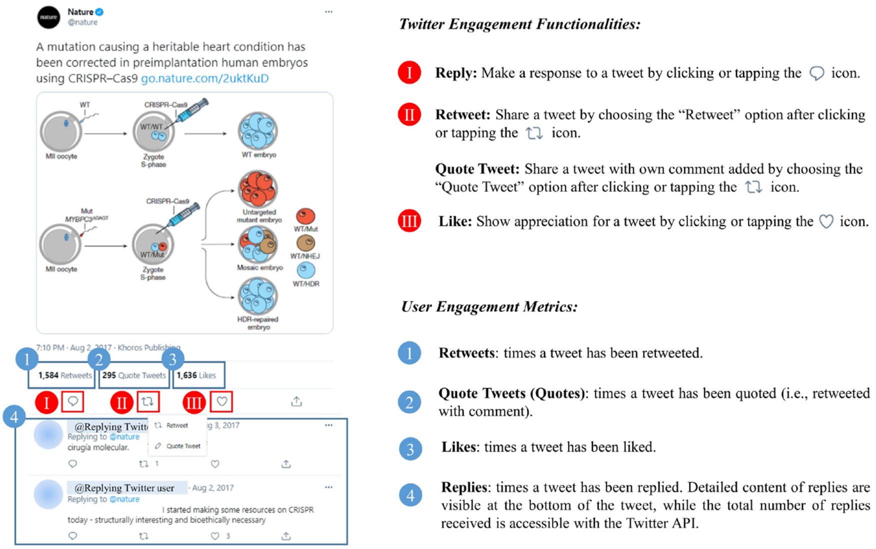

Figure 1. User engagement functionalities and metrics embedded in a tweet example

In existing altmetric literature, retweeting is the most analyzed user engagement behavior. In 2010, Priem & Costello (2010) found that retweets, as the outcomes of retweeting behavior, only made up 19% of a sample of scholarly tweets posted by 28 academic users. However, in more recent research, it was generally reported that retweets accounted for close to or over half of scholarly tweets of papers (Alperin et al., 2019; Didegah et al., 2018; Haustein, 2019), being a key component of the data base of studies related to scholarly Twitter metrics. For a sample of tweets posted by the Twitter accounts of 25 U.S. health agencies, Bhattacharya et al. (2014) found that about one third of them had zero retweet while the rest were retweeted at least once. As a form of information diffusion in nature, retweets were often analyzed to help capture topics of the public's interest in sharing (Bhattacharya et al., 2014; Kahle et al., 2016), or to construct Twitter dissemination networks of scientific knowledge across communities (Araujo, 2020; Hassan et al., 2019).

Besides retweeting, other types of user engagement behaviors, such as liking, replying, and clicking, were also studied to help yield insights into whether and how the public engages with scientific information on Twitter. For instance, considering a spectrum of user engagement metrics (e.g., retweets, likes, replies, clicks on tweeted URLs), Kahle et al. (2016)

studied the rates of user engagement with the tweets posted by the official Twitter accounts of the European Organization for Nuclear Research (CERN). Mohammadi et al. (2018) surveyed the motivations behind users' liking and retweeting behaviors in scientific contexts and reported that most survey respondents liked a tweet to "inform the authors that their tweets were interesting" and retweeted to disseminate the tweets. Based on the click metrics data provided by Bitly for its generated short links, Fang et al. (2021) analyzed the click rates on Twitter of short links referring to scientific papers and concluded that nearly half of the studied short links were not clicked by Twitter users at all.

4.1.3 Objectives

Although some attempts so far have been made to enhance the understanding of how people react to scholarly tweets, existing literature generally focused on either a specific user engagement behavior or a specific discipline. Little is known about the overall picture of the coverage of diverse types of user engagement with science on Twitter. Against this background, on the basis of a large-scale and cross-disciplinary dataset, the main research objective of this study is to systematically unravel the extent to which scholarly tweets are related to different categories of user engagement. Specifically, this study sets out to address the following research questions:

RQ1. To what extent are scholarly tweets engaged with by Twitter users through different types of engagement behaviors (i.e., liking, retweeting, quoting, and replying)?

RQ2. Which subject fields of science have scholarly tweets attracting higher levels of user engagement on Twitter?

RQ3. How does the presence of user engagement correlate with other science-based and Twitter-based factors of scholarly tweets (e.g., scholarly impact of tweeted papers, use of tweet features, user characteristics)?

4.2 Data and methods

4.2.1 Dataset

We retrieved a total of 6,229,001 Web of Science-indexed (WoS) papers published between 2016 and 2018 from the CWTS in-house database, and searched their scholarly tweets recorded by Altmetric.com until October 2019. For the matching with Altmetric.com data, WoS papers are restricted to those with DOI or PubMed ID assigned. On the whole, there are 1,999,199 WoS papers (accounting for 32.1%) with at least one scholarly tweet received, totally generating 7,037,233 unique original scholarly tweets.[2] Note that to explore user engagement behaviors, in this study the analyzed scholarly tweets are limited to original tweets which can be engaged with through the engagement functionalities provided by Twitter.

For the approximately 7 million scholarly tweets in our dataset, we retrieved their engagement metrics (i.e., number of likes, retweets, quotes, and replies received) with the Twitter API in February 2021.

4.2.2 The CWTS publication-level classification

To compare the user engagement situations of scholarly tweets across subject fields of science, we applied the CWTS publication-level classification system (Waltman & Van Eck, 2012) to assign scholarly tweets with subject field information based on their mentioned scientific papers. The CWTS classification clusters WoS papers into micro-level fields based on their citation relationships. These micro-level fields are then algorithmically assigned to five main subject fields, including *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PSE), *Life and Earth Sciences* (LES), and *Mathematics and Computer Science* (MCS).[3] For our dataset, there are a total of 5,932,279 scholarly tweets (accounting for 84.3%) referring to scientific papers with the subject field information assigned by the CWTS classification system. This set of scholarly tweets was drawn as a subsample for studying the subject field differences of user engagement. Table 1 presents the distribution of the analyzed scientific papers and scholarly tweets across the five subject fields of science.[4]

Table 1. Five subject fields of the CWTS publication-level classification system

Subject field	Abbreviation	Number of papers	Number of tweets
Social Sciences and Humanities	SSH	188,142	671,490
Biomedical and Health Sciences	BHS	968,605	3,544,755
Physical Sciences and Engineering	PSE	324,559	676,269
Life and Earth Sciences	LES	288,563	881,941
Mathematics and Computer Science	MCS	58,279	159,680

[2] We collected detailed Twitter information (e.g., tweet content and user demographics) in December 2019 for the tweet IDs provided by Altmetric.com (version: October 2019). Unavailable scholarly tweets caused by deletion of tweets, or suspension and protection of Twitter users' accounts (Fang, Dudek, et al., 2020) were not included in our dataset.

