



an open access  journal



Citation: Maleki, A., & Holmberg, K. (2026). Who are tweeting about academic publications? A systematic review and meta-analysis of altmetric studies. *Quantitative Science Studies*, 7, 485–520. <https://doi.org/10.1162/QSS.a.464>

DOI:
<https://doi.org/10.1162/QSS.a.464>

Peer Review:
<https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/QSS.a.464>

Supporting Information:
<https://doi.org/10.1162/QSS.a.464>

Received: 7 April 2023
Accepted: 11 November 2025

Corresponding Author:
Ashraf Maleki
ashraf.maleki@utu.fi

Handling Editor:
Rodrigo Costas

Copyright: © 2026 Ashraf Maleki and Kim Holmberg. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.



REVIEW ARTICLE

Who are tweeting about academic publications? A systematic review and meta-analysis of altmetric studies

Ashraf Maleki^{ID} and Kim Holmberg^{ID}

Department of Social Research, University of Turku, Turku, Finland

Keywords: altmetrics, meta-analysis, scholarly communications, systematic review, user categorization, X/Twitter

ABSTRACT

Understanding who shares academic publications on Twitter is critical to interpreting altmetrics as signals of scholarly or societal impact. Prior studies have used diverse and often incompatible user classification schemes, making synthesis difficult. This study presents a systematic review and meta-analysis of 23 empirical studies (covering 79,014 Twitter users, over 20 million tweets, and more than 5 million tweeted publications) to estimate category-specific engagement across three metrics: user counts, tweets, and tweeted publications. We developed a harmonized categorization scheme encompassing 11 user types and applied both random effects models (REM) and beta-binomial hierarchical models (BBHM) to estimate proportions, account for study-level variation, and model uncertainty. Across all indicators, individual users were the most active, comprising 66% of users, 55% of tweets, and 50% of tweeted publications. BBHM further enabled in-category vs. out-of-category comparisons and revealed engagement differences not detected by REM. *t*-tests on study-level means confirmed significant differences between academic individuals and other user types. Despite methodological heterogeneity, results consistently show that academic and nonacademic individuals statistically equally dominate Twitter engagement with scholarly content. Our findings support the need for standardized user classification schemes and demonstrate the value of Bayesian modeling for synthesizing altmetric data in study variation and sparsity.

1. INTRODUCTION

Investigating the online mentions of research outputs, or the altmetrics of research outputs, can reveal where, how, and by whom research has been shared, disseminated, and discussed on various online platforms. Such information could lead to an increased understanding of how research may have had some type of impact or influence on wider audiences beyond academia, and with that, how altmetrics could be used for impact assessment (Holmberg, 2015). But this requires that we understand who shares, disseminates, or discusses scientific documents online and for what reason. Twitter, or X as the service is currently called, is one of the main sources of altmetrics. and several earlier studies have focused on investigating the people who share scientific articles on Twitter. Because this study investigates earlier studies that were conducted before the name change, we will keep referring to the service as Twitter. To understand whether Twitter can reflect a wider societal interest or attention towards scientific content, it is important to understand the diverse roles, motivation, and behaviors of Twitter users who share and interact with scientific articles on the platform. Without user characterization, altmetrics

studies risk missing out on valuable insights into the impact and reach of scholarly content on social media platforms. Because different user types may interact with content in distinct ways, without a Twitter user categorization it would be difficult to interpret the engagement metrics.

Despite Twitter's open structure, categorizing its users based on their occupation, expertise, or motivation remains challenging. While some studies have identified scientists on Twitter (Schmitt & Jaschke, 2017) and their professional use of the platform (Álvarez-Bornstein & Montesi, 2019), others highlight the mix between personal and professional engagement (Bowman, 2015; Toupin, Millerand, & Larivière, 2019). This variability in user classification approaches complicates cross-study comparisons and generalization of findings, potentially leading to inconsistent conclusions about who actually drives research dissemination on Twitter. Without clear and standardized user categories, altmetrics studies struggle to accurately contextualize engagement patterns, leading to difficulties in distinguishing between academic vs. nonacademic attention and scholarly vs. social impact.

While numerous classification techniques have been used to explore Twitter users engaging with research, a systematic, meta-analytical synthesis of these studies is lacking. Previous studies have reviewed and condensed research findings on various topics, such as the utilization of Twitter by medical journals (Erskine & Hendricks, 2021), the association between altmetrics and citations (e.g., Bornmann, 2015; Erdt, Nagarajan et al., 2016), and the comparison of altmetric data providers (Ortega, 2020). However, no study has systematically consolidated findings regarding who tweets about scientific content and to what extent. This meta-analysis aims to bridge this gap by quantitatively synthesizing findings from prior user classification studies, and providing a clearer picture of the composition and behaviors of Twitter users engaging with academic publications.

Meta-analysis methods traditionally rely on frequentist random effects models, but such methods have limitations when study-level variability is high or when the number of studies is small. These characteristics are particularly relevant to the meta-analysis of Twitter user categorization research, where the number of studies is limited, yet many studies analyze millions of tweets. Bayesian hierarchical models (BHM) have become increasingly popular for the meta-analysis of social science and biomedical data, where heterogeneity across studies and small study counts pose challenges (Gelman, Carlin et al., 2013; Schmid, 2001). A Bayesian hierarchical model is a probabilistic framework that estimates group-level effects by pooling information across studies while accounting for uncertainty and variability at multiple levels of analysis, therefore offering several advantages over frequentist approaches, including the ability to incorporate weakly informed priors as binary variables (Negrín-Hernández, Martel-Escobar, & Vázquez-Polo, 2021) and greater flexibility in modeling study-level differences. Hierarchical Bayesian models have been applied in capturing intensity and continuity of scholars' productivity over time (Mezzetti & Negri, 2024) and publication bias corrections (Schmidt & Hunter, 2015), which demonstrate their robustness in meta-analytical settings.

The beta-binomial hierarchical model (BBHM) is particularly useful in studies where success and failure events are being measured. This approach has been a widely applied meta-analysis method for a few studies involving rare events (Günhan, Röver, & Friede, 2020; Jansen & Holling, 2023), and identifying individual differences in population studies (Doebler, Frick, & Doebler, 2024). The beta-binomial structure is advantageous because it naturally models count data and gives a success rate in proportion; accounts for study-level variability, ensuring that studies with different sample sizes contribute proportionally to the meta-analysis; and more importantly allows for robust inference even when the number of studies is small despite large data sets in each study, a common limitation faced in meta-analysis of altmetrics

research. This approach enables us to estimate the likelihood of previous studies identifying different user types by better estimation of uncertainty and credible intervals. In Bayesian models, full posterior distributions are provided rather than relying on point estimates and confidence intervals, which enhances the understanding of engagement patterns. Given the diverse nature of Twitter studies, Bayesian methods allow for the inclusion of weak evidence, reducing the risk of bias.

2. BACKGROUND

Earlier studies have had different objectives for their user categorizations, leading to different approaches. Profile descriptions of Twitter users are often the basis for the coding, resulting in category dependence on the information that users share about themselves. Many studies (13) have utilized two or three coders to enhance intercoder reliability (Yu, Xiao et al., 2019; Zhou & Na, 2019), but there are also studies that have tried to analyze broader patterns, adopting automated or semiautomated user categorization methods on larger data sets. The precision and accuracy of automatic categorizations based on keywords in Twitter profiles have been shown to have some limitations. Barthel, Tönnies et al. (2015) for instance, used an automated approach to detect scientists among a group of tweeters, reaching an accuracy of 88%, while another more recent study showed that over 95% precision can be obtained by matching the names of 500,000 scientists on ORCID and Twitter usernames on Crossref Event data (Mongeon, Bowman, & Costas, 2023). Manual coding of studies originates in researchers' interest to more accurately indicate diversity of user categories across either their individual and organizational association (Orduña-Malea & Costas, 2021) or various professions (de Melo Maricato & de Castro Manso, 2022). Many studies have been interested in bringing the focus to only particular user types that are actively involved in sharing scientific publications, such as academics (Zhou & Na, 2019), health practitioners, such as surgeons (Coret, Rok et al., 2020), publishers and journal accounts (Ortega, 2017), and bots (Haustein, Bowman et al., 2016a), or a combination of these (Haustein, 2019; Ye, Na, & Oh, 2022). Other data from profile descriptions, such as occupation, gender (e.g., Ke, Ahn, & Sugimoto, 2017; Tsou, Bowman et al., 2015), and geographic location (Dudek, Bowman, & Costas, 2018; Haustein, 2019; Yu et al., 2019) have also been used for user categorization. Some studies have used a binary categorization, such as *bot* and *nonbot* (Didegah, Mejlgaard, & Sørensen, 2018; Haustein et al., 2016a; Haustein, Tsou et al., 2016b; Yu et al., 2019); *academic* and *nonacademic* (Didegah et al., 2018; Htoo & Jin-Cheon, 2017; Ke et al., 2017; Zhou & Na, 2019); *male* or *female* (Orduña-Malea & Costas, 2021); and *expert* or *nonexpert* (Pandian, Na et al., 2019). Other studies have used three categories, such as *individual*, *organization*, and *the rest*, which have included unknown or unidentifiable accounts (Haustein et al., 2016b), or *professional*, *nonprofessional*, and *composite*, where the categorization was based on the professionalism expressed in the profile description (Yu et al., 2019). A few studies have taken a more detailed approach, categorizing Twitter users into more numerous categories in order to broaden the perspective of the types of users involved in sharing scientific articles, such as the nine categories used by Haustein and Peters (2020) and the 27 categories used by Vainio and Holmberg (2017). On the other hand, many studies have used the categorization provided by the altmetric data provider Altmetric.com, which identifies only some specific types of users, such as *scientists*, *practitioners*, and *science communicators*, everything else being categorized as *members of the public*.

Overall, earlier studies have shown that users who tweet scientific content can come from a range of backgrounds and include both scholars and members of the public (Haustein, Costas, & Larivière, 2015), but some studies have shown that academics may be overrepresented among those that tweet scientific journal articles. Tsou et al. (2015) found that about a third

of Twitter users who had tweeted research articles had a doctoral degree. In another study, which surveyed 1,912 people who had tweeted journal articles, it was found that over half (55%) of the respondents belonged to academia (Mohammadi, Thelwall et al., 2018).

We aim to conduct the first meta-analysis of altmetric data and to combine the results from previous studies and provide a more generalizable understanding of who tweets scientific articles. Understanding who tweets scientific articles is crucial because it allows researchers and institutions to better understand the patterns for dissemination of scholarly information and assess the impact of research beyond traditional academic circles. Central to meta-analysis is quantitative integration of studies with room for the addition of goals of critiquing a field of study and identifying key directions for future conceptual, methodological, and empirical work (Card, 2015, p. 17). A meta-analysis allows for the systematic synthesis and comparison of findings from multiple earlier studies. Integration involves either combining effect sizes of studies or comparing them if the effect sizes found in studies systematically differ depending on study characteristics (Card, 2015). By pooling data from earlier studies, meta-analysis enables us to identify patterns, trends, and discrepancies across different studies, while providing a more comprehensive understanding of the factors influencing who tweets scientific articles. Our chosen approach also enhances statistical power, enabling more robust conclusions to be drawn regarding the characteristics of individuals who engage in tweeting academic research. In addition, our approach will involve a novel and comprehensive user recategorization that can better accommodate the categorizations used in earlier studies and which can be used in future studies.

3. METHOD

To answer the question of “who” is tweeting academic publications, studies that offered statistics about the categorization of Twitter users posting academic publications were included in our research. Both academic journal articles and relevant conference and workshop papers were considered eligible for the current work.

3.1. Review Framework

To ensure transparency, reproducibility, and comparability, this study followed systematic review and meta-analysis guidelines. The PRISMA checklist (Page, McKenzie et al., 2021) was used to structure the review, and the GRADE framework (Schünemann, Brożek et al., 2013) was adopted for evaluating the quality of evidence. The protocol was designed to facilitate future continuous reviews, especially given the diversity of various social media users and evolving nature of engagements across platforms like X/Twitter. Figure 1 gives an overview of the steps taken in the study.

3.2. Data Collection and Recategorization

The search terms “altmetric” and “Twitter” or “tweet” were entered into six academic publication databases (Web of Science, Scopus, Google Scholar, PubMed, Dimensions, and OpenAlex), two publishers’ websites (Wiley and Springer Nature), and the Altmetric workshops during November 2021 (more in Figure 3). In addition to browser searches, the software Publish or Perish (Harzing, 2007) was also used, and the search results were merged. The first round of screening was based on the titles and abstracts of the results. The second round of screening consisted of checking the full text of publications that reported some statistics at the Twitter user level, and data extraction was performed on those that met the criteria. The articles that contained a *Twitter user categorization* and enough information to conduct a meta-analysis on the

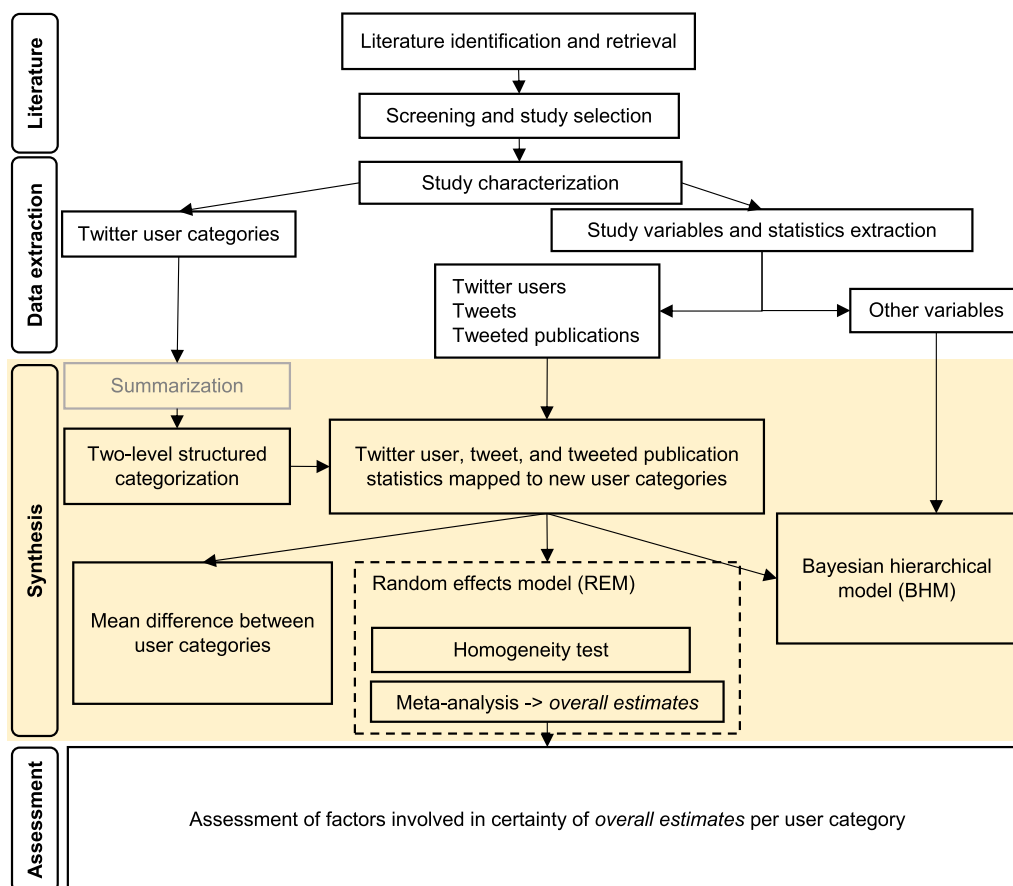


Figure 1. The steps taken for systematic review and meta-analysis of literature on contribution of different categories of Twitter users in dissemination of scientific publications.

statistical tests were then used for further analysis. Meta-analysis can be done with at least two studies (Valentine, Pigott, & Rothstein, 2010), but technically the power of tests in meta-analysis depends on a number of factors (Pigott, 2012, p. 37). Earlier studies had taken different approaches for user categorization, that also influenced the research design of the present study. As the earlier studies had collected their data with Twitter users, tweets, or tweeted publications, as a starting point, these three approaches were also used here to group the data from the chosen earlier studies. Thus, all the analyses were conducted separately on the studies in these groups: Twitter users, tweets, or tweeted publications. Relevant statistics were extracted from the content, or the tables and figures presented in the articles. Due to varied user categorizations across studies, we reclassified user types (Table 1) into a unified two-level structure:

1. Unit level: Individual, Organization, Science Communicator; and
2. Function level: Academic, Professional, Bot, Mixed/Other.

