Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-023-02331-4 OPEN

Check for updates

A solid camp with flowing soldiers: heterogeneous public engagement with science communication on Twitter

Bin Yang[1,2], Naipeng Chao[3]✉ & Cheng-Jun Wang[1,2]✉

The science communication community is constantly evolving. However, limited research has explored the relationship between engagement heterogeneity and fluctuations in science communication. This study aims to address this issue by examining the dissemination of scientific research on Twitter using network analysis. The findings reveal the sensitivity of low-engagement users in two distinct aspects. First, low-engagement users' dissemination of scientific information is positively associated with the overall trend of scientific communication on social media, suggesting their heightened susceptibility to fluctuations and disengagement compared to other users. Second, low-engagement users show decreased attention to health-related topics during fluctuation periods. In light of these findings, an analytical model is developed to integrate the heterogeneity of information acceptance thresholds and external shocks. The simulation results of the model are consistent with empirical observations, highlighting the heterogeneity of information acceptance thresholds in science communication. This study contributes to the understanding of fluidity as the essence of science communication. As the proverb goes, a solid camp is guarded by everchanging soldiers. The solid camp stabilizes science communication communities while flowing soldiers enable the influence of science communication to cross communities.

[1] Computational Communication Collaboratory, Nanjing University, Nanjing 210093, China. [2] School of Journalism and Communication, Nanjing University, Nanjing 210093, China. [3] School of Media and Communication, Shenzhen University, Shenzhen 518060, China. ✉email: npchao@szu.edu.cn; wangchengjun@nju.edu.cn

Introduction

Public engagement has become an increasingly critical element of science communication (Bubela et al., 2009; Miah, 2017; Powell & Kleinman, 2008; Wynne, 2006; Nisbet & Scheufele, 2009). Prior research has identified the crucial role of scientific information dissemination in disease prevention (Funk et al., 2010; Funk et al., 2009; Wang et al., 2019; Wu et al., 2012; Yang et al., 2020), especially in promoting self-protective measures among individuals (Wakefield et al., 2010) and encouraging people to comply with epidemic prevention policies (Van Bavel et al., 2020; Nan et al., 2021). Due to the intricate nature of the population and communication environment (Van Bavel et al., 2020), spreading scientific research on social media during a pandemic outbreak poses significant challenges to policymakers and science communication experts. However, it is challenging to engage the public in disseminating scientific information (Nisbet & Scheufele, 2009) due to a lack of interest (Burns et al., 2003; Maltese et al., 2014; Baram-Tsabari & Osborne, 2015), trust (Hyland-Wood et al., 2021; Nisbet & Scheufele, 2009), and understanding (Burns et al., 2003; Bullock et al., 2019).

This study primarily draws on the *Public Engagement with Science* (PES) model (Bucchi & Trench, 2014; Burchell, 2015; Kessler et al., 2022). According to the PES model, the science community is only one part of all social actors. Compared to the deficit model, which emphasizes the public awareness of science (Bucchi & Trench, 2014), the PES model asserts that the interactions or dialogues among stakeholders play an essential role in science communication. Thus, the PES model promotes a two-way dialogue between science and the public. Nevertheless, there are still research gaps in the PES model. Particularly, the PES model is criticized for not specifying the outcomes or only using the dialogic approach to fill perceived knowledge deficits (Kessler et al., 2022). Following the tradition of the PES model, we define public engagement with science as the interaction or dialogues among different social actors of science communication. Further, PES is a multidimensional concept, and we primarily focus on public engagement with science communication on social media.

We claim that the public's engagement heterogeneity plays a crucial role in shaping the fluctuation of science communication on social media (Hyland-Wood et al., 2021). Although the public can engage in science communication, their level of engagement significantly varies. Only a few participants can stably engage with science communication for a long time. Most participants tend to pursue hot topics and can only engage with science communication for a short time. This study primarily focuses on the heterogeneity of participants' engagement and its potential impact. Just like the other communities, the science communication community is constantly flowing. As a Chinese proverb goes, the solid camp is guarded by soldiers like flowing water. New soldiers arrive, serve, but eventually leave. However, the fluctuations in science communication can have unintended consequences, such as pre-existing systemic inequity and implicit social biases (Perry et al., 2021; Gray et al., 2020; Bonaccorsi et al., 2020).

The intricate interplay between heterogeneous engagement and external shocks requires careful examination. According to the analytical framework established by Crane and Sornette (2008), public attention is influenced by individual preference, social influence, and external shocks. For example, Park et al. (2021) show that entertainment and human interest frames (e.g., humor and colloquial language) can effectively attract public attention. In the light of Crane and Sornette (2008), we define external shocks as the exogenous impacts outside of individuals and social networks in this study. Science communication could be influenced by various factors characterizing external shocks, such as competitive events, conspiracy theories, and misinformation (Scheufele & Krause, 2019; West & Bergstrom, 2021; B. Yang et al., 2023). Although different audiences may have distinct attitudes to the same external shock (Byrne & Hart, 2009), users who share similar levels of engagement may demonstrate homogeneity and uniform responses to external shocks. The confluence of external shocks and population heterogeneity can intensify the disparities in science communication (Hart & Nisbet, 2014). Specifically, certain social groups may experience a more significant decline in science communication or increase their implicit biases during external shocks, exacerbating their vulnerability to infection (Bonaccorsi et al., 2020; Kaim et al., 2021; Gray et al., 2020).

To analyze the relationship between engagement heterogeneity and macroscopic fluctuations of science communication, we conduct this study on the spread of scientific research on COVID-19 on Twitter, using the Altmetric database, which records Twitter users' retweets of scientific research. The participants were categorized into subgroups based on their levels of engagement with science communication. We find that users with different engagement levels exhibit distinct patterns in audience flow and topic content preferences. Our findings emphasize the significance of external shocks, which introduce fluctuations and escalate the existing imbalances and biases in science communication. Nevertheless, our findings reveal a significant positive correlation between low-engagement users' behaviors and the fluctuations in science communication. On the one hand, low-engagement individuals are more likely to disengage during the decline of science communication. On the other hand, they are more likely to engage in science communication during issue escalation. This observation contradicts the belief that low-engagement users are uninterested in scientific information (Burns et al., 2003; Maltese et al., 2014; Baram-Tsabari & Osborne, 2015).

To better explain this phenomenon, we developed a theoretical model incorporating engagement heterogeneity and external shocks into the SIRS model. The model posits that individual engagement with science communication depends on whether the popularity of the information exceeds one's information acceptance threshold. Our model suggests that individuals with low information acceptance thresholds are more likely to engage with science communication in the early stage and disengage when the novelty of science communication decreases. The predictions generated by our model explain the dissemination of scientific information among low-engagement individuals.

The remainder of this paper is organized as follows: Section 2 defines parameters for user engagement. Dividing users into subgroups based on user engagement, we analyzed the evolution of subgroups' flow, the importance of user engagement in communication networks, and their propagation preferences in external shock. In Section 3, we described the details of the model and conducted simulation experiments on a scale-free network to test its validity. Finally, we summarize these results in Section 4.

Data analysis

Subgroup mobility analysis. To analyze the dissemination of COVID-19-related research on Twitter, we filtered Altmetric data using the keyword "COVID-19" and captured 13,806,356 retweets of COVID-19-related research between March 2020 and May 2022, 27 months in total. This period covers the relatively complete spread of the topic, providing a detailed overview of public engagement. We measure engagement heterogeneity by dividing users into subgroups based on their degree of engagement—the number of months a user has participated in the science communication—within 27 months. Our measurement of