[3] See more introduction to the CWTS classification system (also known as the Leiden Ranking classification) at: https://www.leidenranking.com/information/fields (Accessed April 28, 2021).

[4] Full counting was applied for scholarly tweets that cite multiple papers belonging to different subject fields.

4.2.3 Science-based and Twitter-based factors of scholarly tweets

To explore how the presence of user engagement behaviors associates with scholarly, tweet, and user-related factors of the scholarly tweets, we extracted a total of ten factors from the following three dimensions: (1) *scholarly impact* of tweeted papers, (2) use of *tweet features*, and (3) *user characteristics* of those who posted scholarly tweets.

As listed in Table 2, in the dimension of scholarly impact of tweeted papers, we selected WoS citations and Mendeley readers to reflect the impact of tweeted papers in the science environment. In the dimension of tweet features, we focused on the use of hashtags and user mentions. Number of hashtags used, and number of users mentioned in tweets were analyzed to reflect how users edit their scholarly tweets with such interactive tweet features. Last but not least, in the dimension of user characteristics, we studied six factors related to users' demographics and behaviors on Twitter. Thus, number of followers and number of lists in which users are listed represent social media capital held by users, because these two factors largely affect how broad the audiences can be reached for posted tweets. Number of friends and number of likes given tell the story of how active users interact with other users by following others or liking their tweets. Number of tweets posted by users and their science focus depict users' overall tweeting behavior. The former indicates how frequent users post all kinds of tweets, while the latter implies how concentrated users are on tweeting scientific papers.

Table 2. Analyzed factors related to scholarly tweets

Dimension	Factor	Description
Scholarly impact of tweeted papers	Citations	Total number of WoS citations received by the papers mentioned in a tweet. Citation counts were retrieved from the CWTS in-house WoS database (version: March, 2020).
	Readers	Total number of Mendeley readers received by the papers mentioned in a tweet. Mendeley readership data were collected with the Mendeley API in July, 2020.
Tweet features	Hashtags	Number of hashtags used in a tweet.
	Mentioned users	Number of Twitter users mentioned in a tweet.
User characteristics	Followers	Number of Twitter users following a user.
	Lists listed	Number of lists in which a user is listed.
	Friends	Number of Twitter users followed by a user.
	Likes given	Number of likes given by a user since the account was created.
	Tweets posted	Number of tweets posted by a user since the account was created.
	Science focus	Proportion of scholarly tweets (recorded by Altmetric.com) among all tweets posted by a user. This indicator is equivalent to "ptws to papers" in Díaz-Faes et al. (2019). The higher the value of science focus of a user, the more concentrated the user is on tweeting scientific papers.

In the Results section, the correlations between the four analyzed user engagement metrics and the ten factors were studied based on not only the Spearman correlation analysis of counts (performed by R), but also the visualized change trend of the coverage of user engagement among scholarly tweets aggregated at the different levels of each studied factor (coverage of user engagement refers to proportion of scholarly tweets with at least one specific user engagement received).

4.2.4 Regression analysis: hurdle model

To further investigate how different factors can predict user engagement behaviors, we conducted regression analysis for each of the four user engagement metrics as a response variable. As presented later in the Results section, in consideration of that all of the four user engagement metrics are count data and the data are over-dispersed (data with the variance much greater than the mean value) and zero-inflated (data with excess zero values), we adopted hurdle models (Mullahy, 1986) as our regression models. Given that social media engagement data were generally found to be count data with the characteristics of zero-inflation and over-dispersion, hurdle models have been applied by many previous research to model user engagement on different social media platforms like Twitter (Bhattacharya et al., 2014), Facebook (Bhattacharya et al., 2017; Bohn et al., 2014), and Weibo (Fu & Chau, 2013).

Hurdle model is a two-part regression model that processes the count data as two separate components: one is the *zero portion* modeling whether an observation takes zero value or non-zero value (typically a binary logit model), while the other is the *count portion* fitting those non-zero values (a zero-truncated negative binomial model used in this study in consideration of the over-dispersion of the count data). In our case, the zero portion in the hurdle models determines whether a scholarly tweet gets at least one specific user engagement or not, while the count portion models how many times it is engaged with through certain behavior. Therefore, the hurdle models of user engagement metrics allow for the simultaneous modeling of both the likelihood for scholarly tweets of being engaged with, as well as the frequency of being engaged with by users. We employed the *pscl* package in R (Zeileis et al., 2008) to construct four hurdle models (mode 1: likes; model 2: retweets; model 3: quotes; model 4: replies).

4.3 Results

The Results section consists of four parts. The first part exhibits the overall presence of the four types of user engagement (i.e., likes, retweets, quotes, and replies) among the 7 million scholarly tweets in our dataset. The second part compares the presence of user engagement across scholarly tweets in different subject fields. The third part investigates how the presence of user engagement relates to different factors with respect to scholarly impact of tweeted papers, use of tweet features, and characteristics of users. The last part focuses on the hurdle regression of user engagement metrics.

4.3.1 Overall user engagement with scholarly tweets

Figure 2 illustrates the coverage of the four types of user engagement among the 7 million scholarly tweets. About 52% have been engaged with through at least one of the four analyzed engagement behaviors, namely, the overall coverage of user engagement among scholarly tweets is 52%. Around 20% of scholarly tweets were engaged with by users through only one type of engagement behavior, while as low as 2% got all the four types of engagement. More specifically, the coverage of likes is 44%, followed by retweets 36%. Liking and retweeting appear to be the most prevalent engagement behaviors around scientific information. In contrast, the coverage of both quotes and replies is relatively scarce. Only 9% of scholarly tweets got quoted by users, and as low as 7% received at least one reply.