An additional Mixed Groups category was used for undefined or ambiguous users. This process reduced 376 categories (Table S1 in the Supplementary material) across the studies to 11 harmonized categories.

The category of *Individuals excluding Academics* was later added to make comparisons with *Academic Individuals*. This categorization prioritized the most significant categories

Table 1. Twitter user categories as in our recategorization with examples from the earlier studies (the new categories in **bold**, the abbreviations in parentheses, and examples of categories used in earlier studies in normal text)

<i>User Units (ownership)</i>	<i>User Functions (professional associations)</i>		<i>Examples/definition</i>
Individuals (I)	Academic (A)		scientist, researcher, faculty member
	Nonacademic or Individuals excluding Academics (I & A)	Professional (P)	practitioner, entrepreneur
		Mixed/Other (M)	student, citizen, nonacademic, unaffiliated, unspecified individual
Organizations (O)	Academic (A)		university, research institutions
	Professional (P)		enterprise, corporation, association, business
	Mixed/Other (M)		organizational accounts
Science communicators (SC): Accounts systematically disseminating scientific articles	Academic (A)		library, librarian
	Professional (P)		journal, publisher, media/news/journalist
	Bot (B)		bot, fully automated bots
	Mixed/Other (M)		partially automated bots, science communicator
Mixed groups (MG)	Mixed groups (MG)		unspecified accounts and a mix of unit level accounts: member of the public, nonbot, nonacademic, nonprofessional, unknown, undefined, unable to tell

while also accommodating both broad and detailed categorization approaches across the two levels, thereby enabling meta-analysis to encompass a wide range of studies.

Categorization variability across studies led to possible overlaps and inconsistencies. For example, practitioner researchers could appear in both academic and professional groups to due their hybrid roles as practicing researchers. Similarly, science communicators could include individuals, bots, and media outlets, making clear boundaries difficult. Some user roles (e.g., cyborgs in Haustein (2019)) were excluded due to limited representation. Overlaps and inconsistencies were mitigated through structured reclassification and separate treatment of ambiguous or hybrid user types (such as in Haustein (2019) and Didegah et al. (2018)). See the established structure at study level in Table 5 (Section 4.2).

Only three key metrics were analyzed across studies: count/proportion of Twitter users; count/proportion of tweets; and count/proportion of tweeted publications. These were derived from user-linked count data, either directly from studies or calculated from reported values. Other metrics, such as follower counts or median tweets, were excluded due to insufficient data across studies.

Additional variables were extracted to assess presence of systematics sources of heterogeneity and are used as binary variables in a beta-binomial hierarchical model to account for the variability across studies. These variables included details such as the method of coding Twitter users (manual or automated), number of coders in the manually coded studies, sample size of each subject of measurement (count of publications, tweets, and Twitter users), information about the tweeted publications (subject field, source of publication including database and journal, publication year, and geographic location of author), and tweets (collection date, types of tweets, and their content). Table 2 provides a list of these factors and their subgroups with definitions.

Table 2. List and definition of other variables examined in BBHM meta-analysis

Factors	Subgroups	Definition
Twitter user coding method	Manual	Studies with manual Twitter user categorization
	Automated/ manual-automated	Studies with automated Twitter user categorization
No. of user categories	≤4 categories	Studies with broad and not more than four user categories
	4 < categories	Studies offering over four or detailed Twitter user categorization
Sample size	<1,000	Sample sizes of tweeted publications or Twitter users below 1,000
	1,000 <	Sample sizes of tweeted publications or Twitter users above 1,000
	<10,000	Sample sizes of tweets below 10,000
	10,000 <	Sample sizes of tweets above 10,000
Activity level	High activity	Studies on highly tweeting users, or highly tweeted publications or fields
	Mixed	Studies covering a mix of data sets or not specifying activity level
Number of fields	Single subject	Studies on publications in one subject field
	Multiple subject fields	Studies on publications in more than one subject field
	Not specified	Studies on publications without subject field specification
	Multiple subject fields + unspecified	Studies on publications in more than one subject field or without subject field specification
Locations	One country/institution	Studies on publications of authors from one country or institution
	Not specified	Studies on publications without author location specification
Number of publication years	One year	Studies on tweeted articles published in a single year
	A range of years	Studies on tweeted articles published in a range of years
Twitter data collection year	2013–2015	The year Twitter mentions to publications have been collected
	2016–2018	
	2019–2021	

3.3. Synthesis Method

The synthesis aimed to identify cross-study patterns in the involvement of different Twitter user types in disseminating scholarly content. The proportions of Twitter users, tweets, and tweeted publications served as the primary outcome variables. Extracted metrics were standardized by recalculating proportions based on reported user counts relative to sample sizes. When studies presented results across multiple topics (Zhou & Na, 2019) or time periods (Xia, Su et al., 2016), a single representative metric per study was retained to avoid bias. In cases where fine-grained user categories needed merging, values were combined to align with our recategorized classification scheme. Figure 2 illustrates the process used to calculate the percentages of tweets and Twitter users.

When studies used differing data sets or coding schemes (e.g., Abhari, Vincent et al., 2022; Dambanemuya, Abhari et al., 2022; Ye et al., 2022), they were treated as separate data sets.

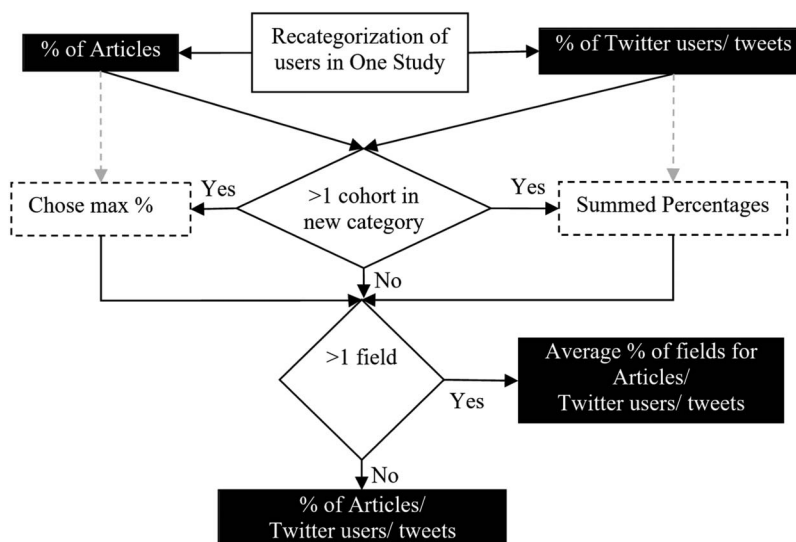


Figure 2. The process scheme for choosing and calculating the statistics based on the new user categorization.

Some studies reported proportions of tweeted articles (such as Didegah et al., 2018; Maleki, 2016), while others reported on the proportion of all articles, including articles with zero tweets (Maleki, 2018; Xia et al., 2016). For tweeted publications, only records with at least one tweet were analyzed. Zero values were not imputed for missing categories.

Homogeneity test: To assess statistical consistency, Cochran's Q test and I^2 statistics were used. A p -value < 0.05 and $I^2 > 50\%$ were interpreted as indicators of substantial heterogeneity, justifying the use of random effects models (Mikolajewicz & Komarova, 2019). H^2 was calculated using $H^2 = Q_{\text{total}}/df$, and I^2 using $I^2 = (Q - df)/Q$.

3.4. Meta-Analysis Approach

3.4.1. Random effects model (REM)

A random effects model using restricted maximum likelihood (ReML) estimation was used to calculate pooled estimates (Bartlett, 1937). Random effect models is assumed to represent a random sample from a particular distribution of the effect sizes and assist in controlling for unobserved heterogeneity (Wooldridge, 2010). REM has historical relevance in meta-analysis studies and could be useful if future research needs to reference its estimates.

Study weights, w_i , in REM are calculated using the inverse variances derived from standard error plus between-study variance or heterogeneity, τ^2 (Mikolajewicz & Komarova, 2019). Because most of the earlier studies did not provide the standard deviation, the standard error, $se(\theta_i)$, is used, which is approximately the inverse root of the study-level sample size, n_i . Study weights (w_i) were derived using:

$$se(\theta_i) \sim \frac{1}{\sqrt{n_i}}$$

$$w_i = \frac{1}{se(\theta_i)^2 + \tau^2}$$

In the current study, except for the standard error, other measures such as study weights, and confidence intervals were derived and illustrated using SPSS version 28.0.0.0 (IBM, 2021). All

the meta-analysis estimates were made with a 95% confidence interval. Forest plots were used for visualizing study estimates and their confidence intervals.

Meta-regression was initially considered to evaluate the impact of study-level variables. However, due to limited data and sparse subgroup representation, it yielded unstable and nongeneralizable estimates. Consequently, metaregression was not pursued. Instead, binary-coded study-level variables were embedded as weakly informative priors in a Bayesian hierarchical model. This allowed us to incorporate potential sources of heterogeneity without inflating error margins.

Traditional REM and proportion-weighted approaches are constrained by their sensitivity to large-sample outliers and sparse category coverage. Weighted means, for example, disproportionately favor studies like Yu (2017), which reported over 5 million tweeted publications, and can obscure meaningful variation from smaller studies. BBHM addresses these limitations by modeling engagement as a probabilistic distribution rather than a fixed point estimate. It accommodates between-study variance through hierarchical priors and fully propagates uncertainty, which is particularly critical when category representation is low. It also treats the engagement outcomes as count data, making it compatible with the binomial structure of the problem. This model is ideal for scenarios like ours, where user classification is nonexhaustive, overlapping, or not mutually exclusive, and sometimes ambiguous. BBHM simultaneously estimates the probability of a user being assigned to a category and not assigned to that category, enabling clearer insights into classification dynamics.

3.4.2. Beta-binomial hierarchical model (BBHM)

To address high heterogeneity, variable sample sizes, and sparse studies, we employed a beta-binomial hierarchical model (BBHM).

Let x_i be the number of classified users in a category and n_i be the total sample size in study i . The BBHM approach allows engagement outcomes to be modeled as binomial events:

$$x_i \sim \text{binomial}(n_i, p_i)$$

where p_i represents the study-specific user engagement probability and is modeled as a beta-distributed random variable:

$$p_i \sim \text{beta}(\alpha_i, \beta_i)$$

Here, α reflects support for user identification in a given user category (“in-category”), while β_i represents exclusion (“out-of-category”). In this way, BBHM provides a natural framework for interpreting both the likelihood of a user being classified into a specific category and the complementary probability of not being classified into that category. This modeling strategy captures uncertainty and variability in user classification performance across studies.

At the user category level, the BBHM yields a posterior estimate ρ , defined as the mean of the beta distribution from which each study-specific probability p_i is drawn. This value represents the overall expected engagement proportion for a given user category across all studies, accounting for between-study variability. In other words, $\rho = \alpha/(\alpha + \beta)$ reflects the model’s best estimate of how frequently users in that category engage with scholarly publications on Twitter. This estimate, along with its highest density interval (HDI), offers a probabilistic summary of category-level engagement based on all available evidence.

BBHM was implemented in Python using PyMC (draws = 100,000, chains = 4 and target_accept = 0.95) with other binary variables given as weakly informative priors and four

Markov chain Monte Carlo (MCMC) chains of 100,000 draws. Monte Carlo standard error (MCSE) was used to assess sampling precision. Posterior estimates were interpreted as probabilistic engagement rates, allowing category-wise comparisons under a unified modeling structure.

3.5. Mean Difference Between User Categories

To assess differences in engagement across user categories, one-tailed and two-tailed *t*-tests were performed on study-level proportions. This allowed for comparative evaluation of academic individuals versus other groups across all three engagement metrics.

3.6. Certainty of Evidence Assessment

Following GRADE guidelines, results were rated for certainty based on risk of bias, imprecision, inconsistency, indirectness, and publication bias (Schünemann et al., 2013). Automated coding studies were considered less reliable than manual ones, and mixed user categories were rated lower due to definitional ambiguity. Imprecision was judged based on confidence interval width and the number of contributing studies. Inconsistency was measured via heterogeneity indicators, and indirectness by relevance of categories to the core research question. Publication bias was assessed qualitatively.

4. FINDINGS

4.1. Study Selection

A comprehensive search was conducted across major academic databases, publishers' websites, and the *Altmetric.org* platform, yielding 3,760 initial records. After removing 1,138 duplicates, 2,622 studies remained for screening. Of these, 1,429 were unrelated to altmetrics, 910 were not relevant to the research question, and 140 were non-English. After applying inclusion criteria, 65 studies were selected for data extraction.

From these, 23 studies contained sufficient quantitative information on Twitter user categorization to be included in the meta-analysis (Figure 3). Several studies were excluded for methodological or conceptual reasons. These were studies that focused on only a single user group (e.g., scientists, as in Holmberg and Thelwall (2014) and Ke et al. (2017) or journal accounts, as in Kwak and Lee (2014) and Said, Bowman et al. (2019)), or those that used incompatible categorization schemes, such as merging individuals and organizations (e.g., Pandian et al., 2019), or applying unusual identification methods like scientists' name-matching (e.g., Joubert & Costas, 2019). Some data sets exhibited high statistical heterogeneity or presented noncomparable results and were only retained after normalization or excluded entirely.

4.2. Study Descriptions

The final 23 studies reflected a diverse set of research designs, subject areas, and temporal coverage. Nine studies focused on specific topics, such as climate change (Toupin et al., 2019) or COVID-19 (Ye et al., 2022), while psychology emerged as the most frequently studied field (Htoo & Jin-Cheon, 2017; Udayakumar, Senadeera et al., 2018; Zhou & Na, 2019). Some studies were geographically bounded, focusing on publications from Finland (Vainio & Holmberg, 2017), Iran (Maleki, 2014), or Brazil (de Melo Maricato & de Castro Manso, 2022).

A few studies analyzed institutional or journal-based tweeting patterns (e.g., Tsou et al., 2015), while others examined specific data sets such as arXiv (Haustein et al., 2016a) or

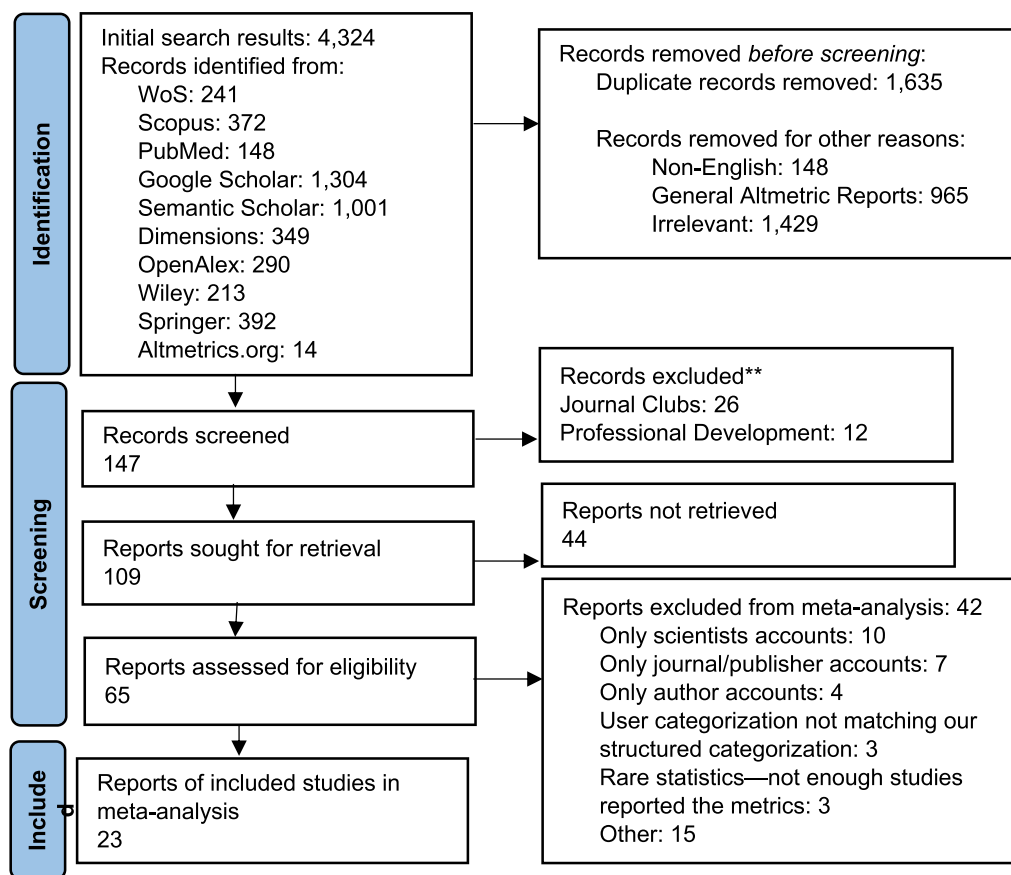


Figure 3. Flow diagram showing the results of the literature search and the screening of Twitter users studies in Altmetrics and their selection process for systematic review and meta-analysis.

university repositories (Sergiadis, 2018). Two large-scale studies (Haustein, 2019; Yu et al., 2019) analyzed millions of tweets without topical or regional restrictions.