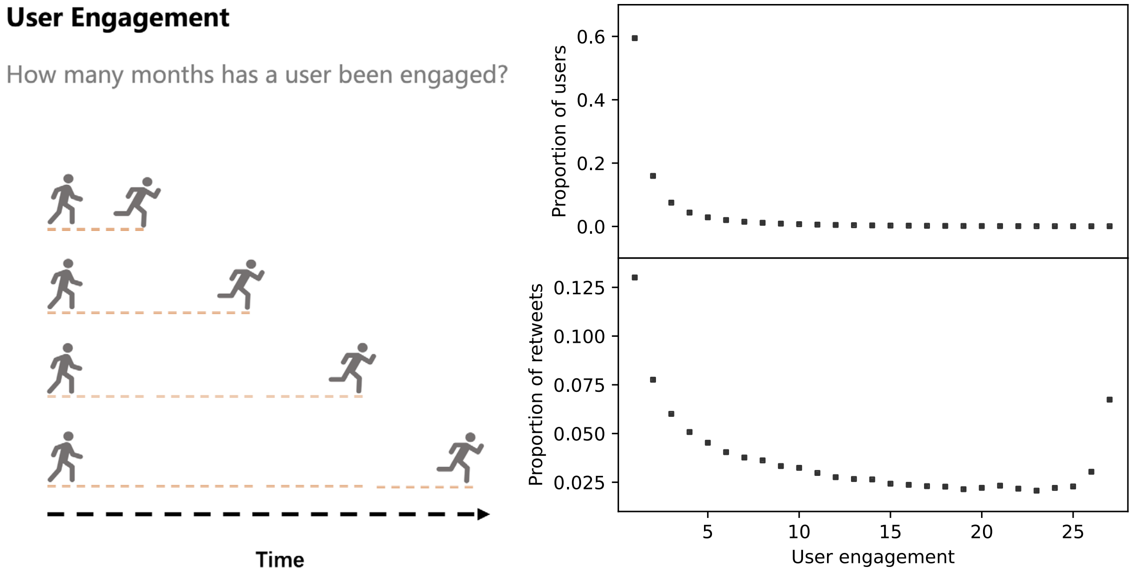

Fig. 1 Quantifying user engagement by its persistence. User engagement is quantified by the number of months a user has engaged with science communication (*left*). The proportion of users in the 27 subgroups (*right upper*) and the proportion of retweets by users in the 27 subgroups (*right bottom*) are shown across the entire time frame.

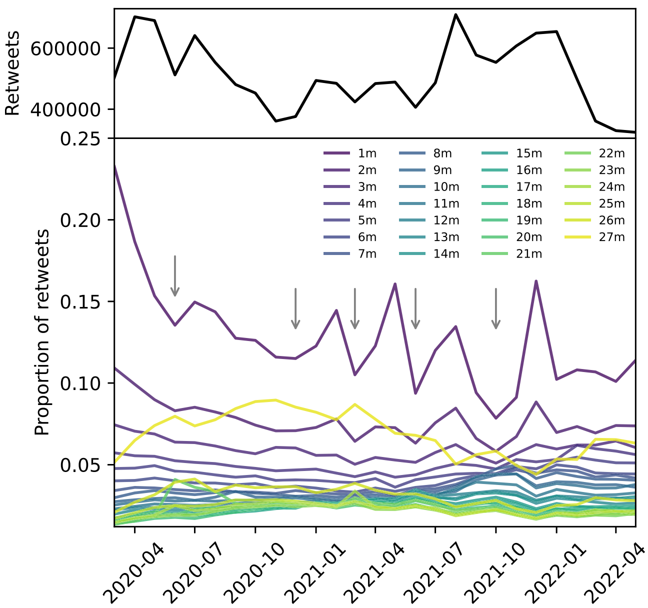

Fig. 2 The fluctuation of public engagement. Overall communication trend of COVID-19-related scientific literature on Twitter (*upper*) and the evolution of the retweet proportion of subgroups (*bottom*). We divide all users into 27 subgroups based on the months they engaged with science communication (ranging from 1 month to 27 months). The proportion of retweets describes the relative information flow of the subgroup.

engagement captures the sustainability of public engagement. Basic statistical data for these subgroups are shown in Fig. 1.

Our investigation began with exploring the dynamics of disparate communication flow among different subgroups. As an initial step, we analyzed the monthly retweet proportion (i.e., the number of monthly retweets in each subgroup divided by the total monthly retweets) across 27 subgroups. We contend that the evolution of retweet proportions within subgroups constitutes a vital metric capable of revealing the heterogeneous user flow within the information system. A pronounced variation in subgroup retweet proportion indicates considerable irregularities in science communication, while a relatively stable retweet proportion suggests a more uniform distribution of information flow among the subgroups.

As demonstrated in Fig. 2, we observed more considerable fluctuations among low-engagement users than high-engagement users, indicating an uneven flow in science communication on Twitter. Additionally, we noted a rapid decline in the proportion of retweets of low-engagement users during the initial spreading phase. This result can be attributed to the decline of low-engagement users after a faster initial increase. Furthermore, we detected a plausible association between the overall retweet trend (represented by the black line in the upper panel of Fig. 2) and the evolution of subgroup users' engagement. Specifically, during a decline in the overall trend, such as in June 2020, December 2020, March 2021, June 2021, and October 2021, low-engagement users experience a dip (see the arrows on the purple line in Fig. 2), while high-engagement users have a minor peak (green line). These findings suggest a potential association between the overall retweet trend and the subgroups' engagement with science communication.

We employ the Maximum Information Coefficient (MIC) (Reshef et al., 2011) and the Pearson correlation coefficient to investigate the relationship between engagement heterogeneity and the overall science communication trends. Pearson correlation coefficient is widely used to quantify the strength of linear correlation between two variables. In addition to linear correlation, the MIC score can also measure the strength of nonlinear relationships between two variables. As depicted in Fig. 3, we find a higher MIC score and significant Pearson correlations for individuals with either very low or very high engagement. Our analysis indicated a significant positive correlation in individuals with low engagement and a significant negative correlation in individuals with high engagement. We also observed a significant linear relationship between user engagement and the previously computed Pearson correlation coefficient ($R^2 = 0.876$, $p < 0.001$). Our results illustrate the non-uniform movement of scientific information during the period of instability and the impact of engagement disparities on subgroup mobility.

In particular, lower-engagement users are more likely to disengage when science communication is at a low ebb. Conversely, low-engagement users are more likely to join during

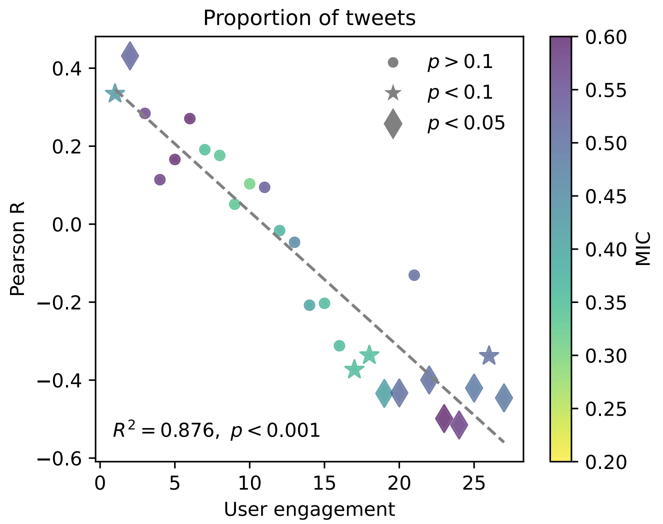

Fig. 3 The association between the subgroup's engagement proportion and the overall engagement. MIC scores are depicted using color, while the statistical significance of the Pearson correlation coefficients is indicated by markers (circles represent non-significant ($p > 0.1$) data points. In contrast, stars and diamonds represent statistically quasi-significant ($p < 0.1$) and significant ($p < 0.05$) data points, respectively). Two scoring metrics are employed to quantify the association between the relative information flow of the subgroup and the overall volatility. The dashed line represents the OLS linear regression line used to fit the Pearson correlation coefficients, with $R^2 = 0.876$ and $p < 0.001$.

periods of increased scientific information spreading. Our findings indicated a higher MIC score and significant Pearson correlations for individuals with either very low or very high engagement. There is a significant positive correlation for individuals with low engagement and a significant negative correlation for individuals with high engagement. On the contrary, low-engagement users are more likely to join during periods of increasing science communication.

Subgroup user centrality analysis. To further understand the implications and impact of the uneven flow of subgroup users, we explore the average centrality of subgroup users in the 27 monthly retweet networks. A higher average centrality score of subgroup users suggests they play a more crucial role in disseminating information (Berry & Widder, 2014). Since the original Altmetric data lacks retweet information, we acquire the data from the Altmetric search page. We match the @username in the tweets with their Twitter user ID, and 85.2% of usernames could be matched. Based on the monthly retweet data, we construct 27 directed and weighted retweet networks, where the nodes represent Twitter users, and the edges indicate retweet relationships. Every edge represents the directed retweet relationship from user i to user j. The weight of the edge represents the frequency of retweets.