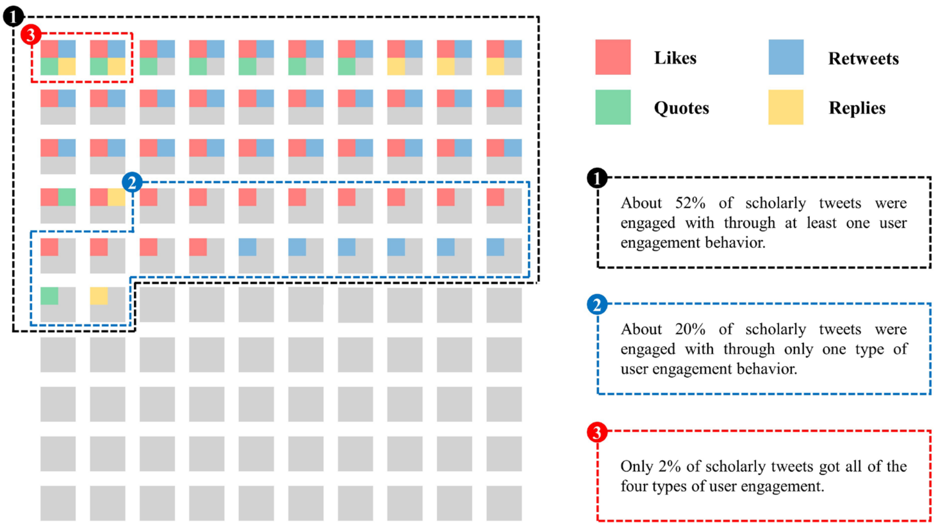

Figure 2. Coverage of the four types of user engagement. Each square represents 1% of scholarly tweets in our dataset. A square tinted with specific color(s) indicates that its represented 1% of scholarly tweets got corresponding type(s) of user engagement

Table 3 presents the descriptive statistics of the four engagement metrics to further reflect the extent to which scholarly tweets are engaged with. The coefficient of skewness and quartiles indicate that the distribution of all of the four types of engagement metrics is highly skewed. Only a few scholarly tweets got considerable user engagement, while the majority were never or rarely engaged with by Twitter users. Liking is the most widespread engagement behavior, contributing the most to user engagement metrics, followed by retweeting. On average, scholarly tweets in the dataset have been liked 2.95 times and retweeted 1.91 times. However, quoting and replying are more rare engagement behaviors, with only 1% of scholarly tweets being quoted for at least 3 times (99^{th} percentile of quotes is 3) or replied for at least 2 times (99^{th} percentile of replies is 2), suggesting that only a very limited share of scholarly tweets successfully aroused users' interest in sharing and communicating their thought within Twitter conversations.

Table 3. Descriptive statistics of the four user engagement metrics

Metrics	Sum	Mean	Min	Q1	Q2	Q3	90^{th} P	99^{th} P	Max	Skewness	SD
Likes	20,755,430	2.95	0	0	0	2	6	39	10,561	156.11	21.17
Retweets	13,429,713	1.91	0	0	0	1	4	26	9,983	218.44	16.89
Quotes	1,179,934	0.17	0	0	0	0	0	3	804	155.91	1.53
Replies	821,176	0.12	0	0	0	0	0	2	1,033	285.41	1.04

Note: Sum = total number of corresponding engagement metrics; Q1, Q2, Q3 = the first, second, and third quartile; 90th P = the 90th percentile; 99th P = the 99th percentile; Min, Max = the minimum and maximum value; Skewness = the coefficient of skewness; SD = standard deviation.

4.3.2 User engagement across subject fields

Figure 3 shows how the coverage of the four types of user engagement varies across the five subject fields of science: *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PSE), *Life and Earth Sciences* (LES), and *Mathematics and Computer Science* (MCS). Overall, scholarly tweets mentioning SSH papers are more likely to be engaged with through any type of engagement behavior studied. For the field of SSH, the proportion of scholarly tweets with at least one engagement record always ranks first, suggesting the higher probability for SSH-related scientific information to be engaged with by Twitter users over other subject fields. Besides, scholarly tweets from the fields of LES and BHS also present a relatively stronger potential in attracting different types of user engagement. In comparison, MCS and PSE are the two subject fields with sparser user engagement behaviors observed, showing the lowest coverage of all kinds of studied engagement metrics.

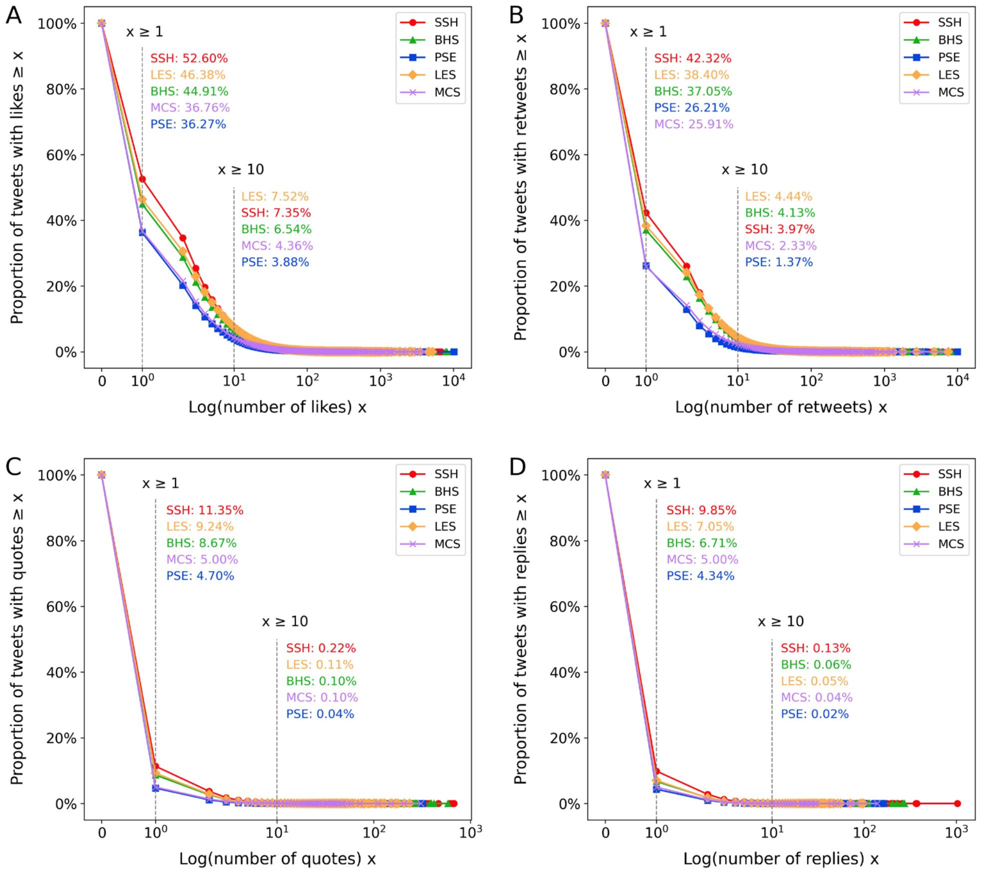

Figure 3. Proportion of scholarly tweets with different levels of **A** likes, **B** retweets, **C** quotes, and **D** replies across the five subject fields of science

The descriptive statistics of user engagement metrics across subject fields presented in Table 4 reinforces the disciplinary differences observed in Figure 3. Considering the greatest values of indicators highlighted in bold for each engagement metrics (if exists), SSH shows the most extensive distribution of all kinds of user engagement, thus acting as the most active subject field in giving rise to engagement with science on Twitter, followed by LES. BHS, as the subject field with the most scholarly tweets, contributes the most to the overall engagement metrics data due to the largest total number of corresponding engagement records. Besides, scholarly tweets of BHS papers also have a relatively higher presence of user engagement. However, user engagement is confirmed to be sparsely distributed among scholarly tweets in the fields of MCS and PSE.