Seven studies targeted high-activity zones on Twitter, including trending disciplines (e.g., astronomy and COVID-19), journals (e.g., *Nature*), or prolific users. While these studies offered rich data sets, their focus on highly active profiles (Haustein, 2019; Haustein et al., 2016a; Yu et al., 2019) may have introduced sampling biases, resulting in underrepresentation of lesser-active user types. Table 3 provides an overview of study-level characteristics, while excluded studies are detailed in Table 4 in accordance with PRISMA guidelines.

4.3. Study-Level Outcome Measures

Table 5 presents the results from the earlier studies that were included for the meta-analysis. The Twitter user types presented in these were examined across three metrics: Twitter users (15 studies); tweets (11); and tweeted publications (8). As some of the included studies presented more than one type of metric, the total number of studies exceeds the number of publications (23) included. Table 5 shows the counts and percentages of the three metrics for each of the proposed Twitter user categories at function level. However, other outcome measures that were reported in the earlier studies, such as count, percent, mean, median, geometric mean, minimum, maximum, and different percentiles of tweets (12 studies), Twitter followers (9),

Table 3. The data characteristics of the earlier studies included in the meta-analysis

ID	Study	(# of) Coder	Twitter user categorization	User category-based metrics	Data set			Publication dates	Data collection date	Subject field	Geography
					Twitter users	Tweets	Publications				
1	Maleki (2014)	One	Manual (10 cohorts)	Tweeted articles	565	1,096	589	1997–2012	July 2013	5 broad fields	Iran
				Tweets							
2	Tsou et al. (2015)	Two	Manual (12 cohorts)	Twitter users	2,000 (500 per journal)	–	–	–	March 2012– March 2013	9 fields	–
3	Haustein et al. (2016a)	Two	Manual (3 cohorts)	Tweets	51 accounts (mentioned arXiv in profile & tweeted)	50,068	–	2012	May 2014	–	–
4	Haustein et al. (2016b)	NR	Manual (22 cohorts)	Twitter users	800 (200 per activity level)	663,547	–	2012	June 2015	–	–
5	Maleki (2016)	Altmetric .com	4 cohorts	Tweeted articles	–	–	8,978 articles	2011–2014	August 2016	–	–
6	Xia et al. (2016)	Altmetric .com	4 cohorts	Articles	–	–	4,276—has duplicates	2010–2015	June 2015	4 fields in Nature	–
7	Htoo and Jin-Cheon (2017)	Two	Manual (4 cohorts and 27 subtypes)	Tweets	1,452	2,016	133 highly tweeted articles	2013	NR	Psychology	–
				Twitter users							
8	Vainio and Holmberg (2017)	Two	Manual (16 + 1 ^a cohorts)	Twitter users	400 (100 per field)	–	20 articles (top 5/field)	2014–2016	2015	4 broad fields	Finland
9	Yu (2017)	Altmetric .com	4 cohorts	Tweeted articles	–	21,255,678	5,184,525	2011–2016	2016	27 fields	–
				Tweets							
10	Didegah et al. (2018)	Three	Manual (14 cohorts)	Tweeted articles	–	6,388	300 (60 per field)	2014	June 2016	5 fields	–
				Tweets							
11	Maleki (2018)	Altmetric .com	4 cohorts	Tweeted Articles	31,051	150,893	29,543 articles	2011–2017	2018	Astrophysics	–
12	Sergiadis (2018)	One	Manual (6 cohorts)	Tweets	–	605	290	1978–2016 Mostly in 2012–2016	October 2016	8 fields	Georgia Southern University
13	Udayakumar et al. (2018)	Two	Manual and Classifier (6 cohorts)	Twitter users	3,000	95,088	37,627	2014–2015	NR	Psychology	–
14	Haustein (2019)	One	Manual (2 + 1 ^a cohorts)	Tweets	2,043 users tweeting ≥ 1,000	24.3 million	3.9 million	2015	June 2016	–	–
				Twitter users							
15	Toupin et al. (2019)	Two	Manual (7 cohorts)	Twitter users	17,837 of 21,965	41,019	2,620 of 4,719	2015–2016	NR	Climate change	–
			Semi-automatic								
16	Yu et al. (2019)	Three	Manual (10 + 17 cohorts)	Twitter users	1,468 (about 500 per activity level in 2016)	88,407 of sampled tweeters	–	–	October 2011– June 2016	–	–

Who are tweeting about academic publications?

17	Zhou and Na (2019)	Two	Manual (3 × 2 cohorts) and Supervised Learning (15 cohorts)	Twitter users	6,000 (3,000 per field) 33,019 in the test set (10,478 in Politics and 22,541 in Psychology)	39,532 in Politics and 252,310 in Psychology	8,265 in Politics and 42,891 in Psychology	2011–2015	March 2017	Psychology & Political Science	–
18	Haustein and Peters (2020)	Two	Manual (11 + 1 ^a cohorts)	Twitter users	464	577	25 of 131	–	October 2018, August 2019	Judit Bar-Ilan's Publications	–
19	Lemke, Bräuer, and Peters (2021)	One	Manual (12 cohorts)	Twitter users Tweets Twitter followers	50 of 1,808	1,161 of 3,677	–	–	July–October 2020	German nuclear repository	Germany
20	Orduña-Malea and Costas (2021)	NR	Manual (4 cohorts)	Twitter users	1,376 (200 originally tweeting, 748 liking, 728 retweeting)	–	–	–	2019	VOSviewer software's website	–
21	Abhari et al. (2022)	Two	Manual and Rule-based classifier (5 cohorts)	Twitter users Tweets	26,260 (20,195 retracted + 6,065 nonretracted)	93,127 (67,124 retracted + 25,997 nonretracted)	3,847 retracted 2,085 nonretracted	2011–2019	September 2021	Retracted articles	–
22	de Melo Maricato and de Castro Manso (2022)	NR	Manual (4 cohorts and 36 + 1 user occupation)	Twitter users Tweeted articles Tweets	2,435	3,653	877	–	2020	22 fields	University of Brasília
23	Ye et al. (2022)	Two	Manual (7 cohorts)	Twitter users Tweeted articles Tweets	659 of 99,619	13,179 of 151,480	417 of 1,252	May 2020	June 2020	COVID-19	–

NR: Not reported; ^a +1 means one category for all the remaining unclassified users; – in Geography: No geographical restriction.

Table 4. Data characteristics of the earlier studies that were excluded from the meta-analysis

ID	Study	Exclusion reason	# of) Coder	Twitter user categorization	User category-based metrics	Data set		Publications/Authors	Publication dates	Data collection date	Subject field
						Twitter users	Tweets				
1	Holmberg and Thelwall (2014)	One-cohort study	No Twitter user coding	One (researchers)	Tweets	447 (ranging between 24–52 across fields)	2,000 (200 per field)	–	–	March–October 2012	10 fields
2	Holmberg, Bowman et al. (2014) ^a	Not only tweeted articles, @user mentions in tweet content	One	Manual (12 cohorts)	Twitter users	518 of 11,252	27,923 of 56,415	32 of 37 astrophysicists	–	May 2013	Astrophysics
3	Kwak and Lee (2014)	One-cohort study	No Twitter user coding	One (journals)	Journals holding an account Twitter followers	57,699	136,139	31,661 articles in top 100 journals	–	January–June 2013	54 domains
4	Bowman (2015)	One-cohort study	No Twitter user coding	One (Professors surveyed)	Scientists holding an account Personal / Professional usage	445 users	–	1910 authors	–	May 2014	–
5	Costas, van Honk, and Franssen (2017)	One-cohort study	No Twitter user coding	One (Authors disambiguated)	Scientists holding a Twitter account	334,856	2,622,116	17,332,510 scholars	1980–2015	April 2016	4 broad fields
6	Hughes, Hughes, and Murphy (2017)	One-cohort study	No Twitter user coding	One (journals)	Journals holding a Twitter account	–	–	50 top journals	–	June 2017	Trauma and orthopaedic surgery
7	Ke et al. (2017)	One-cohort study	No Twitter user coding	One (Scientists)	Scientists holding a Twitter account Tweets Twitter followers	45,867 of 110,708	64,449,234	894,840 scientists in employment	–	May 2014 (employment data)	24 fields 5 broad fields
8	Ortega (2017)	One-cohort study	Targeted groups	3 cohorts of Publisher, Journal, and Owner	Articles	–	–	4,176 from 350 journals	2013	April 2016	5 broad fields
9	Wong, Piraquive, and Levi (2018)	One-cohort study	No Twitter user coding	One (journals)	Journals holding a Twitter account Tweets Twitter followers	18	–	50 journals	2008–2016	NR	Otolaryngology
10	Côté and Darling (2018)	One-cohort study	No Twitter user coding	One (Scientists)	Twitter followers	110 of 167	–	200 researchers	2018	2015	Ecology and evolutionary biology
11	Tur-Viñes, Segarra-Saavedra, and Hidalgo-Marí (2018)	One-cohort study	No Twitter user coding	One (journals)	Journals holding a Twitter account Twitter followers	30	–	70 journals	–	2016	Communication
12	Hassan, Bowman et al. (2019)	Rare metrics	Altmetric.com	4 cohorts	Classification score for predicting highly cited articles	–	–	–	2011–2015	2017	LIS

Who are tweeting about academic publications?

13	Joubert and Costas (2019)	One-cohort study	One rule-based algorithm	WoS disambiguated list of authors matched on Twitter—> Low recall	—	14,174 of 2,622,116	—	—	2010–2016	June 2016	—
14	Pandian et al. (2019) ^a	Unfit for recategorization	Two	Manual (2 cohorts)	Twitter users	2,064	—	2078	2017	NR	Psychology
15	Said et al. (2019)	One-cohort study	No Twitter user coding	One (journals)	Twitter Network PageRank and Eigenvector	19 journals and 149,830	—	77,757	2015	Publications (2017) Tweets (2016)	—
16	Yadav, Nagarkatti et al. (2019)	One-cohort study	No Twitter user coding	One (journal editorial board)	Editors holding a Twitter account	—	—	240 editorial members of 6 high IF journals	—	July–August 2018	Obstetrics and gynecology
17	Coret et al. (2020)	One-cohort study	No Twitter user coding	One (surgeons)	Scientists holding an account Tweets/Tweet Likes Twitter followers/friends	188	—	3,741 surgeons	—	2018	General thoracic surgery
18	Hassan, Aljohani et al. (2020)	Rare metrics	Altmetric.com	4 cohorts	Tweet sentiments	—	5,341,800	—	2011–2016	2016	—
19	Clavier, Occhiali et al. (2021)	One-cohort study	No Twitter user coding, but academics	Two (Academic ranks)	Scientists holding an account Tweets/followers	68 (42 active)	—	162 authors	2016–2020	March 2021	Anaesthesia
20	Haunschild, Bormann et al. (2021) ^b	Irregular bot statistics	One R package	Automated (tweetbotmot)	Twitter users	56,266	173,187	6,433	2011–2019	NR	Opioid
21	Khandelwal and Tagat (2021)	One-cohort study	No Twitter user coding	One (Academics)	—	131	56,512	—	NR	January–February 2020	Development research
22	Sanchez (2021)	One-cohort study	No Twitter user coding, but academics	Three (Academic ranks)	Scientists holding an account Age of accounts Tweets/followers/friends	322	—	1,107 scientists	2019	March 2007–April 2019	Urban planning
23	Wang and Zuccala (2021)	One-cohort study	No Twitter user coding	One (book publishers)	Publishers holding one or multiple Twitter accounts	—	—	6,258 of 15,454 unique book titles	2014–2018	April and May 2019	—
24	Zhang and Wang (2021) ^c	Unfit for recategorization	NR	Two × 2 (official and popular)	Tweeted articles Tweets	—	286,650	20,310 in 6 journals	2009–2019	July 2019	Journals published by the Royal Society

NR: Not Reported; WoS: Web of Science; BKCI: Book Citation Index; IF: Impact Factor; LIS: Library and Information Science.

^a The statistics on Academics were a mix of individuals and organizations.

^b The findings introduced significant bias to *Bot Science Communicators'* meta-analysis.

^c The definition of cohorts did not fit with the new categorization.

Table 5. Outcome measures (percentage values and counts) for Twitter users, tweets, and tweeted publications organized by function-level user categories

Study	Unit	Individual			Organization			Science communicator				Mixed groups	Sample size
	Function	Academic	Professional	Other/ mixed	Academic	Professional	Other/ mixed	Academic	Professional	Bot	Other/mixed		
Twitter users													
Tsou et al. (2015)		26.2%		49.9%	3%	6.7%	8.9%		4.4%			1.1%	100%
		523		997	60	134	178		87			21	2,000
Haustein et al. (2016b)		16.7%	36.1%	5.5%	11.9%	3.2%	5.0%	0.1%	1.2%	8.0%	9.5%	11.7%	100%
		134	288	44	95	25	40	1	10	64	76	93	800
Htoo and Jin-Cheon (2017)		27.5%		55.2%	7.9%		9.4%						100%
		400		801	114		137						1,452
Vainio and Holmberg (2017)		23.8%	21.5%	2.0%	2.8%	9.5%	5.0%	1.3%	7.8%		0.8%	25.8%	100%
		95	86	8	11	38	20	5	31		3	103	400
Udayakumar et al. (2018)		21.4%		46.6%			32%						100%
		641		1,399			960						3,000
Haustein (2019)									14.9%	12.1%		72.9%	100%
									305	248		1,490	2,043
Toupin et al. (2019)		31.7%	2.7%	1.2% ^a			19.5%		12.7%	1.9%		30.4%	100%
		5,651	473	216			3,475		2,265	335		5,422	17,837
Yu et al. (2019)		44.9%	15.8%	22.2%	2.3%	9.1%	16%	2.5%	3.2%	1.2%	0.6%		— ^b
		659	232	326	34	134	235	37	47	18	9		1,468
Zhou and Na (2019)		30.8%		45.7%	7.2%		16.5%						100%
		9,073		14,899	2,489		6,558						33,019
Haustein and Peters (2020)		39.4%	13.8%		5.8%		7.1%	17.0%	26.3%	1.5%		20.5%	— ^b
		183	64		27		33	79	122	7		95	464
Lemke et al. (2021)				44%			54%					2%	100%
				22			27					1	50
Orduña-Malea and Costas (2021)				86.5%			11.2%					2.3%	100%
				1,207			138					31	1,376
Abhari et al. (2022) Retracted		54.3%	17.6%							2.8%	12.3%	13.0%	100%
		5,166	1,674							266	1,170	1,237	9,514

Who are tweeting about academic publications?