We investigate the relative average centrality of subgroup users within the monthly retweet networks. The relative average centrality is calculated by dividing the average centrality of subgroup users by the average centrality of all users in the network. To evaluate the centrality of nodes, we utilize two measures: degree centrality and closeness centrality, which capture distinct aspects of node importance (Freeman, 1978). Degree centrality assesses a node's number of direct connections, indicating its local influence. Closeness centrality identifies nodes that sustain the overall structure by linking different parts of the network by measuring the average distance between a node and all other nodes in the network (Borgatti, 2005).

Furthermore, we consider the differentiation between inward and outward information. In-degree centrality and closeness centrality were employed to assess inward information. The closeness distance function calculates the incoming distance in directed networks. In comparison, the out-degree centrality and the reverse network's closeness centrality were employed to evaluate outward information. The nodes with higher scores in in-degree and closeness centrality indicate that these subgroup users are crucial in providing information, granting them more significant influence and prominence within the network. In contrast, the nodes with higher scores in the out-degree centrality and the reverse network's closeness centrality imply that subgroup users are more proactive and demonstrate outstanding capabilities in disseminating information than the others.

As illustrated in Fig. 4, the analysis of outward information reveals a significant positive correlation and higher MIC scores between the relative average centrality of user subgroups, whose engagement ranges from 6 to 20, and the overall propagation trend. The association is observed in both the out-degree centrality (top left panel) and the closeness centrality on the reverse network (top right panel). In comparison, no significant correlation was found among other users. These findings suggest that external shocks affect participants' mobility and influence their proactive involvement in propagation, particularly for those whose engagement ranges from 6 to 20.

Further examination of inward information, as shown in Fig. 4, reveals that the high-engagement subgroups exhibit a more pronounced association between their relative average centrality and the overall communication trend. High-engagement users usually include vital intermediaries, such as professional science journalists or science communication experts, dedicated to enhancing or maintaining the overall communication structure. Thus, this finding underscores an additional underlying risk during the decline of science communication. It emphasizes the potential weakening of the importance of professionals within the network, thereby posing a challenge to achieve effective science communication.

Subgroup user bias analysis. The inconsistent changes in subgroup users could exacerbate disparities in science communication. This issue became salient amid the COVID-19 pandemic, which created an unprecedented public health emergency and economic crisis. A critical concern is the health-economy dilemma (Kaim et al., 2021), which involves balancing the impact of epidemic prevention measures (e.g., lockdowns) on the economy and public health. Therefore, understanding the general public's preferences for COVID-19 content is critical to effectively addressing this multifaceted challenge (Escandón et al., 2021). To address this issue, we investigate the relationship between subgroups' spreading fluctuations and the evolution of their content preferences.

We analyze the proportion of retweeted scientific research within areas classified under the corresponding Field of Research (FoR) categories. Focusing on the health-economy dilemma, we explore the fields of "11 Medicine and Health Sciences" (corresponding to the focus on public health crises), as well as the fields of "16 Human Social Studies" and "14 Economics" (corresponding to the focus on socio-economic impacts). As shown in Fig. 5, scientific research during the early stages of COVID-19 mainly focused on medical and health research up until September 2020. Afterward, there is a relatively stable distribution of scientific research among medical and health research, human social research, and economics fields. Therefore,

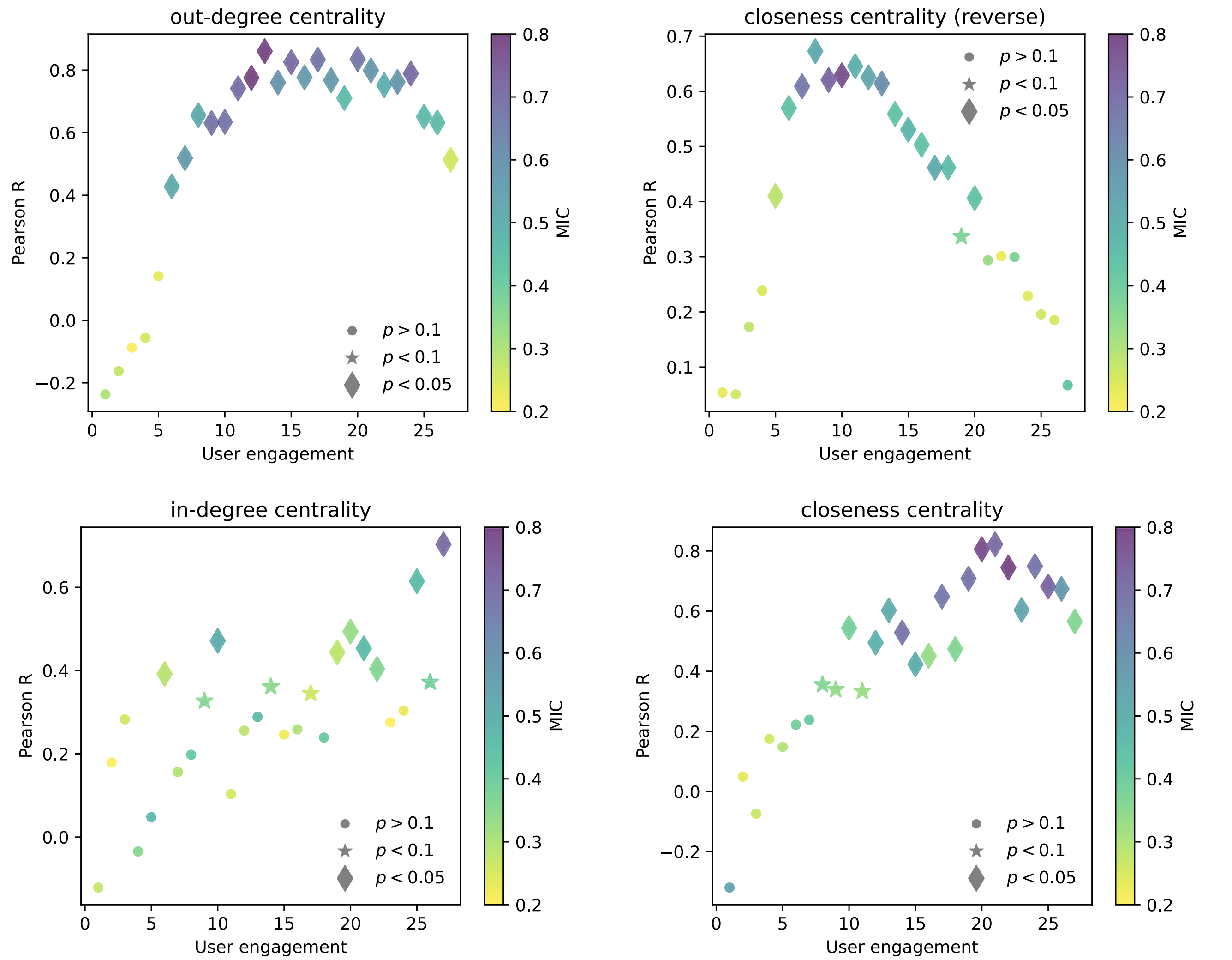

Fig. 4 The association between the subgroup's network centrality and the overall engagement. We measure the association with the Pearson correlation coefficient and the MIC scores. MIC scores are depicted using color, while the statistical significance of the Pearson correlation coefficients is indicated by markers. The average relative centrality of the subgroup is measured by out-degree centrality (*upper left*), closeness centrality on the reverse network (*upper right*), in-degree centrality (*bottom left*), and closeness centrality (*bottom right*). The average relative centrality is the subgroup users' average centrality divided by the average centrality of all nodes in the network. The use of the relative measure enables cross-temporal comparisons of network measures.

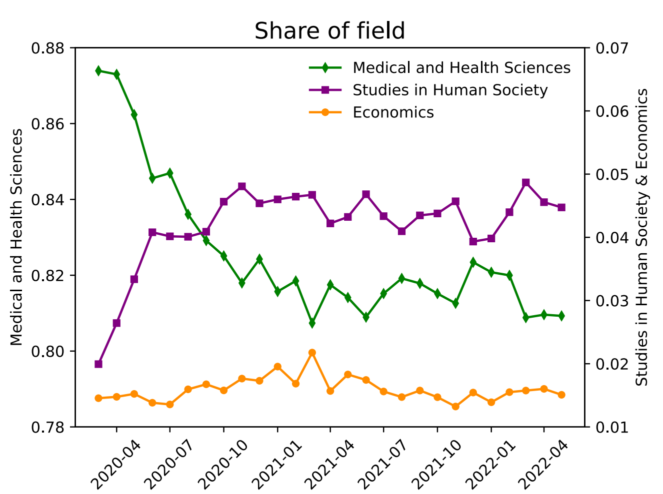

Fig. 5 The evolution of retweet proportion in three research fields. These three research fields include medical and health sciences (green), studies in human society (purple), and economics (yellow).

we select a period spanning 22 months from September 2020 to May 2022 for our data analysis.