Table 4. Descriptive statistics of the four engagement metrics across the five subject fields

Metrics	Field	Sum	Mean	Min	Q1	Q2	Q3	90th P	99th P	Max	SD
Likes	SSH	2,439,687	**3.63**	0	0	**1**	**3**	**7**	**43**	6,796	**29.03**
	BHS	**10,092,255**	2.85	0	0	0	2	6	38	9,336	17.66
	PSE	1,302,420	1.93	0	0	0	1	4	25	**10,105**	21.67
	LES	2,845,167	3.23	0	0	0	2	**7**	42	5,127	19.10
	MCS	370,053	2.32	0	0	0	1	4	32	3,904	22.73
Retweets	SSH	1,463,790	**2.18**	0	0	0	**2**	**5**	25	8,492	**24.06**
	BHS	**6,692,517**	1.89	0	0	0	1	4	25	8,317	15.59
	PSE	642,819	0.95	0	0	0	1	2	12	**9,983**	15.58
	LES	1,822,153	2.07	0	0	0	1	5	**26**	7,495	14.88
	MCS	209,015	1.31	0	0	0	1	2	19	6,255	20.70
Quotes	SSH	159,924	**0.24**	0	0	0	0	**1**	**4**	**673**	**2.49**
	BHS	**551,606**	0.16	0	0	0	0	0	3	584	1.22
	PSE	51,124	0.08	0	0	0	0	0	2	319	0.83
	LES	145,313	0.16	0	0	0	0	0	3	235	1.09
	MCS	15,316	0.10	0	0	0	0	0	2	308	1.31
Replies	SSH	121,918	**0.18**	0	0	0	0	0	**3**	**1,033**	**1.94**
	BHS	**388,313**	0.11	0	0	0	0	0	2	274	0.76
	PSE	43,566	0.06	0	0	0	0	0	1	161	0.58
	LES	95,471	0.11	0	0	0	0	0	2	98	0.63
	MCS	12,399	0.08	0	0	0	0	0	2	139	0.72

Note: Sum = total number of corresponding engagement metrics; Q1, Q2, Q3 = the first, second, and third quartile; 90th P = the 90th percentile; 99th P = the 99th percentile; Min, Max = the minimum and maximum value; SD = standard deviation.

4.3.3 Correlation analysis of user engagement metrics

In this part, to study how different types of user engagement metrics correlate with the factors listed in Table 2, through the lens of each factor, we graphically show the coverage of specific user engagement of scholarly tweets aggregated at the different levels of the corresponding factor, and further interpret the observed relationships by combining the results of the Spearman correlation analysis between studied factors and user engagement metrics at the tweet level.

From the perspective of scholarly impact of tweeted papers, Figure 4 plots the change trend of the coverage of user engagement with the increase of **A** citations and **B** Mendeley readers of tweeted papers. Overall, the coverage of all kinds of user engagement is slightly higher for scholarly tweets mentioning papers with higher levels of citation counts and Mendeley readers accrued, although the uptrends are not that strong, particularly for citations. According to the Spearman correlations (see Figure 7 in Appendix), the four types of user engagement metrics are all positively but negligibly correlated with citations and readers (the coefficient r_s ranges from 0.016 to 0.048 for citations, and ranges from 0.051 to 0.107 for readers).

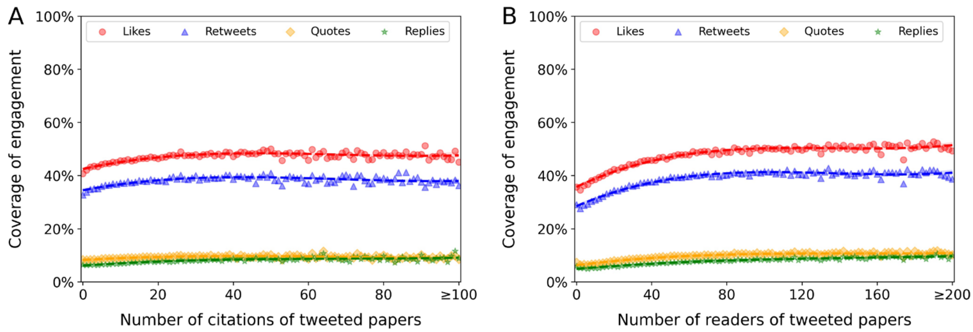

Figure 4. Coverage of the four types of user engagement among scholarly tweets with different levels of **A** WoS citations and **B** Mendeley readers received by tweeted papers

Regarding tweet features used in scholarly tweets, Figure 5 shows the coverage of user engagement when different **A** numbers of hashtags are used, and different **B** numbers of users are mentioned in tweets. These two tweet feature factors present different patterns in their relationships with user engagement. As the number of hashtags per tweet increases, a slight rise can be observed in the coverage of likes, retweets, and quotes, but not for replies. This is confirmed by the positive and negligible correlations found between number of hashtags and number of likes, retweets, and quotes received by tweets (r_s ranges between 0.042 and 0.113), whereas nearly no correlation found between number of hashtags and number of replies (r_s = -0.001). By comparison, the uptrend of the coverage of user engagement is stronger with the increasing number of mentioned users in tweets, especially for the coverage of likes and retweets. For scholarly tweets with more than two users mentioned, their likelihood of being liked exceeds 80% and the likelihood of being retweeted reaches 70%. Similarly, the coverage of both quotes and replies is relatively higher for scholarly tweets with more users mentioned than those without any mentioned users. Correspondingly, the Spearman correlations between user engagement metrics and number of mentioned users are comparatively stronger than other factors mentioned earlier, particularly for likes (r_s = 0.237) and retweets (r_s = 0.229).