Abhari et al. (2022)	54.6%	15.6%					5.3%	20.3%	4.2%	100%	
Not retracted	1,362	389					132	506	105	2,495	
de Melo Maricato and de Castro Manso (2022)	29.6%	12.4%	26.0%		25.2%		1.4%	4.0%	1.5%	100%	
	719	306	633		614		32	97	36	2,437	
Ye et al. (2022)	22.2% ^c	27.3%					5.5%	13.2%	5.2%	39.9%	100%
	146	180					36	87 ^a	34	263	659
Tweets											
Maleki (2014)	13.0%	17.0%	22.0%	11.0%	3.0%		11.0%		23.0%		100%
	110	189	248	137	30		126		256		1,096
Haustein et al. (2016a)									9%	91%	100%
									~4,506	~45,562	50,068
Htoo and Jin-Cheon (2017)	29.2%		47.6%	14.2%		9.0%					100%
	577		937	280		177					1,971
Yu (2017)	16.5%	7.5%							4.9%	71.1%	100%
	3,503,640	1,587,158							1,041,701	15,108,440	21,255,678
Didegah et al. (2018)	21.33%	10.7%	34.7%	1.5%	20.5%	0.1%	5.2%	21.1%	0.04%	6.0%	100%
	1,363	685	2,214	93	1,308	6	333	1,336 ^c	3 ^d	384	6,388
Sergiadis (2018)	14.2%		51.7%		11.2%	22.8%					100%
	86		313		68	138					605
Haustein (2019)							3%	7%		90%	100%
							0.73 million	1.7 million		21.9 million	24.3 million
Lemke et al. (2021)			54.3%			45.0%				0.7%	100%
			631			522				8	1161
Abhari et al. (2022)	18.2%	7.7%	23.8%					8.3%	5.3%	36.7%	100%
Retracted	4,933	2,079	6,435					2,250	1,427	9,949	27,073
Abhari et al. (2022)	35.2%	6.3%	6.5%					4.9%	21.1%	25.9%	100%
Not retracted	4,050	729	750					564	2,723	2,984	11,500

Table 5. (continued)

Study	Unit	Individual			Organization			Science communicator				Mixed groups	Sample size
	Function	Academic	Professional	Other/mixed	Academic	Professional	Other/mixed	Academic	Professional	Bot	Other/mixed		
de Melo Maricato and de Castro Manso (2022)		29.6%	12.3%	26.0%			25.1%		1.3%		4.2%	1.5%	100%
		723	308	638			617		32		102	36	2,456
Ye et al. (2022)		19.8%	26.7%						6.1%	17.6%	6.8%	40.6%	100%
		2,613	3,523						801	2,320 ^c	896	5,346	13,179
Tweeted publications													
Maleki (2014)		11.4%	19.4%	3.4%	16.6%	4.1%			17.7%		31.9%	20.9%	— ^b
		67	114	20	98	24			104		188	123	589
Maleki (2016)		69.6%	49.1%								49.3%	97.5%	— ^b
		6,251	4,407								4,426	8,755	8,978
Xia et al. (2016) ^e		83.9%	26.4%								47.2%	90.6%	— ^b
		3,100	975								1,745	3,348	~3,694 ^e
Yu (2017)		33.1%	16.2%								16.4%	85.1%	— ^b
		1,292,686	631,927								641,232	3,323,066	5,184,525
Didegah et al. (2018)		31.8%	6.8%	13.4%	4.0%	17.2%			4.8%	32.0%		10.7%	— ^b
		74	20	40	12	56			43	96		32	300
Maleki (2018)		50.4%	2.2%								25.1%	78.5%	— ^b
		14,877	649								7,423	23,196	29,543
de Melo Maricato and de Castro Manso (2022)		30.7%	16.2%	70.8%			46.3%		1.9%		6.8%	3.9%	— ^b
		430	226	621			406		26		60	34	877
Ye et al. (2022) ^f		6.72%	2.6%	13.2%					9.6%	5.3%		62.6%	— ^b
		28	11	55					40	22		261	417

^a Calculated based on the proportion of “personal” accounts not overlapping with any other category.^b Overlapping categories (over 100%).^c This study is imprecise because it might include academic institutions.^d The study identifies bots and overlaps with other user categories (excluded when summing up user categories to get the sample size).^e This study reported statistics including papers with zero tweets. The statistics are estimated by the author based on articles with nonzero tweets.^f Only if articles were originally tweeted by the cohort.

Twitter users (5), Twitter friends (4), and similar metrics were not analyzed in this study due to too low a number of studies to perform a meta-analysis with.

4.4. Synthesis by Random Effect Model (REM)

Figures 4–9 show the results of 42 meta-analyses (3 variables $\times$ 11 function-categories (–3 missing due to too few studies) + 3 variables $\times$ 4 unit-categories).

4.4.1. Meta-analysis of Twitter users

Fifteen studies contributed 138 outcome measures as user counts: 48 at the user unit level and 90 at the function level. From a total of 79,014 Twitter accounts, individuals were the dominant contributors to scholarly tweeting, with an average of 66% (95% CI: 59–74%, $I^2 = 98.9\%$)

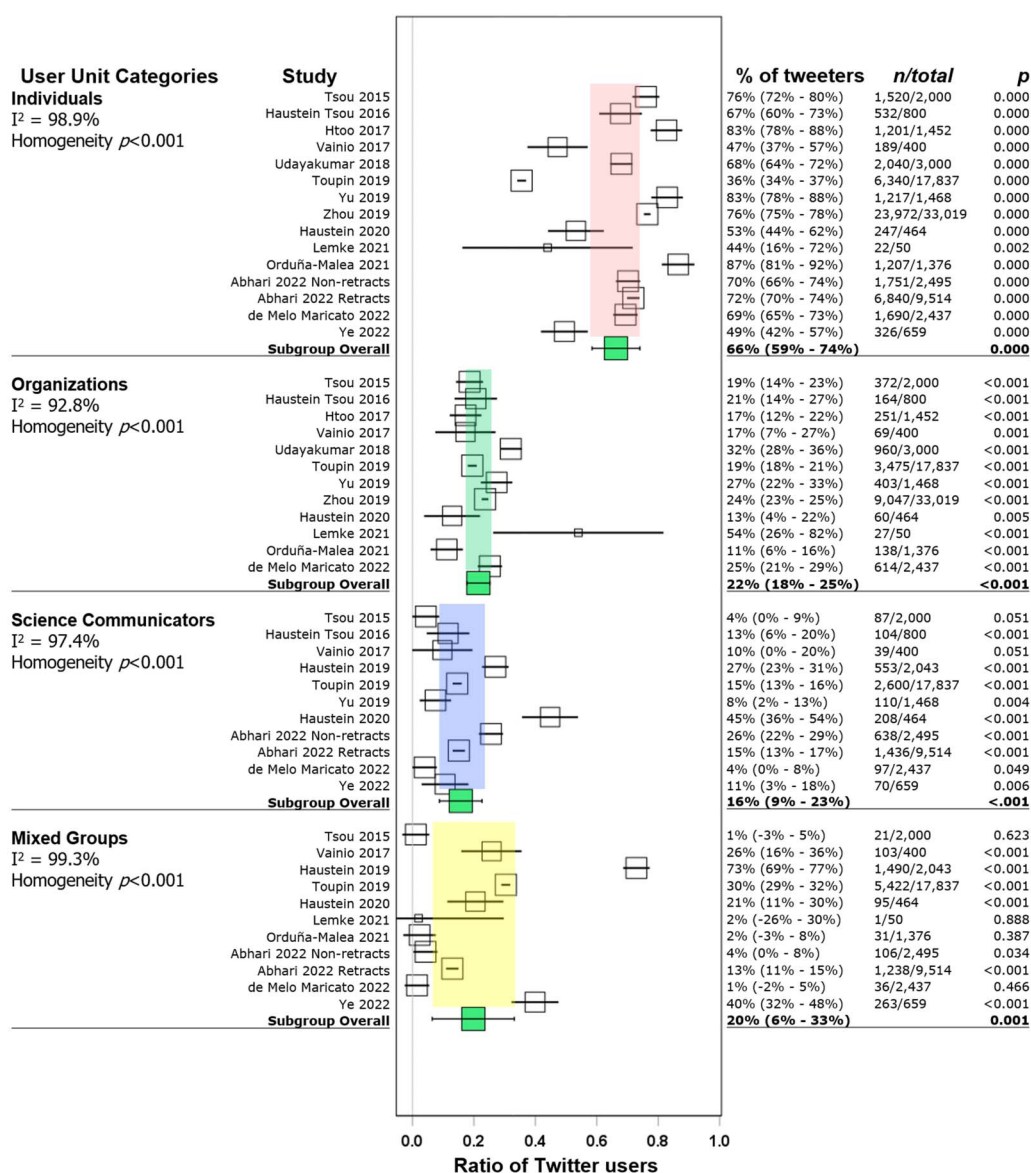


Figure 4. Meta-analysis of proportion of users at User Unit level. The colored ribbons identify the (non)overlapping studies with the confidence intervals of the overall estimates for each user category.

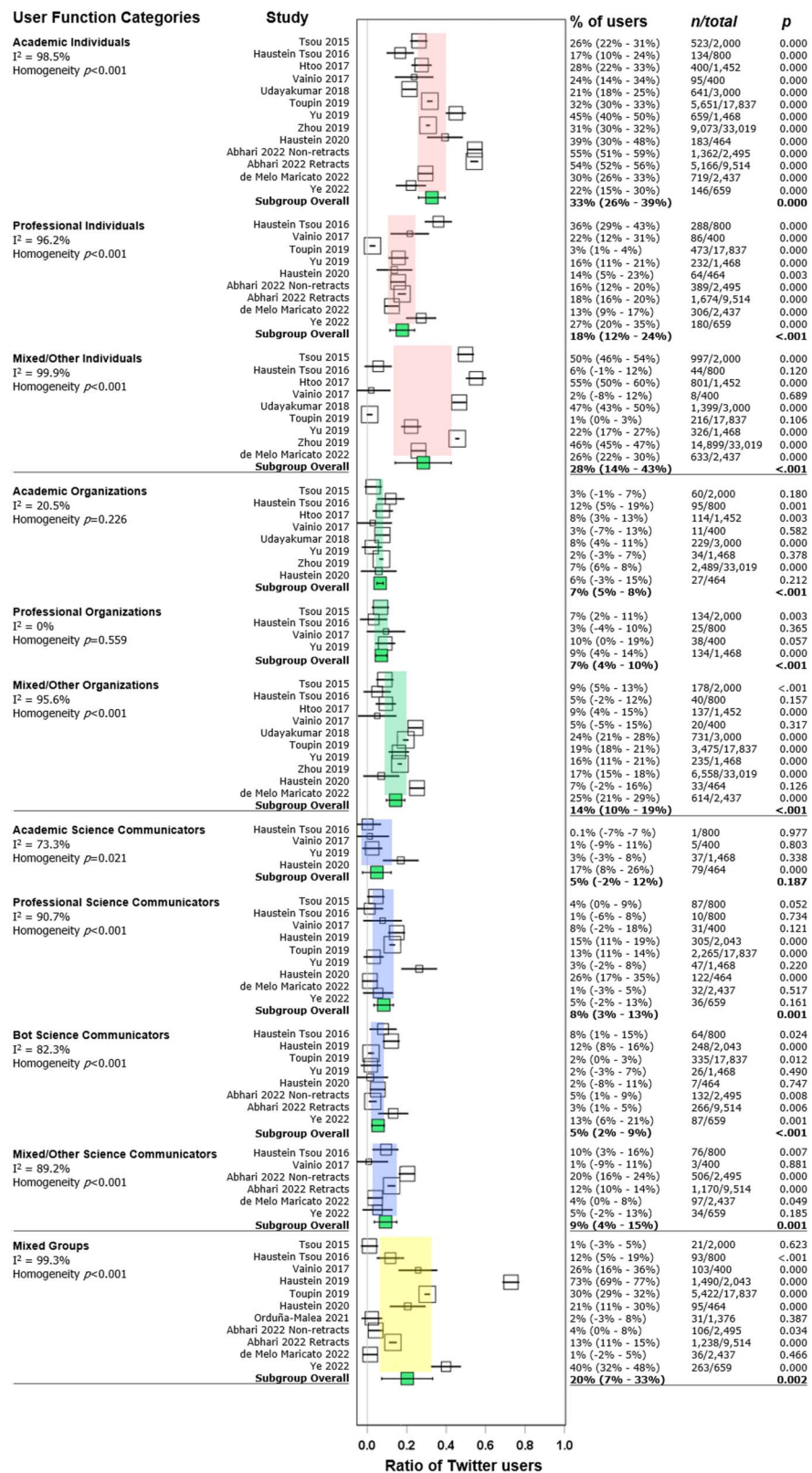


Figure 5. Meta-analysis of proportion of Twitter users at User Function level. The colored ribbons identify the (non)overlapping studies with the confidence intervals of the overall estimates for each user category.

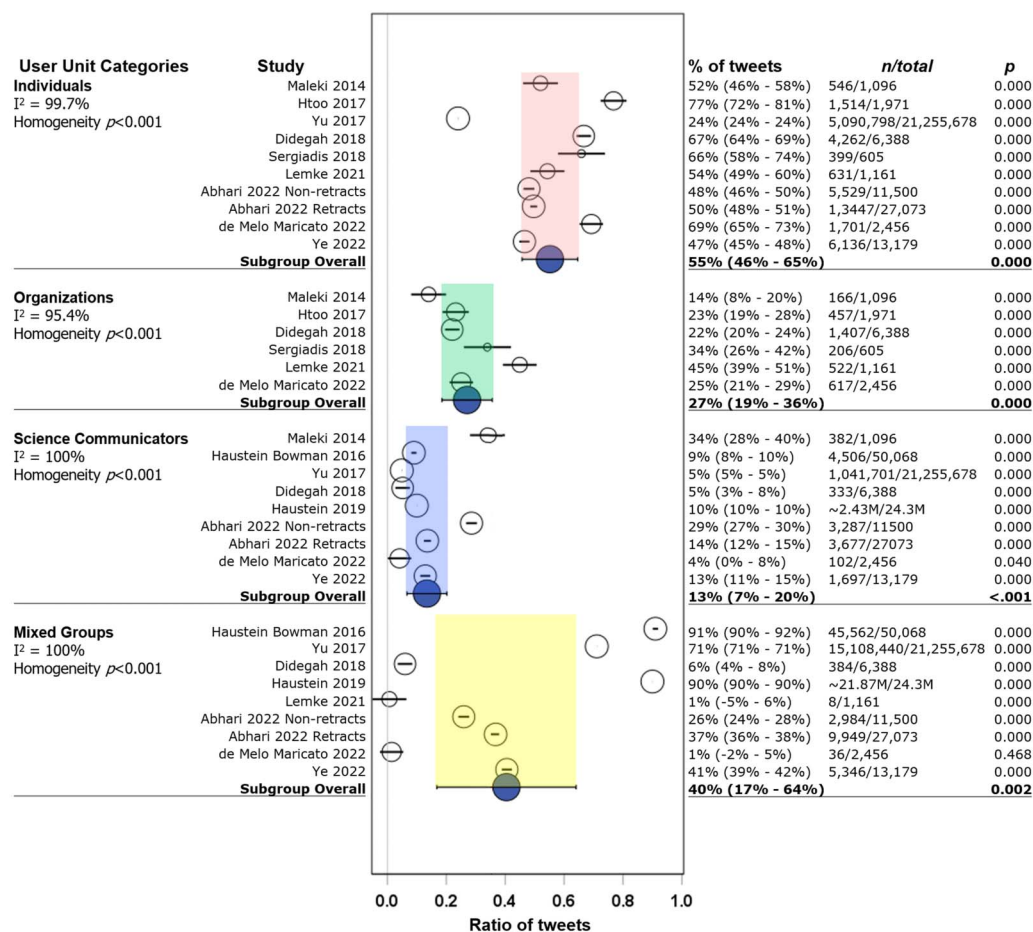


Figure 6. Meta-analysis of proportion of tweets at User Unit level. The colored ribbons identify the (non)overlapping studies with the confidence intervals of the overall estimates for each user category.

(Figure 4). Organizations followed with 22% (CI: 18–25%, $I^2 = 92.8\%$), and science communicators 16% (CI: 9–23%, $I^2 = 97.4\%$).

At the function level (Figure 5), academic individuals represented 33% of Twitter users (CI: 26–39%, $I^2 = 98.9\%$). Academic and professional organizations each had 7% of users (both $I^2 < 50\%$, high certainty). Other categories showed lower proportions and high heterogeneity (e.g., professional individuals at 18%, mixed individuals at 28% ($I^2 = 99.9\%$), and bot science communicators at 5%).

4.4.2. Meta-analysis of tweets

Eleven studies contributed 92 outcomes, covering a total of 104,112 tweets from nine small-scale studies and over 2 million tweets in two large-scale studies. Individuals produced the largest share of tweets (55%, CI: 46–65%, $I^2 = 99.7\%$), followed by organizations (27%, CI: 19–36%, $I^2 = 95.4\%$) and science communicators (13%, CI: 7–20%, $I^2 = 100\%$) (Figure 6).