Figure 6 shows that low-engagement users have a significant correlation and high MIC score between their engagement with a specific topic and their retweet trends across three fields. Thus, when their overall communication trends fluctuate, their engagement with three research fields (i.e., medical and health science, human society, and economics) also changes. Specifically, low-engagement users show a positive correlation (with a Pearson correlation coefficient of around 0.5) in medical and health science. In contrast, negative correlations exist in the human society and economics fields (with Pearson correlation coefficients around −0.6 and −0.5, respectively). A decrease in the spread of low-engagement users results in a shift in their attention from healthcare to the impact of infectious diseases on society and the economy and vice versa.

On the contrary, the information preferences of the high-engagement user were not significantly associated with their retweet trends. Notably, users with engagement levels of 25 and 27 have a significant positive correlation in the medical and health science fields. We conjecture that since these high-

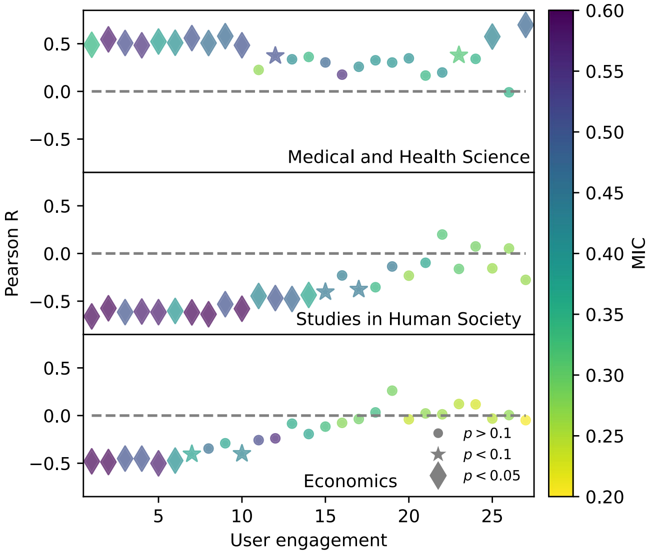

Fig. 6 The association between the subgroup's engagement in three fields and the overall engagement. We measure the association using the Pearson correlation coefficient and the MIC score. MIC scores are depicted using color, while the statistical significance of the Pearson correlation coefficients is indicated by markers. The three fields include medicine and health sciences (*upper*), studies in human society (*middle*), and economics (*lower*).

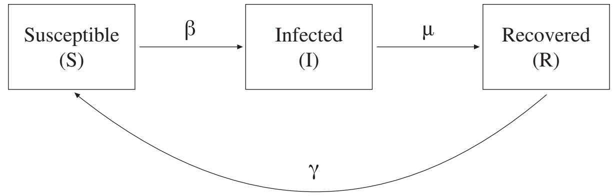

Fig. 7 The SIRS model. There are three states: Susceptible (S), Infected (I), and Recovered (R). Accordingly, there are three parameters in the model: the transmission rate β determines the probability of transmission from an infected individual to a susceptible one; the recovery rate μ depicts the frequency with which an infected individual recovers and returns to the susceptible state; the re-susceptible rate γ determines the probability of individuals returning to the susceptible state after entering the recovered state.

engagement users are deeply embedded in scientific fields, they will not be easily distracted by the outside world.

The findings highlight the vital need to understand the mechanisms of public engagement. The results suggest that low-engagement users are susceptible to the topic of scientific research. If so, a decrease in low-engagement users' engagement with science communication could worsen the issue of inadequate health information among them, thereby increasing their vulnerability to infectious diseases. Taken together, it is crucial to improve public engagement in disseminating scientific research during the pandemic.

Model description and simulation. Our model draws inspiration from Twitter users' engagement and disengagement with science communication. We find that low-engagement participants display a high sensitivity compared to their high-engagement counterparts. On the one hand, low-engagement participants exhibit faster propagation during information dissemination's early and peak phases. On the other hand, when the novelty of science information diminishes over time, low-engagement participants would quickly leave science communication and redirect their attention to other emerging trending topics. In contrast, high-engagement participants are characterized by a slower pace of engagement. They initiate their engagement with science communication gradually, and their disengagements also unfold at a slower pace. In a nutshell, the online community of science communication is like a solid camp guarded by flowing soldiers.

In the context of public engagement, it is commonly argued that the general public's passive involvement in scientific matters is due to a lack of interest in scientific information. Therefore, the enhancement of public engagement relies on increasing public interest in scientific information. Nonetheless, our analytical findings offer evidence that deviates from this prevailing perspective. Our results suggest that, compared to highly engaged users, low-engagement users demonstrate greater sensitivity towards scientific information, characterized by more rapid acceleration during the initial and peak stages. Consequently, their attention flows are more bursty and display a pronounced inclination towards trending topics. These outcomes challenge the idea that low-participation users are generally disinterested in science. If a lack of interest were the primary factor, one would expect to observe a more uniform distribution of attention allocation instead of the explosive attention demonstrated by low-engagement users.

To elucidate this phenomenon, we developed a network-based analytical model that incorporates the concept of engagement heterogeneity by considering information acceptance thresholds. In our framework, whether the public decides to share information depends on their thresholds. Individuals with lower thresholds are more inclined to share information. Furthermore, owing to the inherent limitations on individual attention, individuals tend to disengage with science communication over time. This intrinsic mechanism naturally results in low-engagement users' unsustainable participation.

Model description. We analyze the dissemination of scientific information through a single-layer network where social connections are represented by links and the audience is represented by nodes. With the utilization of the generalized SIRS infectious disease model, individuals within the network are categorized into three groups: susceptible (S), infected (I), and recovered (R). These three states indicate individuals who have not received the information, those who have received the information, and those who are either not interested or not available to receive the information. The SIRS model supposes that individuals no longer interested in the information can regain their interest and move cyclically between the susceptible and the recovered states. This process is depicted graphically for better understanding, as shown in Fig. 7. Our model proposes that each node i in the network has a fixed information acceptance threshold xi, which determines their probability of information acceptance.

We first describe the dynamics of the information propagation process in the absence of external shocks. At each time step, a susceptible individual (S) that comes into contact with an infected individual (I) within the propagation network undergoes information acceptance. The probability of acceptance is given by the transmission rate $\beta = \beta_b \times (I_0 - x_i)$, where β_b represents the base acceptance probability, and I_0 stands for the information's popularity (set at 1). Furthermore, x_i represents the individual's acceptance threshold. A lower acceptance threshold corresponds to a higher probability of individual information acceptance.

Simultaneously, an infected (I) individual i has a probability of leaving the propagation network and entering the recovered state (R) with recovery rate μ. This probability is influenced by the individual's information acceptance threshold, x_i, and base

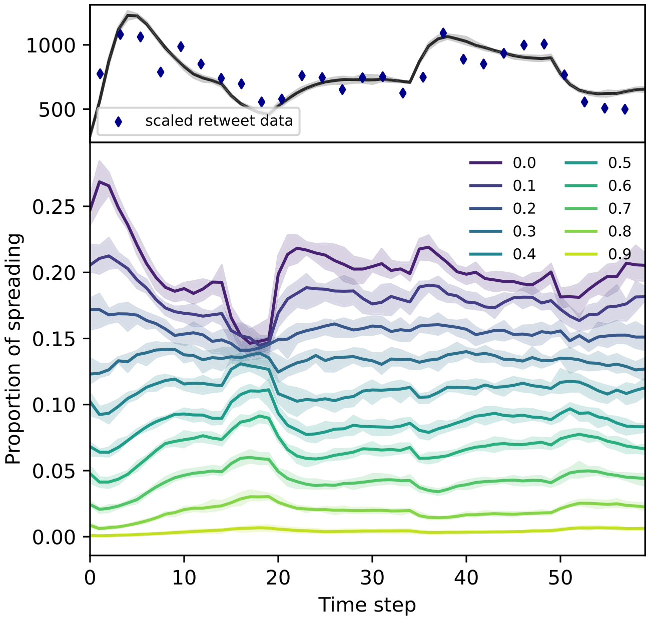

Fig. 8 The evolution of infected nodes in the SIRS model. The evolution of infected nodes was obtained by adding two external shocks to our model (*upper*), roughly consistent with the scaled retweets of scientific research on COVID-19 on Twitter (represented by blue diamond dots). The evolution of the proportion of infected nodes in 10 subgroups to the overall population varies (*bottom*). These two external shocks include a negative event and a positive event, with an intensity I_e of 0.5 and 0.8, and occur during 15 to 20 time steps and 35 to 50 time steps, respectively. We set the base disease spreading rate $\beta_b = 0.2$, the base recovering rate $\mu_b = 0.5$, and the re-susceptible rate $\gamma = 0.15$. The color band shows the 1σ confidence interval of our prediction.

recovery rate, μ_b. Specifically, $\mu = \mu_b \times (I_0 - x_i)$. Due to limited attention capacity and the spread of multiple types of information, individuals with lower acceptance thresholds are more likely to leave the scientific information propagation network than those with higher thresholds.