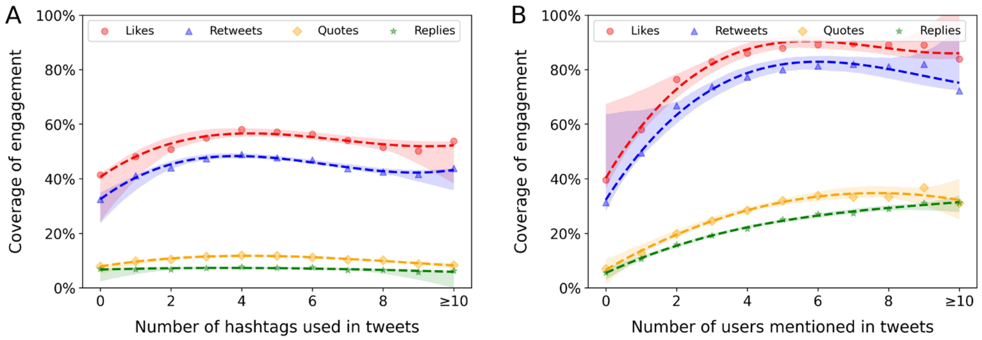

Figure 5. Coverage of the four types of user engagement among scholarly tweets with different numbers of **A** hashtags used and **B** users mentioned in tweets

In terms of user characteristics, Figure 6 shows the coverage of user engagement for scholarly tweets posted by users with different characteristics. Specifically, Figure 6**A** and 6**B** exhibit the change trend of the likelihood of being engaged with for scholarly tweets from users with different levels of followers and lists listed, respectively. These two factors, which to a large extent determine the audience size of posted tweets, are positively associated with the coverage of user engagement. The more followers that users accrue or the more lists that users are listed in positively correlate to the higher probability for their tweets of getting engagement. Based on the Spearman correlation analysis, number of followers of users is moderately correlated with both number of likes and retweets received (r_s = 0.349 and 0.368, respectively).

Figure 6**C** and 6**D** show the relations between user engagement and the two factors about users' interaction activity: number of friends and number of likes given. Similarly, these two factors also keep positive relationships with the coverage of user engagement. Overall, scholarly tweets posted by more active users (who interact more frequently with others by following other users and liking other users' tweets) tend to show a relatively higher probability to be engaged with. The Spearman correlation analysis proves that there exist weak to moderate correlations between user engagement metrics and the two factors about users' interaction activity (r_s ranges from 0.086 for the correlation between number of friends and number of quotes up to 0.313 for the correlation between number of likes given and number of likes received).

Different from the patterns observed for the above factors, as shown in Figure 6**E** and 6**F**, number of tweets posted and science focus of users, which depict the overall tweeting activity of users, show negative relationships with the coverage of user engagement among scholarly tweets. In general, the greater number of tweets posted, as well as the stronger science focus of users, the lower levels of coverage of user engagement. These negative relationships are reinforced by the negative coefficients reported in the Spearman correlation analysis between these factors and user engagement metrics (r_s ranges from -0.147 to -0.028).

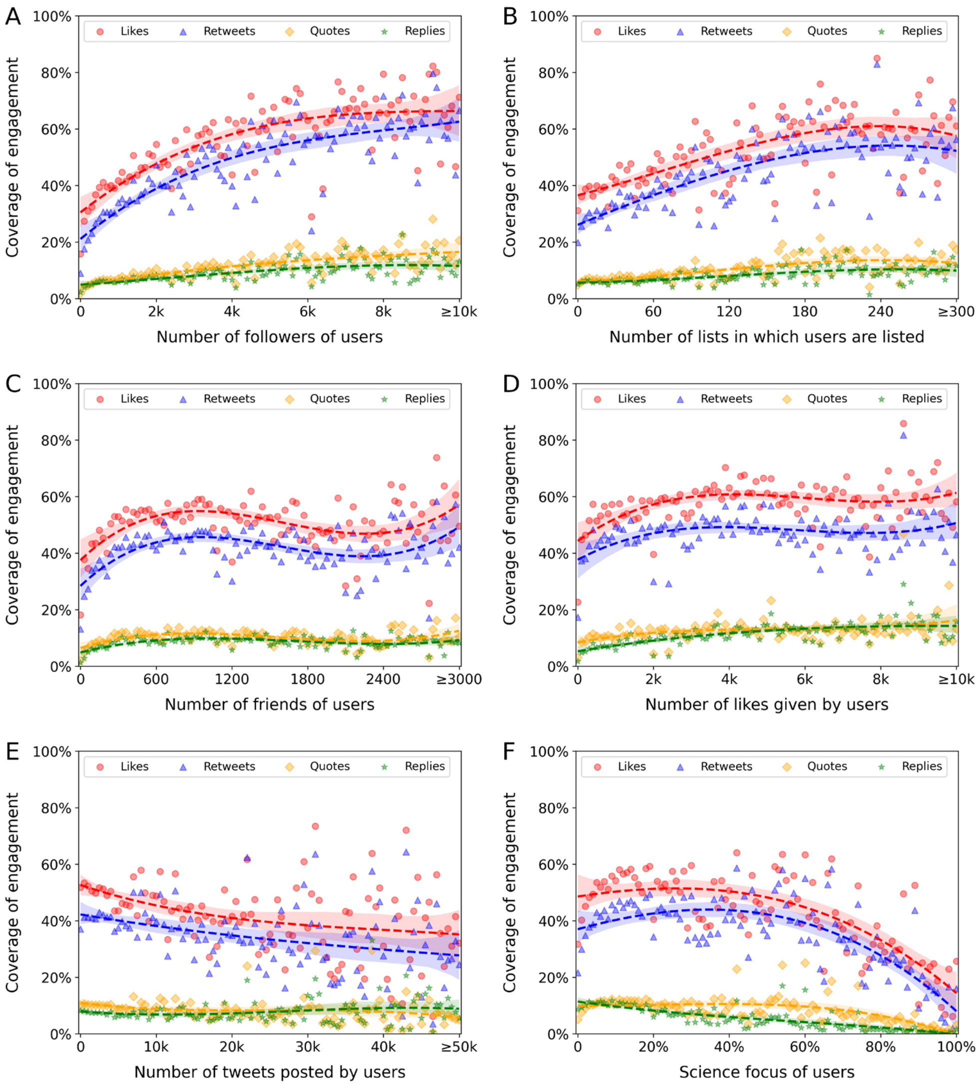

Figure 6. Coverage of the four types of user engagement among scholarly tweets posted by users with different levels of **A** followers, **B** lists listed, **C** friends, **D** likes given, **E** tweets posted, and **F** science focus

4.3.4 Regression analysis using hurdle models

To further compare how different science-based and Twitter-based factors serve as predictors of user engagement behaviors, we conducted regression analyses of the four types of user engagement using hurdle models. In order to avoid multicollinearity in the regression models, in each of the three dimensions discussed above, we selected several representative factors as the explanatory variables. For instance, in the dimension of scholarly impact of tweeted papers, since citations and Mendeley readers are strongly correlated with each other (r_s = 0.712), we selected citations as one of the explanatory variables. In the dimension of tweet features, both number of hashtags used and number of mentioned users were included because they are weakly correlated (r_s = 0.181). In the dimension of user characteristics, number of followers keeps moderate to strong correlations with other homogeneous user factors (r_s ranges from 0.433 to 0.859) except science focus of users (r_s = 0.015), so we adopted number of followers and science focus as representative variables in this dimension.