At the function level (Figure 7), academic individuals tweeted 22% of the research content (CI: 17–27%, $I^2 = 98.9\%$), and mixed individuals 30% (CI: 19–42%, $I^2 = 99.2\%$). Bot science communicators were below the overall average at 9%, while professional science communicators were the only group with higher certainty (5%, $I^2 = 83.7\%$).

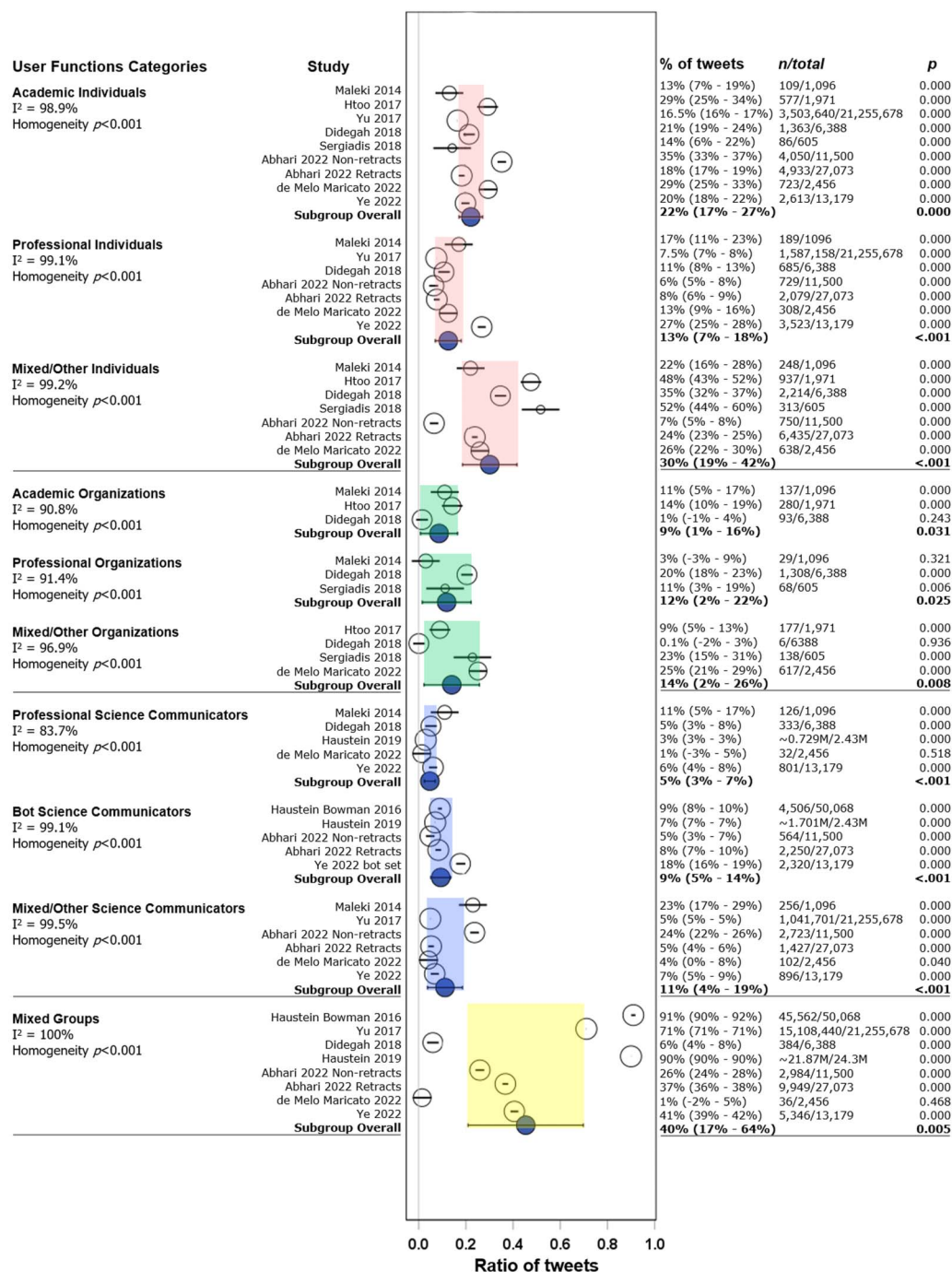


Figure 7. Meta-analysis of proportion of tweets at User Function level. The colored ribbons identify the (non)overlapping studies with the confidence intervals of the overall estimates for each user category.

4.4.3. Meta-analysis of tweeted publications

Eight studies provided 57 outcomes for tweeted publications: 21 at unit level and 36 at function level. From 44,398 publications (plus over 5 million from Yu (2017)), individuals were linked to 50% of tweeted articles (CI: 34–66%, $I^2 = 99.8\%$), organizations 28% (CI: 10–47%, $I^2 = 92.8\%$), and science communicators 30% (CI: 18–42%, $I^2 = 99.7\%$) (Figure 8).

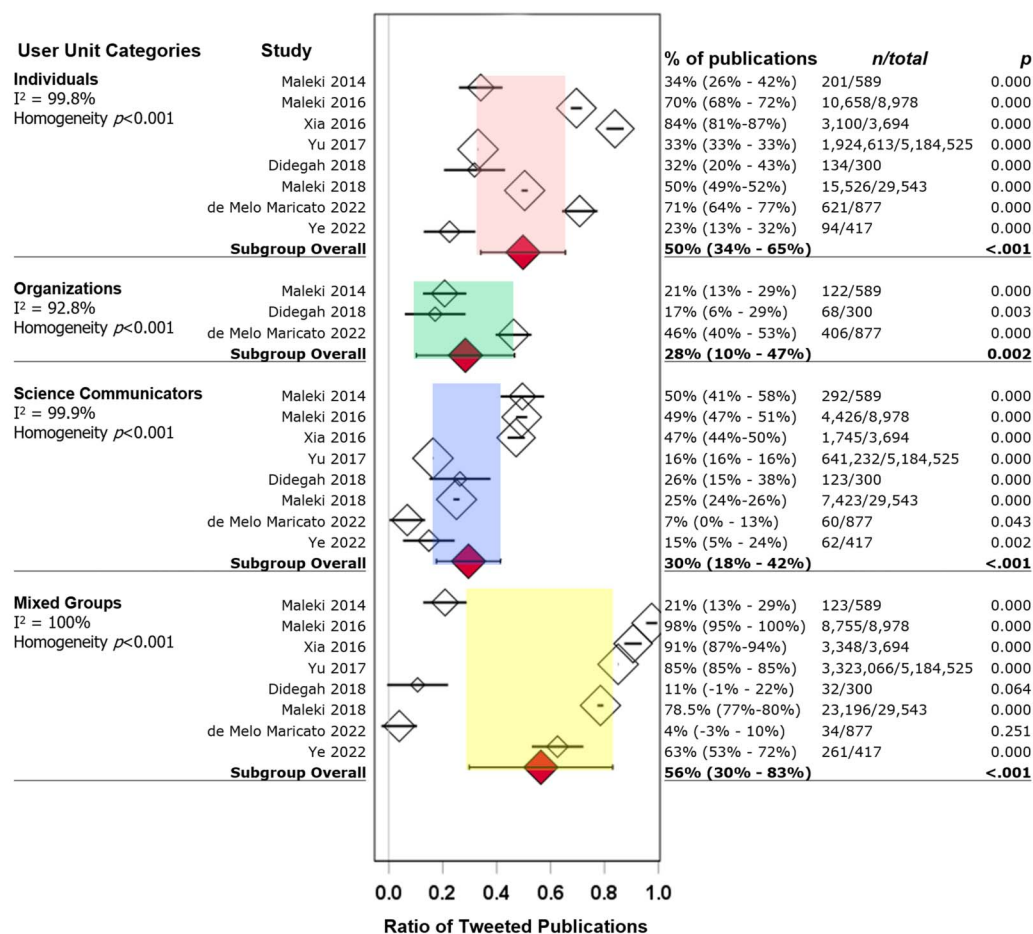


Figure 8. Meta-analysis of proportion of tweeted publications at User Unit level. The colored ribbons identify the (non)overlapping studies with the confidence intervals of the overall estimates for each user category.

At the function level (Figure 9), academic individuals tweeted 42% of publications (CI: 24–61%, $I^2 = 99.9\%$). Professional science communicators tweeted 9% (CI: 2–16%, $I^2 = 61\%$), with other groups showing high heterogeneity and wider uncertainty.

4.5. Synthesis by Beta-Binomial Hierarchical Model (BBHM)

The random effects meta-analyses above showed substantial heterogeneity ($I^2 > 50\%$ in 40 out of 42 variable-user category combinations), suggesting that study-level differences contribute significantly to the variability in results. Our exploratory meta-analyses included only 23 studies with lower study counts (2–13 studies) per category, making traditional random effects models susceptible to unreliable variance estimation. Several studies had wide confidence intervals, especially for proportion of tweeted publications, with extreme heterogeneity in effect sizes (see significance intervals in forest plots, Figures 4–9).

To address these limitations, we implemented a Bayesian hierarchical model (BHM) using a beta-binomial structure. This model allows for robust estimation of between-study variability by incorporating hierarchical priors, mitigating overfitting in small-sample meta-analyses, and accommodating classification uncertainty. The model structure estimates category-specific parameters: α (support for users being in-category) and β (support for users being out-of-

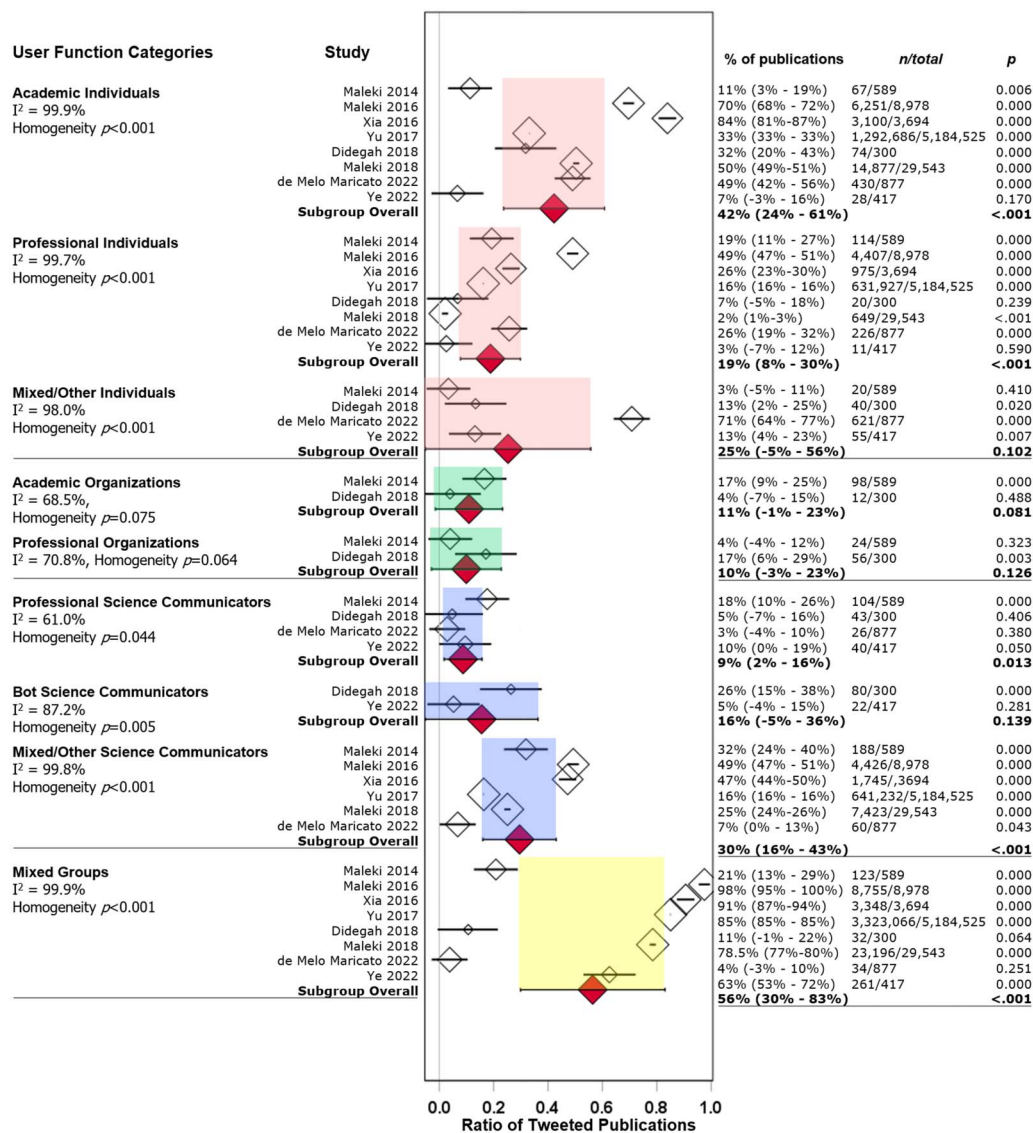


Figure 9. Meta-analysis of proportion of tweeted publications at User Function level. The colored ribbons identify the (non)overlapping studies with the confidence intervals of the overall estimates for each user category.

category), from which the posterior mean engagement proportion (p) is derived. The mean value for study variables that were included as binary priors in all the six implemented BBHM models spanned between -0.008 and 0.007 with all HDI intervals overlapping with zero value, suggesting that these study-level characteristics did not exert a consistent or statistically meaningful influence on user engagement estimates across categories, and thus were not strong explanatory factors for between-study variability in the Bayesian framework.

Figure 10 displays both the posterior distributions for α and β and the sampling trace plots for MCMC convergence for selection of Individuals in the three examined variables. The left panel illustrates the density of each parameter, while the right panel visualizes draw stability across chains, demonstrating the model's robustness and interpretability for single-category inference (see Figure S1 in the Supplementary material for posterior distributions of other user categories and variables).

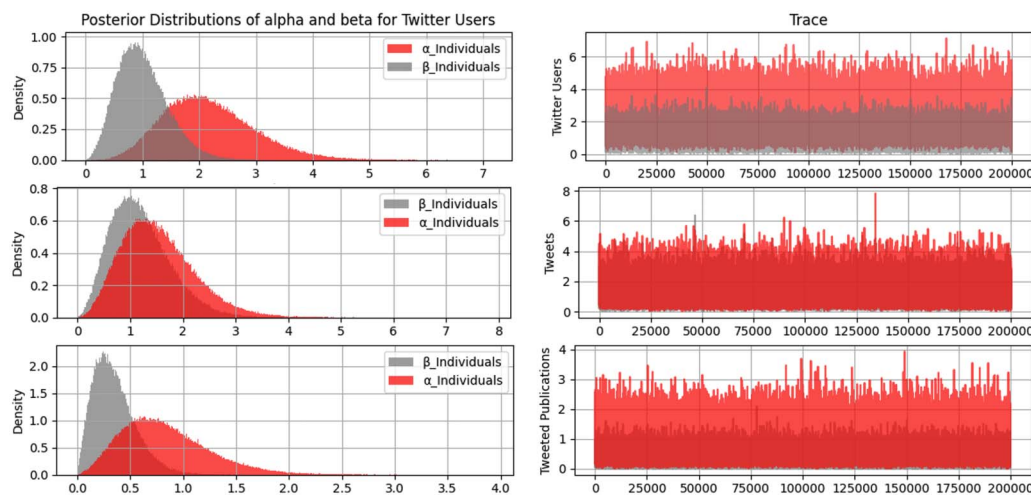


Figure 10. Posterior distributions and sampling traces for alpha (in-category probability) and beta (out-of-category probability) Parameters of Individuals across the three variables of users, tweets, and publications from the BBHM analyses. Left: posterior densities for α (red) and β (gray). Right: MCMC sampling trace plots for 50,000 draws across four chains. These plots illustrate parameter convergence and engagement inference within the BBHM framework.

4.5.1. Meta-analysis of Twitter users

Figure 11 shows the BBHM posterior estimates for Twitter user proportions (values detailed in Table S2 in the Supplementary material). Among individual user types, academic individuals had an estimated engagement proportion of 33% ($\alpha = 1.303$, HDI: 0.433–2.224; $\beta = 2.635$, HDI: 1.012–4.381). Professional individuals followed with 19% ($\alpha = 0.589$, $\beta = 2.554$), and mixed/unclassified individuals with 24%. “Individuals excluding Academics” showed 35%, while the overall Individuals group had the highest engagement at 68% ($\alpha = 2.172$; $\beta = 1.003$).