Individuals in the recovered state (R) tend to lose immunity with the re-susceptible rate γ and return to the susceptible state (S). For the sake of simplicity, the probability of losing immunity γ is assumed to be constant across individuals.

With external shocks, the probability of information dissemination β for a susceptible individual (S) varies with each time step, based on the changes. If there is a negative impact of external shocks on the dissemination of scientific information, the probability of information dissemination β for a susceptible individual (S) is given by the equation:

$$\beta = \beta_b \times \left(I_0 - x_i - \max\left(0,\, I_e - x_i\right)\right)$$

Here, I_0 represents the popularity of the main topic, assumed to be 1, while I_e represents the popularity of the external shock. The likelihood of being affected by this event depends on the individual's acceptance threshold, x_i, and the external shock's popularity, I_e. If the individual's acceptance threshold x_i is greater than the change event's popularity, I_e, their probability of information dissemination β remains unchanged. Otherwise, the probability of information dissemination will decrease.

Similarly, when the external shock has a positive impact on science communication, the probability of information dissemination β for an individual i in the state of ignorance (S) is given by the equation:

$$\beta = \beta_b \times \left(I_0 - x_i + \max\left(0,\, I_e - x_i\right)\right)$$

If the individual's acceptance threshold x_i is greater than or equal to the popularity of the positive external shock, I_e, the probability of information dissemination β for that individual i will remain unchanged. Otherwise, the probability of information dissemination increases.

Our proposed model explains the dynamics of information propagation in a network, considering individuals' varying information acceptance thresholds. The external shock's impact on information dissemination is captured by modifying the acceptance probability according to the external shock's popularity and the individual acceptance threshold. All possible state transitions are shown in Fig. 7.

Simulation. To simulate the information propagation process, we constructed a static scale-free network model utilizing the Barabási-Albert model, which consisted of 5000 nodes with an average degree $< k >$ of 6. Initially, 5% of nodes were randomly selected as spreaders in state I, and the remainder were assigned to state S as ignorant. We varied the threshold for all individuals from 0 to 0.9, selecting ten threshold values at intervals of 0.1. The corresponding distribution pattern exhibited power-law behavior, with an exponent index of 1.5. Consistent with previous analyses than others, most users had lower thresholds, and fewer users had higher thresholds than others.

To investigate the impact of external shock on subgroups with different thresholds, we simulated a simple scenario related to information propagation involving both positive and negative external shocks. We selected the timing and intensity of these events based on trends observed from a real-world dataset analyzed previously (represented by diamond points in Fig. 8). Specifically, the negative external shocks occurred during time steps between 15 and 20 with an intensity of $I_e = 0.5$. In contrast, the positive external shocks happened during time steps between 35 and 50 with an intensity of $I_e = 0.8$. The simulation results depicted in Fig. 8 showed that individuals with low thresholds exhibited a rapid downward trend in transmission proportion at the initial stage, consistent with previously observed real-world phenomena (as depicted in Fig. 2). Furthermore, users with varying thresholds responded to external shocks like that observed in retweets data. During external shocks (at $t \in [15, 20]$), the proportion of retweets by low-threshold users decreases, while high-threshold users exhibit an increase. Conversely, during periods of an overall upward trend caused by external shocks (at $t \in [35, 50]$), the proportion of retweets by low-threshold users increases, whereas high-threshold users experience a decrease.

In addition, as in the previous analysis, we utilize the MIC score and Pearson correlation coefficient to quantify the association between the retweet evolution of individuals (with different thresholds) and the overall retweet trend. As Fig. 9 illustrates, individuals with low thresholds have significant positive correlations, while those with high thresholds have significant negative correlations. Moreover, a significant linear relationship exists between an individual's information dissemination threshold and the correlation coefficient ($R^2 = 0.887$, $p < 0.001$). These findings are consistent with those of earlier analyses of real-world data (as demonstrated in Fig. 3), and individuals with low thresholds have significant positive correlations. In contrast, those with high thresholds have significant negative correlations. Moreover, a significant linear relationship exists between an individual's information acceptance threshold and the correlation coefficient ($R^2 = 0.887$, $p < 0.001$). These results are consistent with our findings with real-world data (as demonstrated in Fig. 3).

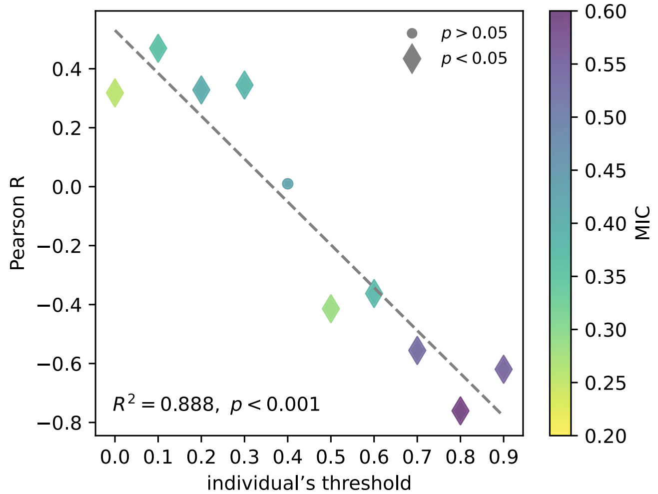

Fig. 9 The association between the subgroup's infection evolution and the overall infection. We measure the association with the Pearson correlation coefficient and the MIC scores. MIC scores are depicted using color, while the statistical significance of the Pearson correlation coefficients is indicated by markers. The subgroup users with a low information acceptance threshold have a more significant infection rate as science communication trends fluctuate. The dashed line represents the OLS linear regression line used to fit the data, with $R^2 = 0.888$ and $p < 0.001$.

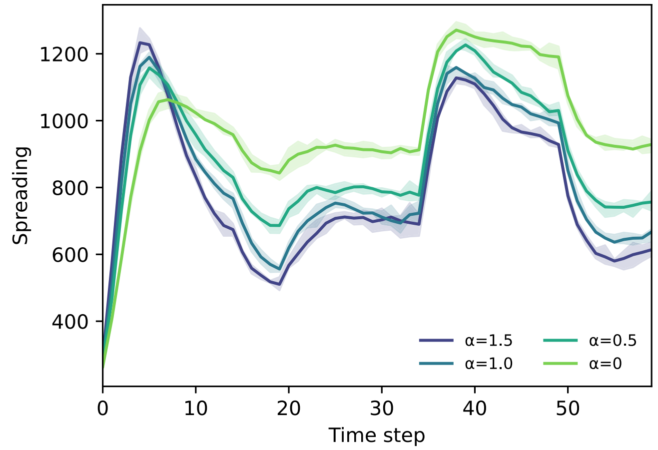

Fig. 10 The evolution of infected nodes in the SIRS model with different power-law exponents. We denote the power law exponent with α and examine the infection evolution with different power-law exponents. When $\alpha = 0$, it represents the network's degree distribution following a uniform distribution. As α increases, the degree distribution of the network becomes increasingly uneven. The colored band shows the 1σ confidence interval of our prediction.