Table 5 reports the results of the zero portion of the four hurdle models of user engagement metrics (logit models): model 1 (likes), model 2 (retweets), model 3 (quotes), and model 4 (replies). Some variables were log-transformed for better model fitting. The zero portion of the models reflects how the selected explanatory variables relate to the likelihood for scholarly tweets of attracting at least one specific user engagement. The four models present a similar pattern in the zero portion, with citations, mentioned users and followers positively associated with the likelihood of getting at least one corresponding user engagement, whereas science focus presents a negative association. The exception is hashtags: in model 1 (likes), mode 2 (retweets) and model 3 (quotes), number of hashtags has a positive relationship with the likelihood that at least one like, retweet or quote occurs, however, in model 4 (replies), number of hashtags presents a negative association. The odds ratios (OR, exponent of regression coefficient in logit model) were calculated to help interpret the results. For example, in model 1 (likes), while all other variables remaining constant, a one-unit increase in the log-transformed number of mentioned users increases the odds of getting at least one like by 185.6% (OR = 2.856). However, while holding all other variables constant in model 1 (likes), a unit increase in science focus decreases the odds of getting at least one like by 42.7% (OR = 0.573).

Table 6 reports the results of the count portion of the hurdle models of user engagement metrics (zero-truncated negative binomial models). The count portion focuses on those scholarly tweets with at least one corresponding user engagement received and indicates how the explanatory variables associate with the increase of the frequency of user engagement. As is evident in all the four models, citations, mentioned users as well as followers are all positively associated with the frequency of user engagement, yet hashtags and science focus are negatively associated with the frequency of user engagement. Similarly, incidence rate ratios (IRR, exponent of regression coefficient in negative binomial model) were computed to help interpret the coefficient of a given variable while holding all other variables constant. For instance, in model 1 (likes), while all other variables remaining constant, a unit increase in the log-transformed number of followers increases the rate of receiving a like by a factor of 1.484, while a unit increase in the log-transformed number of hashtags decreases the rate of obtaining a like by a factor of 0.908.

Table 5. Results of the zero portion of the hurdle models of user engagement metrics

Variable	Model 1 (likes)		Model 2 (retweets)		Model 3 (quotes)		Model 4 (replies)	
	Estimate (SE)	OR	Estimate (SE)	OR	Estimate (SE)	OR	Estimate (SE)	OR
Citations (log-transformed)	0.058 (0.001)	1.059	0.060 (0.001)	1.062	0.030 (0.001)	1.030	0.058 (0.001)	1.060
Hashtags (log-transformed)	0.100 (0.002)	1.105	0.164 (0.002)	1.178	0.005 (0.003)	1.005	-0.245 (0.003)	0.782
Mentioned users (log-transformed)	1.049 (0.002)	2.856	1.035 (0.002)	2.815	0.866 (0.003)	2.378	0.791 (0.003)	2.205
Followers (log-transformed)	0.383 (0.000)	1.466	0.446 (0.001)	1.562	0.389 (0.001)	1.475	0.308 (0.001)	1.360
Science focus	-0.557 (0.003)	0.573	-0.378 (0.003)	0.685	-0.582 (0.005)	0.559	-2.228 (0.007)	0.108

Note: All estimates are significant at the 0.000 level. SE = Standard error. OR = Odds ratio (exponent of estimate in logit model).

Table 6. Results of the count portion of the hurdle models of user engagement metrics

Variable	Model 1 (likes)		Model 2 (retweets)		Model 3 (quotes)		Model 4 (replies)	
	Estimate (SE)	IRR	Estimate (SE)	IRR	Estimate (SE)	IRR	Estimate (SE)	IRR
Citations (log-transformed)	0.068 (0.001)	1.070	0.137 (0.001)	1.147	0.105 (0.002)	1.110	0.009 (0.002)	1.009
Hashtags (log-transformed)	-0.096 (0.002)	0.908	-0.067 (0.002)	0.936	-0.182 (0.005)	0.833	-0.117 (0.006)	0.890
Mentioned users (log-transformed)	0.661 (0.003)	1.936	0.475 (0.003)	1.609	0.238 (0.005)	1.268	0.189 (0.006)	1.208
Followers (log-transformed)	0.395 (0.001)	1.484	0.395 (0.001)	1.484	0.369 (0.001)	1.446	0.339 (0.002)	1.403
Science focus	-1.122 (0.005)	0.326	-0.922 (0.005)	0.398	-0.950 (0.011)	0.387	-1.595 (0.017)	0.203

Note: All estimates (coefficients) are significant at the 0.000 level. SE = Standard error. IRR = Incidence rate ratio (exponent of estimate in negative binomial model).

4.4 Discussion

As discussed by Brossard and Scheufele (2013), in the era of mass media, science stories as well as their accuracy, importance and popularity are no longer just "presented in isolation but instead are embedded in a host of cues that accompany nearly all online news stories", such as comments on blog posts and user engagement on social media. Such cues, according to Brossard and Scheufele (2013), "may add meaning beyond what the author of the original story intended to convey". In the context of scholarly Twitter metrics, this argument, on the one side, highlights the importance of the examination of user engagement in studying science-social media interactions, but on the other side, poses a question about how many scholarly tweets indeed triggered user engagement which are believed to contain extra meaning added to science stories.

Although user engagement with scholarly tweets have long been seen valuable for characterizing the interactions between scholarly objects and social media (Wouters et al., 2019), there is still an overall lack of evidence which can be drawn upon to mirror how effectively scholarly tweets attract the public's attention and further stimulate public engagement in Twitter conversations around science. Based on a large-scale and cross-disciplinary dataset, this study unravels the coverage of diverse types of user engagement among scholarly tweets, thus offering an answer to the question about the overall presence of public engagement with scientific information on Twitter.