Organizational user groups contributed far less: academic organizations (10%), professional organizations (19%), and mixed organizations (16%). Among science communicators, academic (12%), bot (8%), and other types (ranging from 9–13%) showed limited engagement.

Mixed groups—unclassified accounts—had engagement proportions of 19% (function level) and 12% (unit level), which shows the prevalence of ambiguous or hybrid user identities across studies.

4.5.2. Meta-analysis of tweets

Figure 12 presents BBHM results for tweet proportions (values detailed in Table S3 in the Supplementary material). Academic individuals contributed 22% of tweets ($\alpha = 0.764$; $\beta = 2.639$), professional individuals 16%, and mixed individuals 30%. “Individuals excluding Academics” had 34%, and the broader Individuals group reached 56% ($\alpha = 1.529$; $\beta = 1.197$). Organizational user categories remained low: academic organizations (18%), professional organizations (21%), and mixed organizations (15%). Science communicators showed the lowest engagement overall (10%), with bots, mixed, and professional communicators around 12–15%. Mixed groups had a moderate share (31%), with narrower HDIs indicating greater consistency in tweet activity across studies that remain unclassified.

4.5.3. Meta-analysis of tweeted publications

Figure 13 summarizes BBHM estimates for tweeted publications (values detailed in Table S4 in the Supplementary material). Academic individuals had the highest function-specific share

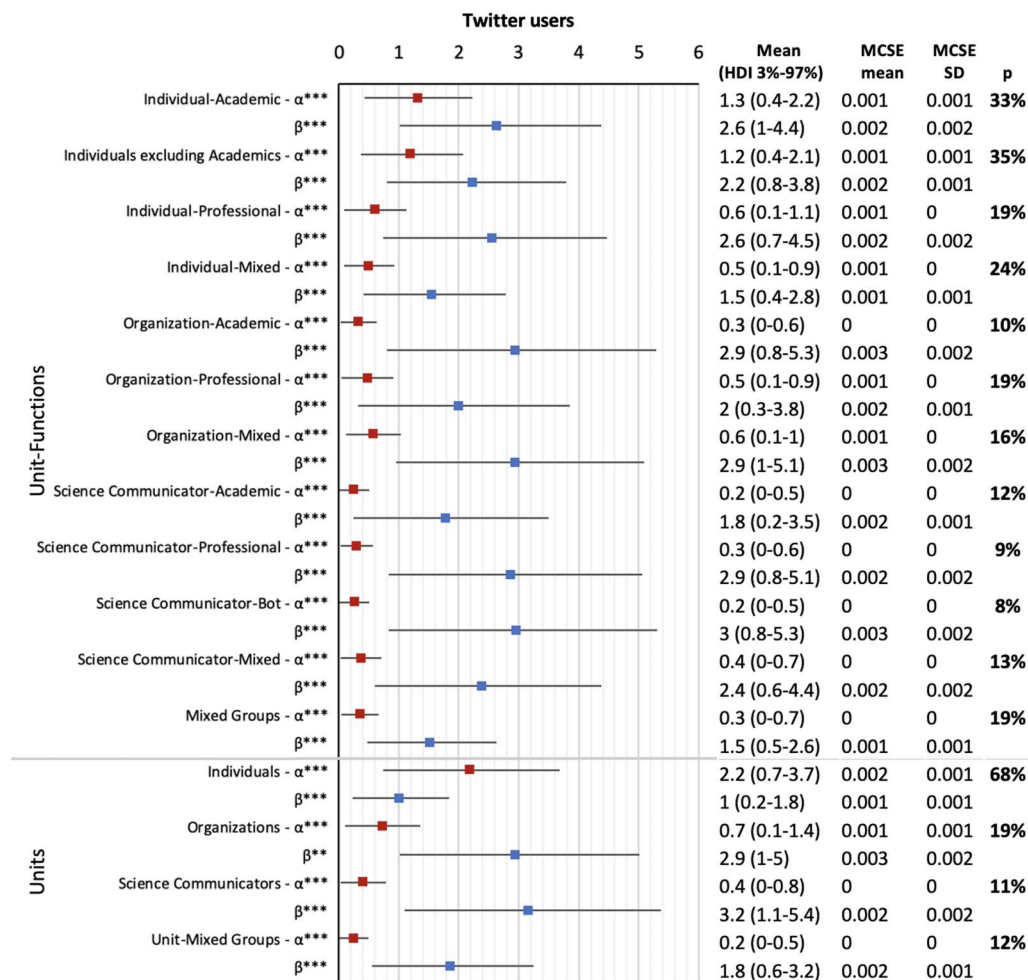


Figure 11. Posterior mean estimates of Twitter users and 94% highest density intervals (HDI 3%–97%) for the parameters α (red) and β (blue) from two BBHM analyses across unit- and function-level user classification schemes. The right column shows the posterior mean proportion p . Asterisks denote Monte Carlo standard error (MCSE) thresholds: *** for $\text{MCSE} \leq 0.001$ (very high precision), ** for $0.001 < \text{MCSE} \leq 0.005$ (high precision). MCSE SD: Standard deviation of Monte Carlo standard error.

(39%; $\alpha = 0.729$, $\beta = 1.157$), followed by mixed (28%) and professional individuals (17%). “Individuals excluding Academics” reached 33%, while all individuals combined accounted for 69% of publications shared ($\alpha = 0.856$, $\beta = 0.392$). Organizational groups posted fewer publications: academic (24%) and professional organizations (25%). Science communicators’ estimates ranged from 19–30%, with bots and mixed communicators each around 28%. Notably, mixed groups outperformed many classified types, contributing 54% of tweeted publications, suggesting great proportion of publications with ambiguous user engagement.

4.6. Comparing Results of BBHM and REM Models

As shown in Figure 14, both BBHM and REM identified the same dominant, well-represented contributors—academic individuals and individuals overall—across all engagement variables. For instance, estimates for academic individuals were aligned: 33% (Twitter users), 22% (tweets), and 39% (tweeted publications). Similarly, individuals overall scored highest in all models: 68% (BBHM) vs. 65% (REM) for users; 56% vs. 55% for tweets; and 69% vs. 50% for tweeted publications.

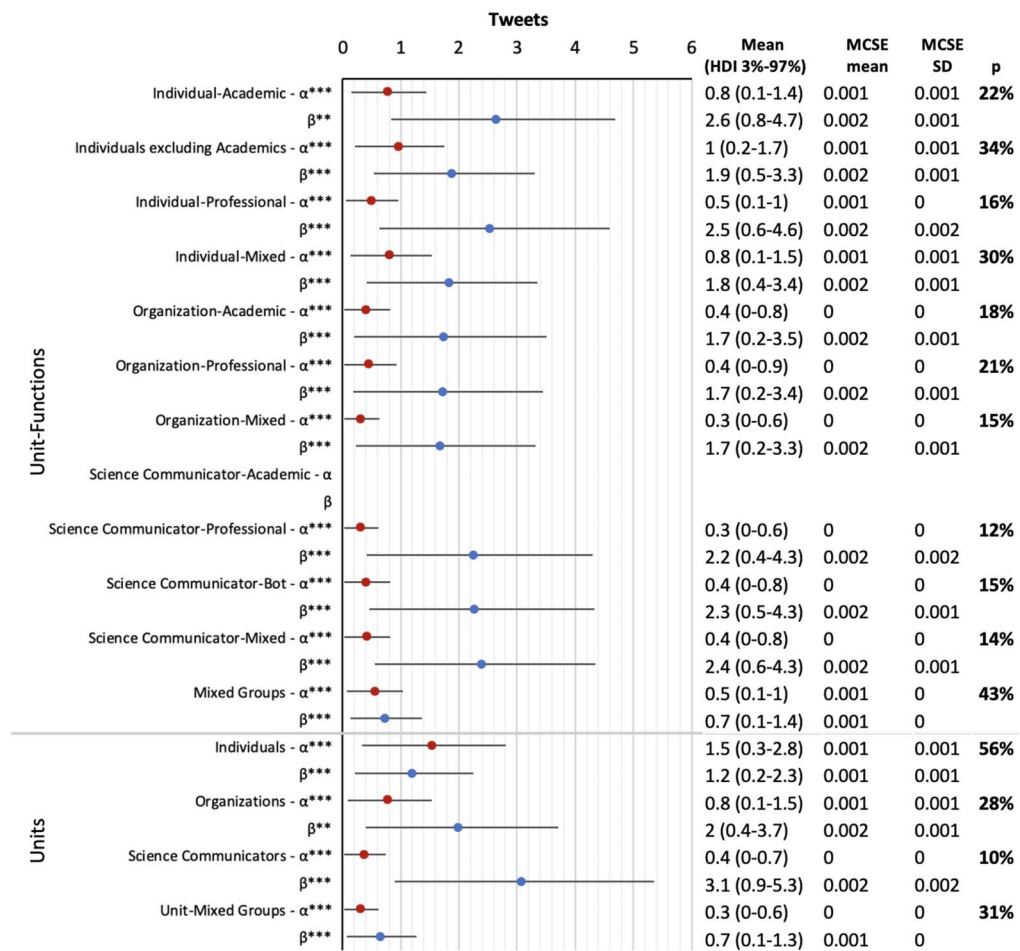


Figure 12. Posterior mean estimates of tweets and 94% HDIs for α and β parameters two BBHM analyses across unit- and function-level user classification schemes. See Figure 11 caption for full explanation.

However, discrepancies between BBHM and REM emerged in groups with sparse data or greater between-study variation. BBHM estimates were typically higher and more stable, while REM often yielded lower proportions and wider confidence intervals, sometimes extending to below-zero values. For academic organizations, BBHM estimated 18% (tweets) and 24% (publications), while REM yielded lower values (9% and 11%). For professional organizations: 21% and 25% (BBHM) vs. 12% and 10% (REM).

Science communicator subtypes showed similar patterns. BBHM estimated 12% (tweets) and 19% (publications) for professional communicators, compared to 5% and 9% (REM). For bots, BBHM reported 15% and 28%, while REM returned 11% and 16%. These differences highlight REM's tendency to underrepresent smaller or inconsistent groups due to its reliance on between-study variance. BBHM estimates still fall in the estimated intervals in REM although much closer to the higher end.

Mixed user groups were consistently high across both models: 19%, 31%, and 54% (BBHM) vs. 20%, 30%, and 56% (REM). This consistency across models shows the significant role of unclassified user accounts in research dissemination—highlighting the limitations of user classification and rigid typologies in capturing social media engagement.

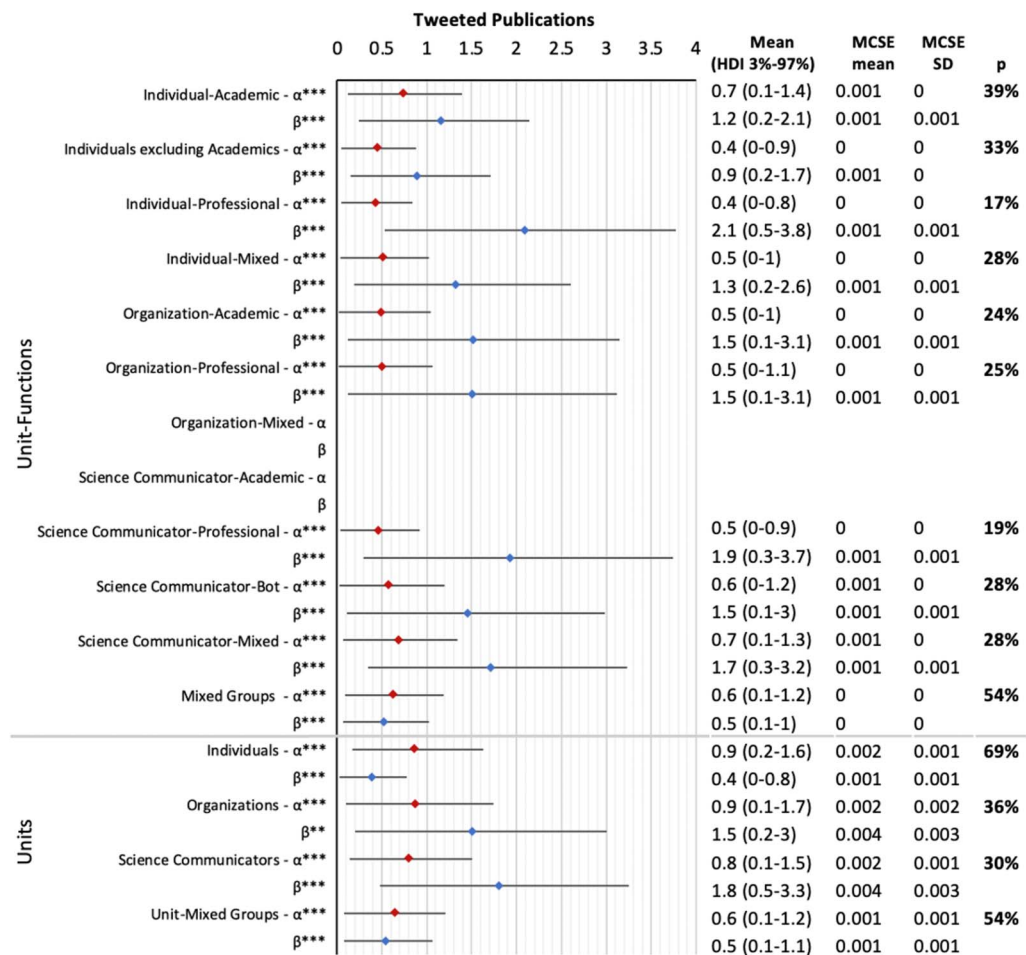


Figure 13. Posterior mean estimates of tweeted publications and 94% HDIs for α and β parameters from two BBHM analyses across unit- and function-level user classification schemes. See Figure 11 caption for full explanation.

Overall, BBHM's hierarchical structure and ability to use the strength across studies enabled it to produce plausible, bounded, and stable estimates, even in categories with limited data. In contrast, REM's performance was constrained by its sensitivity to sample size and between-study heterogeneity, leading to lower estimates in smaller groups and less interpretability in confidence intervals. These findings suggest that while REM remains a valid benchmark, BBHM offers reliable modeling for altmetrics research, especially due to the complexity of data sets where user types were diverse, fragmented, and classification uncertainty was prevalent.

4.7. The Mean Difference Between User Categories

To further assess which user types are more active in disseminating academic publications compared to academic individuals, a series of mean difference analyses was conducted using study-level proportions. Figure 15 illustrates the mean differences between Academic Individuals and all other user categories. In terms of user account proportions, Academic Individuals significantly outnumbered Organizational accounts by 12% and Science Communicators by 15% ($p < 0.01$ for both). However, there was no significant difference between Academic Individuals and Individuals excluding Academics ($I \nsubseteq A$), indicating that other types of

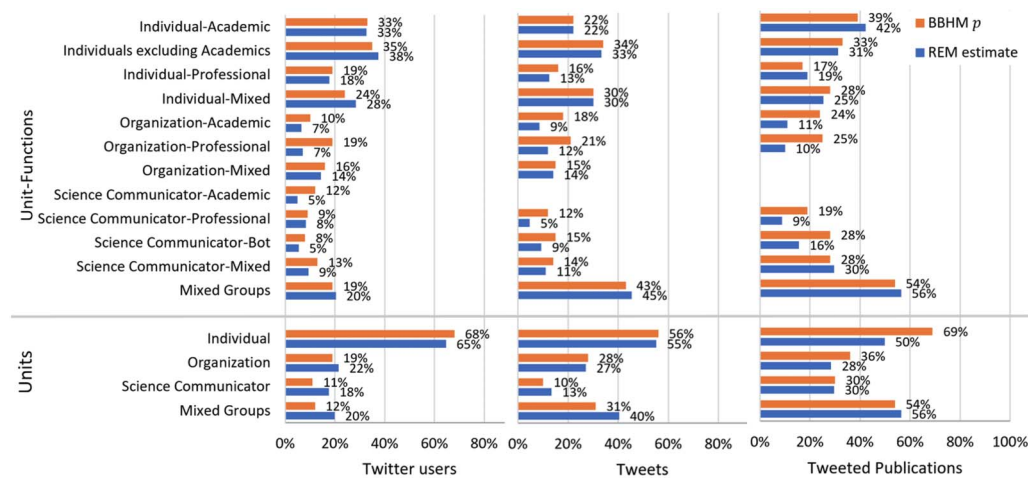


Figure 14. Comparison between the overall proportion estimates of unit- and function-level user categories in the three variables across REM and BBHM models.

individuals (e.g., professionals or unaffiliated users) are just as present on Twitter as academics themselves.