We further explore how engagement heterogeneity influences the fluctuation of science communication. We manipulated the power-law exponent of individuals' threshold distribution and examined various distributions to investigate these effects. Figure 10 shows that when the population had a uniform distribution ($\alpha = 0$), there was less initial spread, and the overall science communication effectiveness (the cumulative number of infected individuals within a time step of 50) improved by 23.9% compared to the previously used long-tailed distribution ($\alpha = 1.5$). Furthermore, networks with uniform distributions exhibited greater resilience to external shocks than those with long-tailed distributions; under negative interference, the decrease in the number of infected individuals before and after the propagation was 11% (when $\alpha = 0$) and 24.3% (when $\alpha = 1.5$), respectively.

Furthermore, we analyze the relative average centrality of subgroups within a simulated dissemination network, considering both outward and inward information propagation (see Fig. 11). The network was constructed at every time step of the simulation, resulting in 60 independent diffusion networks. All nodes in the network are in state I, including inherent state I nodes and newly entered state I nodes. Edges represent the information propagation relationships between the newly entered nodes and the inherent nodes.

In the investigation of outward information propagation, we use measurements of node centrality, such as out-degree centrality and closeness centrality, in the networks. Our findings reveal noteworthy patterns: We observe significant negative correlations between the evolution of relative average centrality (out-degree centrality and closeness centrality) for nodes with an information acceptance threshold below 0.4 and the overall propagation trend. In contrast, nodes with an information acceptance threshold above 0.4 show a significant positive correlation between the evolution of their relative average centrality (out-degree centrality and closeness centrality) and the overall propagation trend. These results are consistent with our analysis of real-world data (see Fig. 4), demonstrating a significant positive correlation for participants with engagement levels ranging from 6 to 20. However, for users with engagements above 20, the Pearson correlation coefficients showed a declining trend in real-world data analysis that our model could not fully explained. We conjecture that this may be attributed to the intricate propagation mechanisms associated with high-engagement participants.

A plausible explanation for this observed phenomenon can be outlined as follows: Since the nodes with a high threshold have a lower probability of leaving, inherent nodes (i.e., the nodes still in the dissemination network) are more susceptible to these high-threshold nodes. Consequently, when science communication declines and decreases the number of newly entered nodes, the proportion of inherent nodes increases. This shift results in a decline in the proactivity of high-threshold nodes (e.g., outward relative average centrality) and an increase in the proactivity of low-threshold nodes. Conversely, when science communication surges, the proportion of the network's inherent nodes decreases. Consequently, there is an upswing in the proactivity of high-threshold nodes and a corresponding decrease in the proactivity of low-threshold nodes.

In investigating inward information propagation, we observe a positive correlation between the in-degree and closeness centrality for subgroups with higher thresholds. These findings closely parallel the results derived from real-world data (see Fig. 4). This congruence suggests that the diminished influence of users with high information acceptance thresholds may be related to the overall decline in propagation. Our model provides a compelling rationale for this phenomenon: During periods of reduced science communication, inherent nodes that persist within the network are more likely to become isolated or receive fewer connections. Conversely, when science communication escalates, these inherent nodes garner more connections. As mentioned earlier, because high-threshold nodes are less likely to exit the propagation, inherent nodes become increasingly dominated by high-threshold nodes. Our results further highlight the potential importance of information acceptance thresholds in shaping PES.

Discussions and conclusions

Drawing on the PES model (Bucchi & Trench, 2014; Burchell, 2015; Kessler et al., 2022), we analyze the Altmetric data of the

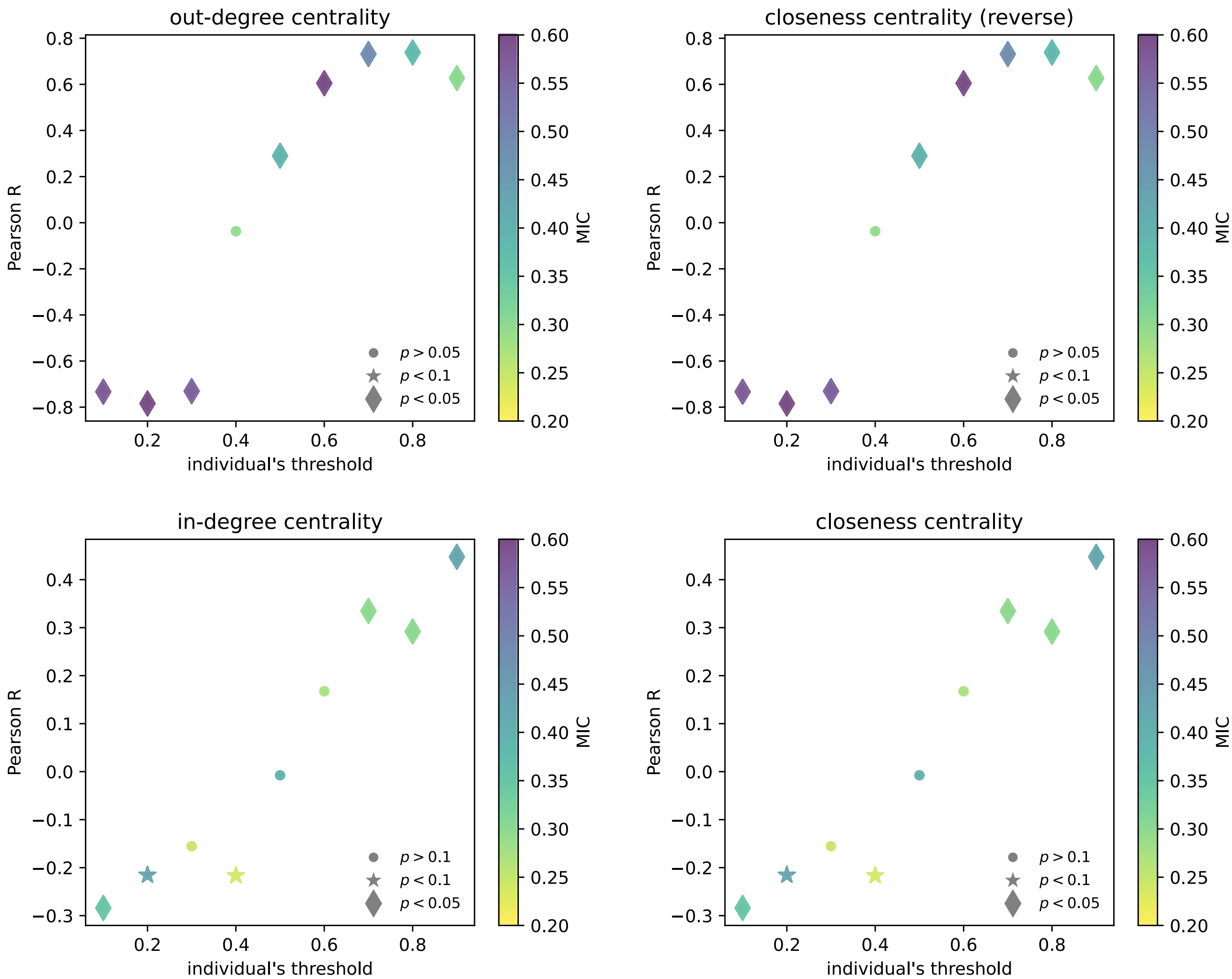

Fig. 11 The association between the subgroup's network centrality and the overall infection. The data shown in the figure is generated using the SIRS model. We measure the association with the Pearson correlation coefficient and the MIC scores. MIC scores are depicted using colors, while the statistical significance of the Pearson correlation coefficients is indicated by markers. The subgroups' average relative centrality is measured by out-degree centrality (*upper left*), closeness centrality on the reverse network (*upper right*), in-degree centrality (*bottom left*), and closeness centrality (*bottom right*). The average relative centrality is the average centrality of a subgroup's users divided by the average centrality of all nodes within the network. The use of the relative measure allows for cross-temporal comparisons of network measures.

spread of scientific research on COVID-19 on Twitter. We contribute to the PES model by conceptualizing participants' engagement heterogeneity and investigating its association with the fluctuation of science communication on social media. First, our analysis of the empirical data reveals that low-engagement users demonstrate greater mobility during these periods compared to high-engagement users. Second, the simulation shows external shocks can significantly influence public engagement with science communication. Third, we identify synchronous shifts in the dissemination preferences of low-engagement users, where their focus moves from medical and healthcare topics toward the social impact of infectious diseases. Our model based on users' engagement threshold and external shocks can well explain the observed disparity in subgroup dynamics. Moreover, our model provides a conceptual framework for understanding the evolution of subgroup centrality in the spreading networks over time.