4.4.1 Overall presence of user engagement with scholarly tweets

As conceptualized by Haustein, Bowman, & Costas (2016) in the context of *primary social media metrics* with scholarly objects as the objects of study, they classified acts referring to scholarly objects to three main categories, including *access*, *appraise*, and *apply*. *Access* refers to acts that involve "accessing and showing interest in the research objects", such as viewing and downloading a scientific paper. *Appraise* stands for acts of "mentioning the research objects on various platforms" like blogs and social media. *Apply* includes acts of "using significant parts of, adapting, or transforming the research objects", such as thoroughly discussing a scientific paper in a blog post or citing it in papers. Therefore, *apply* represents the highest level of engagement with research objects, followed by *appraise* and then *access*. Following this framework, we applied it in the context of *secondary social media metrics* (Díaz-Faes et al., 2019), in which social media users and their online activities become the objects of study, instead of the scholarly objects as in Haustein, Bowman, & Costas (2016). Correspondingly, in the specific case of Twitter engagement metrics, *access* would indicate acts of accessing and showing interest in the scholarly tweets and their constitutive elements, such as viewing a tweet and adding a tweet to bookmarks. *Appraise* would refer to acts of commending and further disseminating the scholarly tweets, such as liking a tweet or retweeting it. Those tweets liked or retweeted by a user would be displayed on the user's homepage and have notifications sent to other users involved in the tweets (e.g., authors of the tweets, and users mentioned in the tweets). Lastly, *apply* would include acts of participating in discussions and expressing views based on the scientific information tweeted, for example, retweeting with extra comments added (i.e., quoting) or making a response to a tweet (i.e., replying), which would contribute to the creation of another Twitter form of engagement that is a *conversation*[5]. The category of *apply* also has records publicly visible on actors' homepage and have notifications sent to users involved, and more importantly, such acts will generate more information (and possibly subsequent engagement) that the original scholarly tweets may not contain. From the standpoint of scholarly tweets, the level of engagement increases from *access* over *appraise* to *apply* as well.

It's not surprising to find that as the level of engagement grows, the coverage of user engagement behavior becomes lower. In this study, with likes, retweets, quotes and replies as the traces of user engagement, we found that likes and retweets, as the acts of *appraise* with the moderate level of engagement, were present for about 44% and 36% of the studied scholarly tweets, respectively. However, the coverage of quotes and replies, the two behaviors with the highest level of engagement (i.e., *apply*), is as low as 9% and 7%, respectively. The globally low presence of user engagement, particularly for the engagement behaviors with more informative outcomes generated, reveals the fact that the attention paid to scholarly tweets varied a lot on the one hand, and puts more emphasis on the significance of more in-depth measurement of Twitter reception of scientific papers on the other hand. For papers with exactly the same number of scholarly tweets accumulated, although the papers' Twitter reception appears to be equal only based on their absolute number of scholarly tweets, those with scholarly tweets being widely engaged with might be disseminated and perceived on Twitter in a more effective way. This is because engagement behaviors provide concrete evidence that they reached out to audiences who also showed further interest.

Moreover, the presence of user engagement differs by subject field. As the subject fields found to be more frequently mentioned in the Twittersphere (Costas et al., 2015a; Haustein, Costas, et al., 2015), SSH, LES, and BHS also have their scholarly tweets more actively engaged with by users through liking, retweeting, quoting, and replying, outperforming the fields of PSE and MCS. Behind the consistency of SSH, LES, and BHS shown in the vitality in the Twitter environment, there are multiple possible reasons such as the lay audiences' preference for topics related to social issues, environmental problems, and healthcare (Haustein, Costas, et al., 2015; Haustein, Peters, Sugimoto, et al., 2014), and the higher degree of Twitter uptake by scholars from these fields (Costas et al., 2020; Mohammadi et al., 2018). In addition, it has been reported that scholars from the field of social sciences and

[5] As defined by Twitter (https://help.twitter.com/en/using-twitter/twitter-conversations) (Accessed April 28, 2021), a conversation on Twitter is composed of an original tweet and its replies, as well as replies to those replies.

humanities more frequently communicate their research with the public as an important audience (Bentley & Kyvik, 2011) and more often devote to popularization activities than scholars from natural sciences and technology (Kreimer et al., 2011), which can be partly explained by the "less strict demarcation between internal scientific and public communication and between scientific and general knowledge" existed within social sciences and humanities than natural sciences (H. P. Peters, 2013). This might be another possible reason for the superiority of SSH in obtaining further engagement. To further interpret the subject field differences, future research is needed to scrutinize the contexts in which user engagement takes place (e.g., engaging users' identity and motivations) across subject fields.

4.4.2 Factors related to user engagement with scholarly tweets

On the basis of both correlation analysis and regression analysis, we investigated the relationships between user engagement and a spectrum of science-based and Twitter-based factors. Similar to previously reported weak or no correlations between citations and tweeting activities (Bardus et al., 2020; Zahedi et al., 2014), we found that user engagement with scholarly tweets was also negligibly correlated with scholarly impact factors of tweeted papers (i.e., citations and Mendeley readers), thus adding more empirical evidence to the idea that science and social media have different concerns about research outputs and conform to different spaces of engagement (Fang et al., 2021).

In contrast to science-based factors, Twitter-based factors generally tend to be more related to user engagement. Specifically, from the perspective of tweet features, although hashtags and user mentions are both tweet features increasing the visibility of tweets, the former is utilized to label and broadcast tweets to potential users interested in the same topics, while the latter is targeted to specific users with notifications delivered to them, showing a more conversational nature than the former. As a result, number of users mentioned in tweets is more related to user engagement and more effective in predicting user engagement. From the perspective of user characteristics, both users' social media capital (i.e., followers and lists listed) and interaction activity (i.e., friends and likes given) were positively correlated with user engagement around their tweets. Nevertheless, flooding the screen (i.e., too many tweets posted) and attaching to tweeting only scientific papers (i.e., too strong science focus) were found to be related to lower levels of user engagement. From a practical point of view, as suggested by Cheplygina et al. (2020) for scientists getting start on Twitter, building a community by interacting with others, as well as sharing something personal and non-academic can also be relevant to get support in science communication on Twitter.

This study took into account scholarly tweets written in all languages. With this we provide a relatively complete picture of user engagement regardless of the language of the tweets. However, more detailed tweet content analysis should be applied, considering different linguistic contexts, as well as more local topics and sentiment, which were not included in this study. Given that specific tweet content like those including awe-inspiring imagery and newsworthy items frequently attract high levels of engagement across social media platforms (Kahle et al., 2016), future research should also focus on how different tweet content might be related to subsequent user engagement in scholarly contexts.