For tweet activity, Academic Individuals tweeted significantly more than Science Communicators by 8%, though the confidence interval (8%–18%) indicates that the result was significant only in one-tailed testing ($p < 0.05$). However, they did not tweet significantly more than Organizational accounts. Interestingly, Academic Individuals posted 12% fewer tweets than Individuals excluding Academics (significant from one side with 95% CI: –24% to –12%), indicating that while academics are visible, they may not be the most prolific tweeters.

With regard to tweeted publications, most comparisons lacked sufficient evidence due to limited studies. The only statistically reliable finding was that Academic Individuals tweeted 33% more publications than Professional Science Communicators ($p < 0.01$), suggesting that researchers are more engaged in direct research dissemination than publishers, media outlets, or journals.

5. LIMITATIONS

This systematic review and meta-analysis is subject to several limitations, which stem from the nature of the included studies, methodological heterogeneity, and limitations in available data.

A key limitation involves the possibility of overlapping data sets across studies. Due to the lack of access to study-level data sets, we could not determine the extent of duplication in user or publication samples. For instance, Zhou and Na (2019) analyzed 42,891 psychology publications from 2011–2015, which could partially overlap with the 37,627 psychology publications in 2014–2015 examined in Udayakumar et al. (2018), despite differences in sampling strategies and user categorization methods. Similarly, Yu (2017) included data from 27 disciplines between 2011–2016, likely overlapping with multiple studies during the same period.

Another significant limitation lies in the inconsistent and subjective categorization of Twitter users across earlier studies. Some merged individuals and organizations, others used automated vs. manual classification, and hybrid roles (e.g., “practitioner researchers”) were often categorized differently. Despite efforts to recategorize users into a unified scheme (unit and

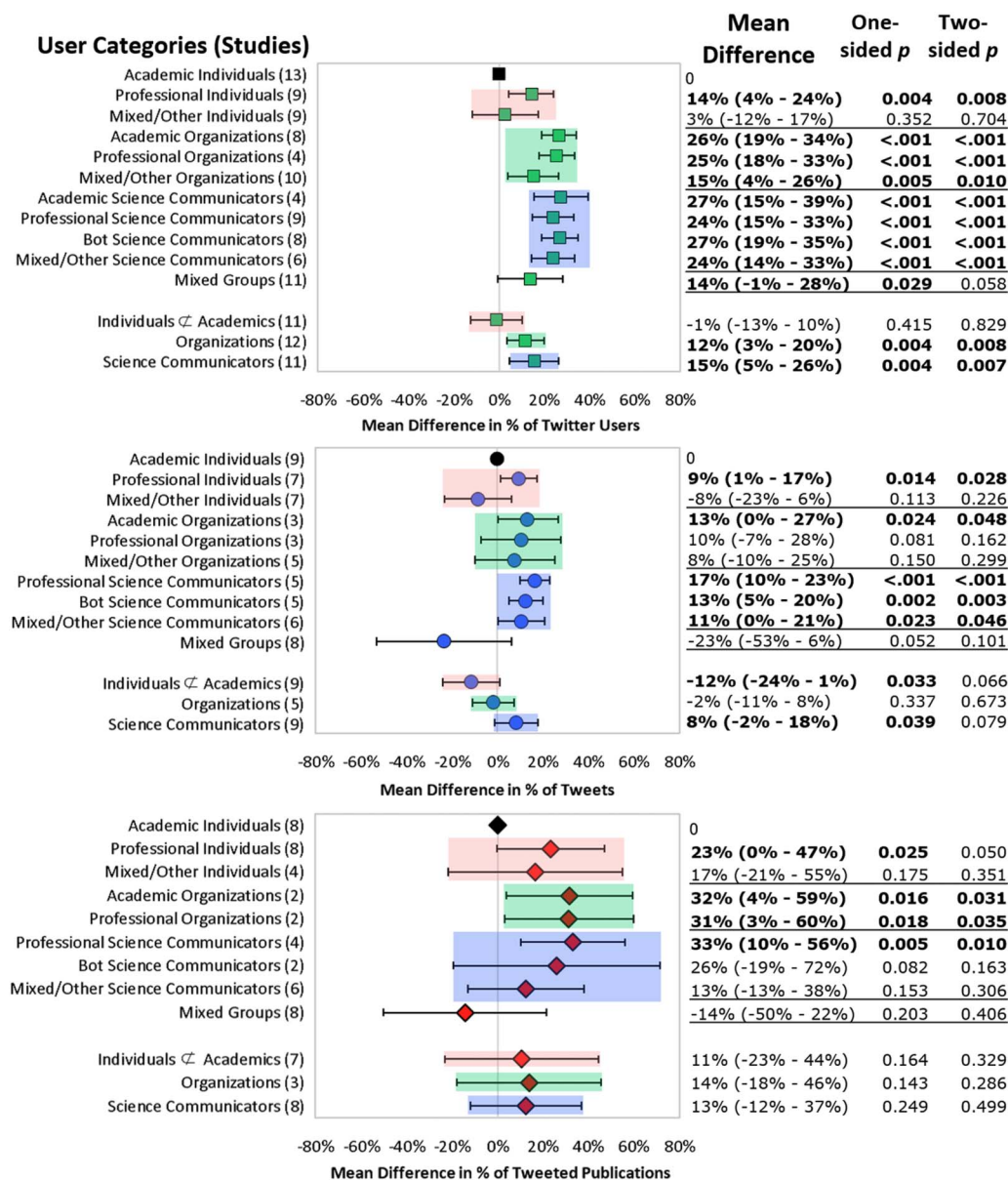


Figure 15. The mean difference between proportion metrics for Academic Individuals and all the other user categories (both function and unit categories). Significant mean differences are shown in bold. Academic Individuals outnumbered the other categories in positive mean differences and vice versa in negative mean differences. Only when the percentages fall on the side of the mean that error bar has no overlapping with the 0% is significant in one-tailed significant differences. $\nsubseteq$: not contains. The colored ribbons indicate significance intervals across user units.

function), ambiguities remained—particularly within the Mixed/Other categories. These inconsistencies may have introduced misclassification errors, especially where studies differed in how bots, media, journalists, or science communicators' accounts were defined and grouped.

We consistently observed high statistical heterogeneity ($I^2 > 90\%$) across most user categories and engagement metrics. Although the primary metaregression revealed some contributing factors—namely the number of user categories, coding method, sample size, and user

activity level, for many user groups, insufficient studies limited our ability to fully model and explain the sources of variability.

Several included studies explicitly focused on highly tweeted articles (e.g., Htoo & Jin-Cheon, 2017) or prolific Twitter users (e.g., Haustein, 2019; Haustein et al., 2016b; Yu et al., 2019). This may have introduced bias, as such users represent only a small proportion of Twitter's user base (Haustein, 2019). The findings indicate that studies centered on high-activity users often reported more bots and higher proportions of unclassified users, suggesting that these may fail to represent the full diversity of Twitter user types. Consequently, findings may be biased toward specific engagement patterns, limiting generalizability to broader Twitter populations.

Initial attempts to use meta-regression to evaluate study-level covariates were limited by small category sizes and failed to produce stable estimates. As a result, binary representations of study-level variables were incorporated into the beta-binomial hierarchical model (BBHM) instead. While BBHM offers a flexible framework for modeling sparse data and accounting for uncertainty, it still relies on assumptions about beta-binomial distributions and may oversimplify the real-world complexity of social media behavior. Furthermore, estimates for some categories (e.g., bots, professional science communicators) still carry wide highest density intervals (HDIs) due to limited observations.

6. DISCUSSION

We set out to address a persistent and unresolved question in altmetrics: who is sharing academic research on Twitter, and how does this reflect academic or societal impact? Despite a decade of empirical inquiry, the lack of standardized methodologies and inconsistent user classifications has prevented the field from reaching a conclusive answer. We conducted the first meta-analysis focused specifically on Twitter user categories involved in scholarly communication, synthesizing 23 studies that offered user-level data, and analyzing three distinct engagement metrics—Twitter users, tweets, and tweeted publications—using both random effects models (REM) and a beta-binomial hierarchical model (BBHM).

6.1. Individuals Dominate, Academics Lead

The findings show a clear pattern: Individual accounts dominate the sharing of academic publications on Twitter. On average, individuals accounted for 66% of all Twitter users, 55% of tweets, and 50% of tweeted publications. Within this group, academic individuals consistently emerged as the largest category, with some exceptions, in tweet frequency. This pattern remained stable across both REM and BBHM approaches, indicating the robustness of the results. However, our comparisons between academic and nonacademic individuals revealed near parity in their contributions: Both groups averaged around 30% across most indicators, although academic users showed slightly higher proportions in tweeted publications (~40%) and slightly lower frequency of tweets overall (22%). This result aligns with Zhang, Gou et al. (2023), who found similar parity when aggregating Twitter users into academic and societal categories, although their broader and overlapping classification scheme complicates direct comparisons.

6.2. Science Communicators and Organizations Play Smaller Roles

Organizations and science communicators—including journals, publishers, and bots—were significantly less active, typically accounting for less than 30% of any engagement metric.

In particular, professional science communicators (e.g., publishers, media) had the lowest engagement levels across all metrics.

6.3. The Mixed Groups Always Have a Role

A significant proportion of publication were disseminated by the “mixed” or unclassified user groups, that could not be reliably categorized in earlier studies due to study design, and ambiguous or incomplete user data.

6.4. Meta-Analysis Models

Given the high heterogeneity found across studies ($I^2 > 50\%$ in 44 of 46 meta-analyses), the BBHM provided more stable and interpretable estimates than traditional REM models. It proved particularly more interpretable in modeling user categories with sparse data or inconsistent reporting (e.g., science communicators), where REM often yielded low or uninterpretable, negative estimates. By incorporating study-level priors and hierarchical structuring, BBHM allowed us to estimate population-level proportions while accounting for between-study variability. The resulting posterior distributions demonstrated high stability and low Monte Carlo standard errors, affirming the precision of our estimates even in the face of limited or inconsistent primary data.

6.5. Implications for Altmetrics and Scholarly Communication

Our results raise critical implications for using Twitter as a proxy for societal impact. While individual (nonacademic) users are active participants, academic users still comprise a large share of engagement, especially at the publication level. This balance complicates the narrative that Twitter altmetrics significantly capture public attention beyond academia. The data suggests a hybrid model of dissemination, where academics and engaged publics both play key roles, and institutional actors and bots contribute less than previously assumed.

6.6. Dynamics in Platform Use

Reviewers rightly note that the dynamics of platform use—and the extent of academic versus institutional engagement—may have evolved due to recent shifts in science communication strategies. While our synthesis reflects pre-2021 patterns, future research should revisit these analyses with updated data sets and explore how user engagement patterns evolve across newer platforms (e.g., Bluesky) or institutional outreach strategies. Our findings offer a benchmark, not a ceiling, and emphasize the importance of adaptive classification systems and temporal validation in altmetric research.

7. CONCLUSION

While our results highlight the significance of academics in disseminating research online, and with that, questioning the applicability of Twitter (or X) as a source to investigate societal impact of research, our contributions also have more practical implications for altmetrics research. A major contribution of this study lies in the development and application of a harmonized user categorization scheme, derived from previous studies and refined to include both unit-level (e.g., individual, organization) and function-level (e.g., academic, professional, mixed) categories. This two-level structure enabled the integration of highly heterogeneous data sets and facilitated meaningful category analyses. This harmonized categorization system, built from prior studies, provides a structured foundation for future research and more

consistent cross-study comparisons. However, its application requires careful consideration of known limitations. We recommend that future user classification efforts distinguish between unit-level categories (e.g., individuals vs. organizations) and function-level roles (e.g., academics, applied researchers, nonresearch professionals such as journalists, science communicators including journals, publishers, bots, and aggregators, and other or unclassified users). This two-level scheme helps clarify ambiguities (see Table 6) and offers a scalable approach for improving the precision and interpretability of user-based altmetric analyses.

Methodologically, this study introduces the beta-binomial hierarchical model (BBHM) for altmetric meta-analysis. BBHM proved especially effective in estimating category-specific engagement levels in the presence of data sparsity and heterogeneity. Compared to the random effects model (REM), BBHM generated more stable and interpretable, meaningful results, making it a strong candidate for future meta-analyses in altmetric research. Nevertheless, the study is not without limitations. Overlapping data sets and varying user categorization approaches across studies may have introduced bias. While beta-binomial hierarchical modeling confirmed the stability of our estimates, the structural limitations inherited from earlier studies highlight the need for open data and transparent reporting in future studies.

Our results underscore that design choices—such as selecting tweets, users, or publications as the unit of analysis—can influence interpretations of engagement. For instance, while user-based metrics reflect equal academic and nonacademic involvement, tweet-based metrics often highlight more public engagement due to the substantially larger scale of public engagement in high profile publications. Thus, future research should carefully consider the metric used when drawing conclusions about impact.

In sum, this study contributes a methodological and conceptual framework for future altmetrics research by demonstrating the value of harmonized user categorization, and robust statistical synthesis and Bayesian modeling. While our findings represent pre-2021 patterns on Twitter (now X), they lay the groundwork for tracking temporal changes and guiding classification practices in emerging platforms, such as Bluesky. Meta-analysis remains a vital tool for retrospective analysis of online engagement with research, and we encourage

Table 6. Suggested Twitter user classification scheme for future studies. The table presents a two-level categorization framework aiming to standardize classification across studies and enable consistent treatment of hybrid roles, across both individuals and organizations

Function	Unit	Individuals	Organization
Academic		Academic scientists	University
Nonacademic or applied researchers		Practitioner researchers	Government organizations, companies, hospitals
		Industry researchers	
		Political researcher	
Nonresearch professionals		All other professionals including	Nonresearch organizations
		Students and journalists	Media outlets
Science communicators		Aggregators, bloggers, personal or no identity bots	Journals, publishers, organizational bots
Others		The general public not in above roles	
Unrecognizable		No user information	No user information

future studies to revisit these patterns as the scholarly communication ecosystem continues to evolve.

AUTHOR CONTRIBUTIONS

Ashraf Maleki: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing—original draft, Writing—review & editing. Kim Holmberg: Conceptualization, Funding acquisition, Resources, Supervision, Writing—original draft, Writing—review & editing.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

This study is part of the research project “Applicability of altmetrics in research impact assessment”, funded by the Academy of Finland (332961), currently known as the Research Council of Finland.

DATA AVAILABILITY

All statistics used in the analyses for this systematic review and meta-analysis are fully documented within the main tables presented in the paper. These include user category proportions, study-level outcome measures, and model estimates across both random effects models (REM) and beta-binomial hierarchical models (BBHM). Should additional clarification or access to structured data be required, for purposes such as replication, reanalysis, or extended exploration, these materials can be made available by the corresponding author upon reasonable request.