This study uncovers the underlying link between PES and the flow of public attention. We reframe our study within the research framework of attention flow. In the information age, there is an excess of information supply while human attention is increasingly scarce and has become the new currency intertwined with digital media content production and consumption (Davenport & Beck, 2001; Falkinger, 2007). Consequently, the flow of public attention directly determines which information will spread like a virus and which will sink without a trace. The limited capacity model provides a theoretical foundation for understanding public attention (Kahneman, 1973; Norman & Bobrow, 1975). First, attention and cognitive resources are strongly correlated; Second, cognitive resources are limited. Further, researchers propose that public attention follows a zero-sum game (McCombs & Zhu, 1995; Zhu, 1992). If the public focuses on one issue, it comes at the expense of attention given to other issues. PES is contingent upon the flow and allocation of public attention. Thus, the logic of the zero-sum game of public attention implies that promoting public engagement with science communication may be a significant but never-ending Sisyphean myth.

We emphasize the potential risks of engagement heterogeneity in science communication. Our study finds that reducing the proportion of individuals with lower thresholds in the population can significantly improve the effectiveness of science communication, with a 23.9% increase in propagation efficiency when individuals with different thresholds are evenly distributed. This finding underscores the critical role of targeted and customized information in public health information dissemination, which is more effective in reaching and influencing the audience than non-targeted large-scale movements (Nan et al., 2021). Structural differences in acceptance thresholds may lead to increased polarization when faced with external shocks, further exacerbating low-engagement users' risk concentration and vulnerability.

Our study suggests that the evolution and structure of user engagement are essential for understanding the overall dynamics of online communities. Biddix et al. (2023) examine the discourse about high education on Twitter in the early phases of COVID-19, drawing on the crisis management theory. Using social network analysis, they shows that the discussion with the hashtag *#highered* evolved rapidly (Biddix et al., 2023). Notably, the prominent influencers also change rapidly (Biddix et al., 2023). In contrast, our study reveals that highly engaged users remain relatively stable. The network structure of user engagement also matters. Through analyzing the Twitter networks across three periods of COVID-19, Chong and Park (2021) show that the information carriers on Twitter, including information channels, sources, and messages, were interconnected. Further, Park et al. (2021) reveal that Twitter users tend to engage in clustered discussions about COVID-19 issues without opinion leaders.

Further, institutional actors (e.g., governments) also play a pivotal role in science communication. Tahira (2022) studies how the Korean government employs digital techniques for health diplomacy. Surveying 219 foreigners residing in Korea, Tahira (2022) finds that most respondents believe the Korean government effectively leveraged digital technology during the pandemic. Park et al. (2021) identify the institutional actors who cite COVID-19 research (e.g., the World Health Organization) to analyze policy engagement with science during the pandemic. Western countries demonstrate more effective networking practices in constructing a global scientific network ecology than developing countries (Park & Yoon, 2023). Hence, developing countries deserve more attention. In line with this objective, Vargas Meza & Park, 2023 investigate the spread of scientific research on social media among Spanish-speaking and Caribbean communities. Their findings reveal that these communities predominantly engage with scientific papers on medical and health sciences, and the top scientific authorships are from China (Vargas Meza & Park, 2023).

This study has implications for future research in the following aspects: First, simulation models can effectively help strengthen our data analysis. Our simulation results align with the dissemination outcomes of users with different engagement, including low-engagement users rising more rapidly in the initial stage, being more sensitive to external shocks, and the evolution of subgroup users' centrality in the network. In this sense, ongoing discussions regarding scientific literacy and audience interest often overlook the mechanism of user information acceptance thresholds. Second, dividing participants into subgroups according to their sustainability offers valuable insights for operationalizing public engagement in future scenarios. Third, linking engagement heterogeneity with the fluctuation of science communication also has important implications for future research. Given that the PES model is being criticized for not articulating the outcomes (Kessler et al., 2022), we propose that the fluctuation of science communication as an outcome on the aggregate level deserves more attention.

In all, our theoretical model and empirical analysis contributes to the discussion on public engagement in scientific information dissemination. It offers a novel perspective on how the heterogeneity of users' information acceptance thresholds gives rise to varying degrees of engagement. On the one hand, high-engagement users approach new information with greater caution. Consequently, they engage in disseminating scientific content slower but are also less likely to disengage from science communication. On the other hand, low-engagement users exhibit faster propagation in the early and peak phases of information dissemination but also disengage from science communication at an accelerated rate when the novelty of scientific information diminishes. Instead of emphasizing that there is no country for old members in online communities (Danescu-Niculescu-Mizil et al., 2013), we acknowledge that such fluidity is the essence of online communities and science communication. Our findings are nontrivial in highlighting the role of low-engagement participants. As the proverb goes, a solid camp as iron is guarded by ever-changing soldiers who move like flowing water. The solid camp helps stabilize communities, while flowing soldiers enable the influence of the camp to cross communities.

Data availability

The simulation code employed in this study is available on the Open Science Framework (https://osf.io/8p5am/). The data used in this study can be acquired through Altmetric (https://www.altmetric.com/).

Received: 6 May 2023; Accepted: 27 October 2023;
Published online: 13 November 2023

References

Baram-Tsabari A, Osborne J (2015) Bridging science education and science communication research. J Res Sci Teaching 52(2):135–144. https://doi.org/10.1002/tea.21202

Berry, D, & Widder, S (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front Microbiol 5*. https://doi.org/10.3389/fmicb.2014.00219

Biddix JP, Park H, Collom GD, Bailey MR, Park HW (2023) Discourse about higher education on Twitter in early phases of COVID-19: A crisis management social network analysis. Educ Inf Technol 28(8):9957–9983. https://doi.org/10.1007/s10639-023-11590-2

Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, Schmidt AL, Valensise CM, Scala A, Quattrociocchi W, Pammolli F (2020) Economic and social consequences of human mobility restrictions under COVID-19. Proc Nat Acad Sci USA 117(27):15530–15535. https://doi.org/10.1073/pnas.2007658117

Borgatti SP (2005) Centrality and network flow. Social Netw 27(1):55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Bubela T, Nisbet MC, Borchelt R, Brunger F, Critchley C, Einsiedel EF, Geller G, Gupta A, Hampel J, Hyde-Lay R, Jandciu EW, Jones S, Kolopack P, Lane S, Lougheed T, Nerlich B, Ogbogu U, O'Riordan K, Ouellette CA, Caulfield T (2009) Science communication reconsidered. Nat Biotechnol 27(6):514–518. https://doi.org/10.1038/nbt0609-514

Bucchi, M, & Trench, B (2014) Science communication research: Themes and challenges. In: Bucchi M and Trench B (eds) Routledge Handbook of Public Communication of Science and Technology (Vol. 2). New York, NY: Routledge, pp. 1–14

Bullock OM, Amill DC, Shulman HC, Dixon G (2019) Jargon as a barrier to effective science communication: Evidence from metacognition. Public Understanding Sci 28(7):845–853. https://doi.org/10.1177/0963662519865687

Burchell K (2015) Factors Affecting Public Engagement by Researchers: Literature Review. Policy Studies Institute, London

Burns T, O'Connor DJ, Stocklmayer S (2003) Science communication: a contemporary definition. Public Understanding Sci 12(2):183–202. https://doi.org/10.1177/09636625030122004

Byrne S, Hart PS (2009) The boomerang effect a synthesis of findings and a preliminary theoretical framework. Ann Int Commun Assoc 33(1):3–37. https://doi.org/10.1080/23808985.2009.11679083