4.4.3 Implications for social media studies of science

As an important part of the Twitter information ecosystem, user engagement behaviors leave digital traces of wider public interactions with science, thereby allowing both for investigation of online scholarly communication and civic participation in science-focused discussions, and for exploration of the deeper levels of Twitter reception of science from the standpoint of broader social media audiences. Correspondingly, the implications of studying user engagement for more advanced social media studies of science are two-fold.

On the one hand, in terms of science-social media interactions, user engagement provides additional information beyond what is delivered by scientific papers and original Twitter mentions, especially for those behaviors with higher levels of engagement such as quoting and replying. These engagement behaviors act as sources of information on how users communicate science in non-academic environments and how the public at large receives these messages related to science. For those scholarly tweets with informative commentaries or conversations, they may offer valuable evidence to develop a more comprehensive understanding of science-social media interactions.

On the other hand, in terms of impact measurement on Twitter, the presence of scholarly tweets has been regarded as an important indicator of social (media) attention paid to research outputs (Sugimoto, Work, et al., 2017). However, it only reflects the attention of users who brought scientific information to Twitter, but neglects the attention of those Twitter audiences who engaged with this scientific information through diverse engagement behaviors. User engagement metrics would then capture a more deep-seated reception of science in the Twitter universe, complementing the “science stories” in the social media environment (Brossard & Scheufele, 2013). Therefore, including user engagement metrics in the altmetric toolkit might open a novel window to characterize the popularity of research outputs. This argument, although based on the study of scholarly tweets, can be generalized to other altmetric data sources with user engagement metrics available (e.g., likes and shares on Facebook posts, views on YouTube videos) or potential (e.g., readers of blog posts, comments in news media platforms).

4.4.4 Limitations

There are several limitations in this study. First, there are more than four types of engagement behavior that users can take to interact with scholarly tweets, such as clicking on tweeted scholarly URLs, clicking on users’ profile, and adding tweets to bookmarks. However, these engagement metrics are currently not publicly accessible on a large scale, and they were not included in this study. Should these engagement metrics be made publicly retrievable in the future, a more complete picture of user interaction behavior around science could be drawn. Second, although reply tweets and quote tweets, which are outcomes of replying and quoting behavior, can be further engaged with through Twitter functionalities as well (e.g., liking or retweeting replies and quote tweets), they were not included in the analyzed dataset to avoid double counting. Lastly, we explored how the characteristics of engaged users (i.e., users whose tweets are engaged with) relate to user engagement, while the characteristics of engaging users (i.e., users who engaged with tweets) are also of relevance to the interpretation of the occurrence of engagement behavior. However, due to the lack of detailed information of the engaging users, their demographics and behavioral patterns were not analyzed in this study.

4.5 Conclusions

This study contributes to the expansion of the idea of *secondary social media metrics* (Díaz-Faes et al., 2019) by presenting a large-scale and cross-disciplinary analysis of four types of user engagement (i.e., liking, retweeting, quoting, and replying) around scholarly tweets. Of the 7 million scholarly tweets in our dataset, 52% were engaged with by Twitter users through at least one engagement behavior. Likes and retweets are most widespread, with the highest coverage among scholarly tweets (44% and 36%, respectively). In contrast, the coverage of quotes and replies is much lower (9% and 7%, respectively), suggesting the overall low presence of user engagement amongst Twitter mentions of scientific papers, particularly for those behaviors with higher levels of engagement needed. Scholarly tweets from the fields of SSH, LES, and BHS tend to have more frequent user engagement distributed. Finally, the presence of user engagement is more related to other Twitter-based factors (mentioned users in tweets and number of followers of users in particular) than with science-based factors of papers (e.g., citations and Mendeley readers), implying both the intrinsically connected dynamics of Twitter elements and the distinguishing focuses between scientific and tweeting activities.

Our findings provide a first overview of the extent to which scholarly tweets are related to broader public engagement with science on Twitter, thereby paving the way towards the measurement of Twitter reception of science in a more interactive and comprehensive manner. Based on the exploratory results presented in this study, a series of research questions emerge, which will need to be examined in much greater detail, such as the motivations and behavioral patterns of engaging users, the differential aspects that increase the social media capital of Twitter users (e.g., by increasing their number of followers), and what topic-related factors (e.g., controversial topics) embodied in tweets can be related to be triggering more effective forms of engagement and Twitter communication. All in all, delving into user engagement behaviors may help delineate the role that Twitter plays in facilitating public understanding of science as well as the meaning that Twitter-based indicators may have in research evaluation and science communication.

4.6 Appendix

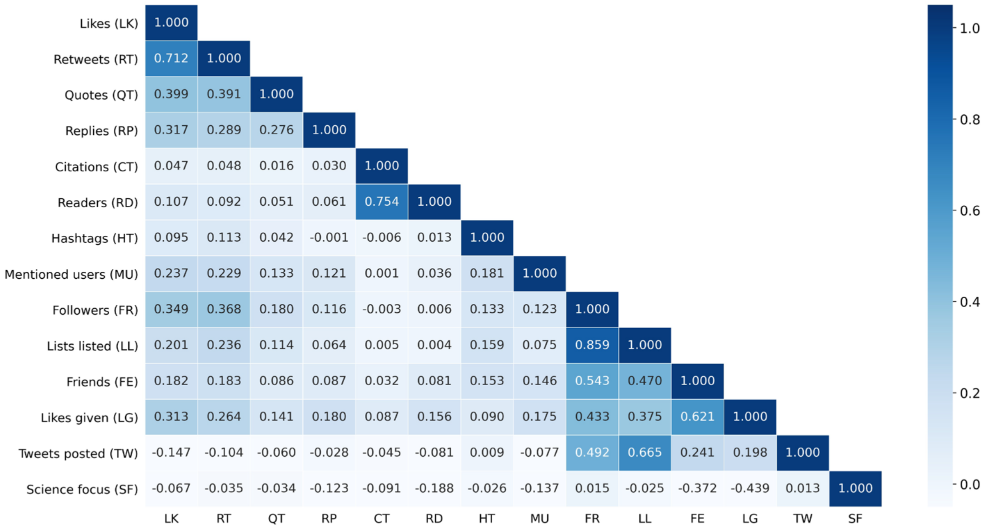

Figure 7. Spearman correlation analysis of the four user engagement metrics and studied factors