REFERENCES

- Abhari, R., Vincent, N., Dambanemuya, H. K., Bodon, H., & Horvát, E.-Á. (2022). Twitter engagement with retracted articles: Who, when, and how? *arXiv*. <https://doi.org/10.48550/arXiv.2203.04228>
- Álvarez-Bornstein, B., & Montesi, M. (2019). Who is interacting with researchers on Twitter? A survey in the field of Information Science. *JLIS: Italian Journal of Library, Archives and Information Science*, 10(2), 87–106. <https://doi.org/10.4403/jlis.it-12530>
- Barthel, S., Tönnies, S., Köhncke, B., Siehndel, P., & Balke, W.-T. (2015). What does Twitter measure? Influence of diverse user groups in altmetrics. In *Proceedings of the 15th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 119–128). <https://doi.org/10.1145/2756406.2756913>
- Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160(901), 268–282. <https://doi.org/10.1098/rspa.1937.0109>
- Bornmann, L. (2015). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics*, 103, 1123–1144. <https://doi.org/10.1007/s11192-015-1565-y>
- Bowman, T. D. (2015). Differences in personal and professional tweets of scholars. *Aslib Journal of Information Management*, 67(3), 356–371. <https://doi.org/10.1108/AJIM-12-2014-0180>
- Card, N. A. (2015). *Applied meta-analysis for social science research*. New York, NY: Guilford Publications.
- Clavier, T., Occhiali, E., Demailly, Z., Compère, V., Veber, B., ... Besnier, E. (2021). The association between professional accounts on social networks Twitter and ResearchGate and the number of scientific publications and citations among anesthesia researchers: Observational study. *Journal of Medical Internet Research*, 23(10), e29809. <https://doi.org/10.2196/29809>, PubMed: 34652279
- Coret, M., Rok, M., Newman, J., Deonarain, D., Agzarian, J., ... Hanna, W. C. (2020). Twitter activity is associated with a higher research citation index for academic thoracic surgeons. *Annals of Thoracic Surgery*, 110(2), 660–663. <https://doi.org/10.1016/j.athoracsur.2019.09.075>, PubMed: 31756321
- Costas, R., van Honk, J., & Franssen, T. (2017). Scholars on Twitter: Who and how many are they? *arXiv*. <https://doi.org/10.48550/arXiv.1712.05667>
- Côté, I. M., & Darling, E. S. (2018). Scientists on Twitter: Preaching to the choir or singing from the rooftops? *Facets*, 3(1), 682–694. <https://doi.org/10.1139/facets-2018-0002>
- Dambanemuya, H. K., Abhari, R., Vincent, N., & Horvát, E.-Á. (2022). Twitter engagement with retracted articles: Who, when, and how? *arXiv*. <https://doi.org/10.48550/arXiv.2203.04228>
- de Melo Maricato, J., & de Castro Manso, B. L. (2022). Characterization of the communities of attention interacting with scientific papers on Twitter: Altmetric analysis of a Brazilian university. *Scientometrics*, 127, 3815–3835. <https://doi.org/10.1007/s11192-022-04442-2>

- Didegah, F., Mejlgaard, N., & Sørensen, M. P. (2018). Investigating the quality of interactions and public engagement around scientific papers on Twitter. *Journal of Informetrics*, 12(3), 960–971. <https://doi.org/10.1016/j.joi.2018.08.002>
- Doebler, P., Frick, S., & Doeblér, A. (2024). Beta-binomial meta-analysis of individual differences based on sample means and standard deviations: Studying reliability of sum scores of binary items. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000649>, PubMed: 38483524
- Dudek, J., Bowman, T. D., & Costas, R. (2018). When do tweeters tweet about science? Exploratory analysis of the Twitter dissemination of scientific publications by weekdays and months. In *STI 2018 Conference Proceedings* (pp. 1566–1578). Centre for Science and Technology Studies (CWTS).
- Erdt, M., Nagarajan, A., Sin, S.-C. J., & Theng, Y.-L. (2016). Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics*, 109, 1117–1166. <https://doi.org/10.1007/s11192-016-2077-0>
- Erskine, N., & Hendricks, S. (2021). The use of Twitter by medical journals: Systematic review of the literature. *Journal of Medical Internet Research*, 23(7), e26378. <https://doi.org/10.2196/26378>, PubMed: 34319238
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC Press. <https://doi.org/10.1201/b16018>
- Günhan, B. K., Röver, C., & Friede, T. (2020). Random-effects meta-analysis of few studies involving rare events. *Research Synthesis Methods*, 11(1), 74–90. <https://doi.org/10.1002/jrsm.1370>, PubMed: 31348846
- Harzing, A.-W. (2007). *Publish or perish* [Computer software]. <https://harzing.com/resources/publish-or-perish>
- Hassan, S.-U., Aljohani, N. R., Idrees, N., Sarwar, R., Nawaz, R., ... Herrera, F. (2020). Predicting literature's early impact with sentiment analysis in Twitter. *Knowledge-Based Systems*, 192, 105383. <https://doi.org/10.1016/j.knsys.2019.105383>
- Hassan, S.-U., Bowman, T. D., Shabbir, M., Akhtar, A., Imran, M., & Aljohani, N. R. (2019). Influential tweeters in relation to highly cited articles in altmetric big data. *Scientometrics*, 119, 481–493. <https://doi.org/10.1007/s11192-019-03044-9>
- Haunschild, R., Bornmann, L., Potnis, D., & Tahamtan, I. (2021). Investigating dissemination of scientific information on Twitter: A study of topic networks in opioid publications. *Quantitative Science Studies*, 2(4), 1486–1510. https://doi.org/10.1162/qss-a_00168
- Haustein, S. (2019). Scholarly Twitter metrics. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 729–760). Cham: Springer. https://doi.org/10.1007/978-3-030-02511-3_28
- Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (2016a). Tweets as impact indicators: Examining the implications of automated “bot” accounts on Twitter. *Journal of the Association for Information Science and Technology*, 67(1), 232–238. <https://doi.org/10.1002/asi.23456>
- Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLOS ONE*, 10(3), e0120495. <https://doi.org/10.1371/journal.pone.0120495>, PubMed: 25780916
- Haustein, S., & Peters, I. (2020). Commemorating Judit Bar-Ilan from bibliometric and altmetric perspectives. *Scientometrics*, 123, 1211–1224. <https://doi.org/10.1007/s11192-020-03448-y>
- Haustein, S., Tsou, A., Minik, V., Brinson, D., Hayes, E., & Sugimoto, C. R. (2016b). Identifying Twitter user communities in the context of altmetrics. In *3rd Altmetrics Conference*. Bucharest.
- Holmberg, K. (2015). *Altmetrics for information professionals: Past, present and future*. Witney, Oxfordshire: Chandos Publishing.
- Holmberg, K., Bowman, T. D., Haustein, S., & Peters, I. (2014). Astrophysicists' conversational connections on Twitter. *PLOS ONE*, 9(8), e106086. <https://doi.org/10.1371/journal.pone.0106086>, PubMed: 25153196
- Holmberg, K., & Thelwall, M. (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*, 101, 1027–1042. <https://doi.org/10.1007/s11192-014-1229-3>
- Htoo, T. H. H., & Jin-Cheon, N. (2017). Who are tweeting research articles and why? *Journal of Information Science Theory and Practice*, 5(3), 48–60. <https://doi.org/10.1633/JISTaP.2017.5.3.4>
- Hughes, H., Hughes, A., & Murphy, C. G. (2017). The use of Twitter by trauma and orthopaedic surgery journals: Twitter activity, impact factor, and alternative metrics. *Cureus*, 9(12), e1931. <https://doi.org/10.7759/cureus.1931>, PubMed: 29464138
- IBM. (2021). *IBM SPSS statistics for windows*, Version 28.0.0.0 [Computer software]. IBM Corporation.
- Jansen, K., & Holling, H. (2023). Rare events meta-analysis using the Bayesian beta-binomial model. *Research Synthesis Methods*, 14(6), 853–873. <https://doi.org/10.1002/jrsm.1662>, PubMed: 37607885
- Joubert, M., & Costas, R. (2019). Getting to know science tweeters: A pilot analysis of South African Twitter users tweeting about research articles. *Journal of Altmetrics*, 2(1), 2. <https://doi.org/10.29024/joa.8>
- Ke, Q., Ahn, Y.-Y., & Sugimoto, C. R. (2017). A systematic identification and analysis of scientists on Twitter. *PLOS ONE*, 12(4), e0175368. <https://doi.org/10.1371/journal.pone.0175368>, PubMed: 28399145
- Khandelwal, A., & Tagat, A. (2021). #DevResearch: Exploring development researchers Twitter use for research dissemination. *Scholarly and Research Communication*, 12(1). <https://doi.org/10.22230/src.2021v12n1a395>
- Kwak, H., & Lee, J. G. (2014). Has much potential but biased: Exploring the scholarly landscape in Twitter. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 563–564). <https://doi.org/10.1145/2567948.2576956>
- Lemke, S., Bräuer, P., & Peters, I. (2021). Does the general public share research on Twitter? A case study on the online conversation about the search for a nuclear repository in Germany. https://epub.uni-regensburg.de/44940/1/isi_lemke_et_al.pdf
- Maleki, A. (2014). Twitter users in science tweets linking to articles: The case of Web of Science articles with Iranian authors. In *SIG/MET Workshop in 77th ASIS&T Annual Meeting*.
- Maleki, A. (2016). Do tweets indicate scholarly communications? In *2016 Altmetrics Workshop*. Bucharest. https://altmetrics.org/wp-content/uploads/2016/09/altmetrics16_paper_8.pdf
- Maleki, A. (2018). Connectedness and consistency of Twitter users posting scholarly articles in astronomy and astrophysics. In *2018 Altmetrics Workshop*. London. https://altmetrics.org/wp-content/uploads/2018/09/altmetrics18_paper_13_Maleki.pdf
- Mezzetti, M., & Negri, I. (2024). Hierarchical Bayesian model to estimate and compare research productivity of Italian academic statisticians. *Scientometrics*, 129, 7443–7474. <https://doi.org/10.1007/s11192-024-05154-5>
- Mikolajewicz, N., & Komarova, S. V. (2019). Meta-analytic methodology for basic research: A practical guide. *Frontiers in Physiology*, 10, 203. <https://doi.org/10.3389/fphys.2019.00203>, PubMed: 30971933
- Mohammadi, E., Thelwall, M., Kwasny, M., & Holmes, K. L. (2018). Academic information on Twitter: A user survey. *PLOS*

- ONE, 13(5), e0197265. <https://doi.org/10.1371/journal.pone.0197265>, PubMed: 29771947
- Mongeon, P., Bowman, T. D., & Costas, R. (2023). An open data set of scholars on Twitter. *Quantitative Science Studies*, 4(2), 314–324. https://doi.org/10.1162/qss_a_00250
- Negrín-Hernández, M.-A., Martel-Escobar, M., & Vázquez-Polo, F.-J. (2021). Bayesian meta-analysis for binary data and prior distribution on models. *International Journal of Environmental Research and Public Health*, 18(2), 809. <https://doi.org/10.3390/ijerph18020809>, PubMed: 33477861
- Orduña-Malea, E., & Costas, R. (2021). Link-based approach to study scientific software usage: The case of VOSviewer. *Scientometrics*, 126, 8153–8186. <https://doi.org/10.1007/s11192-021-04082-y>
- 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. <https://doi.org/10.1108/AJIM-02-2017-0055>
- Ortega, J. L. (2020). Altmetrics data providers: A meta-analysis review of the coverage of metrics and publication. *El Profesional de la Información*, 29(1), e290107. <https://doi.org/10.3145/epi.2020.ene.07>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., ... Moher, D. (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>, PubMed: 33782057
- Pandian, N. D. S., Na, J.-C., Veeramachaneni, B., & Boothaladinni, R. V. (2019). Altmetrics: Factor analysis for assessing the popularity of research articles on Twitter. *Journal of Information Science Theory and Practice*, 7(4), 33–44. <https://doi.org/10.1633/JISTaP.2019.7.4.3>
- Pigott, T. D. (2012). *Advances in meta-analysis*. New York, NY: Springer. <https://doi.org/10.1007/978-1-4614-2278-5>
- Said, A., Bowman, T. D., Abbasi, R. A., Aljohani, N. R., Hassan, S.-U., & Nawaz, R. (2019). Mining network-level properties of Twitter altmetrics data. *Scientometrics*, 120, 217–235. <https://doi.org/10.1007/s11192-019-03112-0>
- Sanchez, T. W. (2021). Urban planning academics: Tweets and citations. *Urban Planning*, 6(1), 146–153. <https://doi.org/10.17645/up.v6i1.3720>
- Sergiadis, A. (2018). Analysis of tweets mentioning works from an institutional repository. *Journal of New Librarianship*, 3(1), 130–150. <https://doi.org/10.21173/newlibs/4/27>
- Schmid, C. H. (2001). Using Bayesian inference to perform meta-analysis. *Evaluation & the Health Professions*, 24(2), 165–189. <https://doi.org/10.1177/01632780122034867>, PubMed: 11523385
- Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781483398105>
- Schmitt, M., & Jaschke, R. (2017). What do computer scientists tweet? Analyzing the link-sharing practice on Twitter. *PLOS ONE*, 12(6), e0179630. <https://doi.org/10.1371/journal.pone.0179630>, PubMed: 28636619
- Schünemann, H., Brożek, J., Guyatt, G., & Oxman, A. (2013). *GRADE handbook*. Cochrane Training. <https://gdt.gradeapro.org/app/handbook/handbook.html>
- Toupin, R., Millerand, F., & Larivière, V. (2019). Scholarly communication or public communication of science? Assessing who engage with climate change research on Twitter. In *17th International Conference on Scientometrics and Informetrics (ISSI 2019) with a special STI conference track*. Rome, Italy.
- Tsou, A., Bowman, T. D., Ghazinejad, A., & Sugimoto, C. R. (2015). Who tweets about science? In *Proceedings of the 2015 International Society for Scientometrics and Informetrics*. Istanbul, Turkey.
- Tur-Viñes, V., Segarra-Saavedra, J., & Hidalgo-Marí, T. (2018). Use of Twitter in Spanish communication journals. *Publications*, 6(3), 34. <https://doi.org/10.3390/publications6030034>
- Udayakumar, S., Senadeera, D. C., Yamunarani, S., & Cheon, N. J. (2018). Demographics analysis of Twitter users who tweeted on psychological articles and tweets analysis. *Procedia Computer Science*, 144, 96–104. <https://doi.org/10.1016/j.procs.2018.10.509>
- Vainio, J., & Holmberg, K. (2017). Highly tweeted science articles: Who tweets them? An analysis of Twitter user profile descriptions. *Scientometrics*, 112, 345–366. <https://doi.org/10.1007/s11192-017-2368-0>
- Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215–247. <https://doi.org/10.3102/1076998609346961>
- Wang, Y., & Zuccala, A. (2021). Scholarly book publishers as publicity agents for SSH titles on Twitter. *Scientometrics*, 126, 4817–4840. <https://doi.org/10.1007/s11192-021-03947-6>
- Wong, K., Piraquive, J., & Levi, J. R. (2018). Social media presence of otolaryngology journals: The past, present, and future. *The Laryngoscope*, 128(2), 363–368. <https://doi.org/10.1002/lary.26727>, PubMed: 28600839
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT Press.
- Xia, F., Su, X., Wang, W., Zhang, C., Ning, Z., & Lee, I. (2016). Bibliographic analysis of Nature based on Twitter and Facebook altmetrics data. *PLOS ONE*, 11(12), e0165997. <https://doi.org/10.1371/journal.pone.0165997>, PubMed: 27906981
- Yadav, G. S., Nagarkatti, N. R., Rohondia, S. O., Erfani, H., Kilpatrick, C. C., & Turrentine, M. A. (2019). Academic tweeting in #ObGyn. Where do we stand? *Journal of Perinatal Medicine*, 47(8), 867–870. <https://doi.org/10.1515/jpm-2019-0166>, PubMed: 31494634
- Ye, Y. E., Na, J.-C., & Oh, P. (2022). Are automated accounts driving scholarly communication on Twitter? A case study of dissemination of COVID-19 publications. *Scientometrics*, 127, 2151–2172. <https://doi.org/10.1007/s11192-022-04343-4>, PubMed: 35370323
- Yu, H. (2017). Context of altmetrics data matters: An investigation of count type and user category. *Scientometrics*, 111, 267–283. <https://doi.org/10.1007/s11192-017-2251-z>
- Yu, H., Xiao, T., Xu, S., & Wang, Y. (2019). Who posts scientific tweets? An investigation into the productivity, locations, and identities of scientific tweeters. *Journal of Informetrics*, 13(3), 841–855. <https://doi.org/10.1016/j.joi.2019.08.001>
- Zhang, L., Gou, Z., Fang, Z., Sivertsen, G., & Huang, Y. (2023). Who tweets scientific publications? A large-scale study of tweeting audiences in all areas of research. *Journal of the Association for Information Science and Technology*, 74(13), 1485–1497. <https://doi.org/10.1002/asi.24830>
- Zhang, L., & Wang, J. (2021). What affects publications' popularity on Twitter? *Scientometrics*, 126, 9185–9198. <https://doi.org/10.1007/s11192-021-04152-1>
- Zhou, Y., & Na, J.-C. (2019). A comparative analysis of Twitter users who tweeted on psychology and political science journal articles. *Online Information Review*, 43(7), 1188–1208. <https://doi.org/10.1108/OIR-03-2019-0097>