Chong M, Park HW (2021) COVID-19 in the Twitterverse, from epidemic to pandemic: information-sharing behavior and Twitter as an information

carrier. Scientometrics 126(8):6479–6503. https://doi.org/10.1007/s11192-021-04054-2
Crane R, Sornette D (2008) Robust dynamic classes revealed by measuring the response function of a social system. Proc Nat Acad Sci 105(41):15649
Danescu-Niculescu-Mizil, C, West, R, Jurafsky, D, Leskovec, J, & Potts, C (2013, May). No country for old members: User lifecycle and linguistic change in online communities. In Proceedings of the 22nd international conference on World Wide Web (pp. 307-318)
Davenport, TH, & Beck, JC (2001). The attention economy: Understanding the new currency of business. Harvard Business School Press
Escandón, K, Rasmussen, AL, Bogoch, II, Murray, E, Escandón, K, Popescu, S, & Kindrachuk, J (2021). COVID-19 false dichotomies and a comprehensive review of the evidence regarding public health, COVID-19 symptomatology, SARS-CoV-2 transmission, mask wearing, and reinfection. *BMC Inf Dis 21*(1). https://doi.org/10.1186/s12879-021-06357-4
Falkinger J (2007) Attention economies. J Econ Theory 133(1):266–294
Freeman LC (1978) Centrality in social networks conceptual clarification. Social Netw 1(3):215–239. https://doi.org/10.1016/0378-8733(78)90021-7
Funk S, Gilad E, Watkins C, Jansen (2009) The spread of awareness and its impact on epidemic outbreaks. Proc Nat Acad Sci USA 106(16):6872–6877. https://doi.org/10.1073/pnas.0810762106
Funk S, Salathé M, Jansen (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. J R Soc Interf 7(50):1247–1256. https://doi.org/10.1098/rsif.2010.0142
Gray DM, Anyane-Yeboa A, Balzora S, Issaka RB, May FP (2020) COVID-19 and the other pandemic: populations made vulnerable by systemic inequity. Nat Rev Gastroenterol Hepatol 17(9):520–522. https://doi.org/10.1038/s41575-020-0330-8
Hart, PS, & Nisbet, EC (2014). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies [Dataset]. In *PsycEXTRA Dataset*. https://doi.org/10.1177/0093650211416646
Hyland-Wood, B, Gardner, J, Leask, J, & Ecker, UKH (2021). Toward effective government communication strategies in the era of COVID-19. *Human Soc Sci Commun 8*(1). https://doi.org/10.1057/s41599-020-00701-w
Kaim A, Gering T, Moshaiov A, Adini B (2021) Deciphering the COVID-19 Health Economic Dilemma (HED): A scoping review. Int J Environ Res Public Health 18(18):9555. https://doi.org/10.3390/ijerph18189555
Kahneman, D (1973). Attention and effort. Prentice-Hall
Kessler SH, Schäfer MS, Johann D, Rauhut H (2022) Mapping mental models of science communication: How academics in Germany, Austria and Switzerland understand and practice science communication. Public Understanding Sci 31(6):711–731. https://doi.org/10.1177/09636625211065743
Maltese AV, Melki C, Wiebke H (2014) The nature of experiences responsible for the generation and maintenance of interest in STEM. Sci Educ 98(6):937–962. https://doi.org/10.1002/sce.21132
McCombs M, Zhu JH (1995) Capacity, diversity, and volatility of the public agenda. Public Opin Quarterly 59(4):495–525
Miah A (2017) Nanoethics, science communication, and a fourth model for public engagement. Nanoethics 11(2):139–152. https://doi.org/10.1007/s11569-017-0302-9
Nan X, Iles IA, Yang B, Ma Z (2021) Public Health Messaging during the COVID-19 Pandemic and Beyond: Lessons from Communication Science. Health Commun 37(1):1–19. https://doi.org/10.1080/10410236.2021.1994910
Nisbet MC, Scheufele DA (2009) What's next for science communication? Promising directions and lingering distractions. Am J Botany 96(10):1767–1778. https://doi.org/10.3732/ajb.0900041
Norman DA, Bobrow DG (1975) On data-limited and resource-limited processes. Cogn Psychol 7(1):44–64. https://doi.org/10.1016/0010-0285(75)90004-3
Park, H, Biddix, JP, & Park, HW (2021). Discussion, news information, and research sharing on social media at the onset of Covid-19. El Profesional de La Información, e300405. https://doi.org/10.3145/epi.2021.jul.05
Park HW, Yoon HY (2023) Global COVID-19 policy engagement with scientific research information: altmetric data study. J Med Int Res 25:e46328. https://doi.org/10.2196/46328
Park, S, Bier, LM, & Park, HW (2021). The effects of infotainment on public reaction to North Korea using hybrid text mining: Content analysis, machine learning-based sentiment analysis, and co-word analysis. El Profesional de La Información, e300306. https://doi.org/10.3145/epi.2021.may.06
Perry, BL, Aronson, B, & Pescosolido, BA (2021). Pandemic precarity: COVID-19 is exposing and exacerbating inequalities in the American heartland. *Proc Nat Acad Sci USA 118*(8). https://doi.org/10.1073/pnas.2020685118
Powell M, Kleinman DL (2008) Building citizen capacities for participation in nanotechnology decision-making: the democratic virtues of the consensus Conference model. Public Understanding Sci 17(3):329–348. https://doi.org/10.1177/0963662506068000
Reshef DN, Reshef Y, Finucane H, Grossman SR, McVean G, Turnbaugh PJ, Lander ES, Mitzenmacher M, Sabeti PC (2011) Detecting novel associations in large data sets. Science 334(6062):1518–1524. https://doi.org/10.1126/science.1205438
Scheufele DA, Krause N (2019) Science audiences, misinformation, and fake news. Proc Nat Acad Sci USA 116(16):7662–7669. https://doi.org/10.1073/pnas.1805871115
Tahira I (2022) Digital technology practices and vaccine campaign in Korea: International perceptions on health diplomacy amid COVID-19 crisis. J Contemp Eastern Asia 21(2):27–46. https://doi.org/10.17477/JCEA.2022.21.2.027
Van Bavel JJ, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, Crockett MJ, Crum AJ, Douglas KM, Druckman JN, Drury J, Dube O, Ellemers N, Finkel EJ, Fowler JH, Gelfand MJ, Han S, Haslam SA, Jetten J, Willer R (2020) Using social and behavioural science to support COVID-19 pandemic response. Nat Human Behav 4(5):460–471. https://doi.org/10.1038/s41562-020-0884-z
Vargas Meza X, Park HW (2023) Information circulation among spanish-speaking and caribbean communities related to COVID-19: social media–based multidimensional analysis. J Med Int Res 25:e42669. https://doi.org/10.2196/42669
Wakefield M, Loken B, Hornik R (2010) Use of mass media campaigns to change health behaviour. Lancet 376(9748):1261–1271. https://doi.org/10.1016/s0140-6736(10)60809-4
Wang W, Liu Q, Liang J, Hu Y, Zhou T (2019) Coevolution spreading in complex networks. Physics Rep 820:1–51. https://doi.org/10.1016/j.physrep.2019.07.001
West, JD, & Bergstrom, CT (2021). Misinformation in and about science. *Proc Nat Acad Sci USA 118*(15). https://doi.org/10.1073/pnas.1912444117
Wu Q, Fu X, Small M, Xu X (2012) The impact of awareness on epidemic spreading in networks. Chaos 22(1):013101. https://doi.org/10.1063/1.3673573
Wynne B (2006) Public engagement as a means of restoring public trust in science – hitting the notes, but missing the music. Public Health Genomics 9(3):211–220. https://doi.org/10.1159/000092659
Yang, B, Shang, K, Small, M, & Chao, N (2023). Information overload: How hot topics distract from news---COVID-19 spread in the US. *Nat Sci Open*, 20220051. https://doi.org/10.1360/nso/20220051
Yang, Y, Zhang, Q, Ruan, Z, Cao, Z, Xuan, Q, & Zeng, D (2020). Effect of heterogeneous risk perception on information diffusion, behavior change, and disease transmission. *Phys Rev. 102*(4). https://doi.org/10.1103/physreve.102.042314
Zhu JH (1992) Issue competition and attention distraction: A zero-sum theory of agenda-setting. Journalism Quarterly 69(4):825–836

Acknowledgements
This research work is funded by the National Social Science Fund of China (Grant No.22BXW032), the Social Science Foundation of Jiangsu Province of China (Grant No. 19JD001).

Author contributions
B.Y. and C.W. designed the research; B.Y., C.W. and N.C performed the research; B.Y. and C.W. designed the model and the computational framework and analysed the data; and B.Y. and C.W. wrote the paper. All authors discussed the results and contributed to the final manuscript.

Competing interests
The authors declare no competing interest.

Ethical approval
This article does not contain any studies with human participants performed by any of the authors.

Informed consent
This article does not contain any studies with human participants performed by any of the authors.

Additional information
Correspondence and requests for materials should be addressed to Naipeng Chao or Cheng-Jun Wang.