

The Psychology of Virality

Steve Rathje¹

Jay J. Van Bavel¹²³

¹New York University, Department of Psychology

²New York University, Center for Neural Science

³Norwegian School of Economics

This paper is now in press at *Trends in Cognitive Sciences*.

Abstract: Why do some ideas spread widely, while others fail to catch on? We review the psychology of information spread, or the psychology of “virality.” Similar types of information tend to spread in many contexts, both online and offline. This is likely because similar psychological processes drive information spread across contexts. We explain how these psychological processes interact with structural features of information environments, including norms, networks, and incentive structures. Surprisingly, widely shared content is often not widely liked—a phenomenon called “the paradox of virality.” We discuss the strengths and limitations of the virality metaphor. We also discuss future directions for the field, such as leveraging recent advances in artificial intelligence to better understand how information spreads across cultures and contexts.

Keywords: Information Sharing, Misinformation, Gossip, Social Media, Diffusion, Virality

Steve Rathje: 0000-0001-6727-571X

Jay J. Van Bavel: 0000-0002-2520-0442

Abstract Word Count: 120

Word Count: 3,989

Acknowledgements: The authors would like to thank members of the Center for Conflict and cooperation for their thoughtful feedback on this paper. The authors gratefully acknowledge financial support from a National Science Foundation grant awarded to JVB and SR (#2334148), a National Science Foundation Postdoctoral Fellowship awarded to SR (#2404649), a Templeton World Charity Foundation Grant awarded to JVB and SR (TWCF-2023-31570, <https://doi.org/10.54224/31570>) and another to JVB (TWCF-2022-30561; doi.org/10.54224/30561).

Highlights

- Similar types of information (e.g., negative, high-arousal, or moral information) tend to spread both online via social media and offline via gossip.
- The information as virus metaphor is often used to describe the process of information spread (e.g., some types of information are more contagious than others, and “superspreaders” contribute to most (mis)information spread). This metaphor has pros and cons.
- Similar underlying psychological processes (e.g., preferential attention to negativity, social motivations, etc.) drive the spread of information across contexts.
- These psychological processes interact with structural features of information environments (e.g., norms, networks, and incentive structures).
- Future work can use recent advances in artificial intelligence to explore how information spreads across different languages, cultures, and contexts.

Toward a Psychology of Virality

Human communication evolved in small, face-to-face social networks [1]. Throughout history, technology has profoundly changed how information is created and shared—for better and for worse. For instance, many argue that the printing press sparked the scientific revolution. At the same time, many of the bestselling books immediately after the invention of the printing press were religious extremist texts and witch-hunting manuals [2]. Similarly, social media platforms like Facebook promised to “connect the world” and “bring people closer together”¹. But, they also facilitated the spread of divisive content [3]. Artificial intelligence is poised to make many aspects of life easier, but also makes it easier to generate misinformation [4,5].

Yet, the extent to which technology has fundamentally altered the spread of information is still unclear. While negativity and moral outrage “go viral” online [6–8], offline gossip is common in regular conversation [9] and often consists of negativity [10,11] or discussion of moral transgressions [12]. While people are more inclined to click on negative news on the internet [13], negative news has always sold well, leading to the popular adage “if it bleeds it leads”². These common patterns suggest that basic psychological processes may lead similar types of information to go viral in most contexts—online and offline. Despite these commonalities, modern technologies may still influence the substance, scale, and spread of information.

This article reviews the psychology of **virality**, or the **psychological factors** and **structural features** that shape information spread in both online and offline contexts. We synthesize the interdisciplinary literature on information spread, noting that similar features predict information spread (or “virality”) across numerous contexts. This can be partially explained by underlying psychological factors (e.g., our attraction to high-arousal and negative information). However, structural features (e.g., norms, networks, and incentive structures) also interact with these psychological factors to influence information spread. While prior work on information spread has been explored in distinct literatures (e.g., literatures on social media, gossip, word-of-mouth marketing), we integrate these diverse literatures to explore key similarities and differences in how information spreads across contexts.

Understanding the psychology of virality is important because the information that spreads widely or goes “viral” shapes the information diet people consume on a daily basis. The

main way people access news today via social media [14], meaning that the primary way people access news is by consuming “viral” content that is widely shared or amplified by algorithms. Similarly, in the offline world, people learn about “viral” information that spreads via their in-person networks. The information people consume online and offline can have important consequences for how people vote, feel, and behave [15,16].

We use the term “the psychology of virality” to acknowledge the ubiquity of the virus metaphor in describing information spread. This metaphor has been widely used by academics, popular science writers, and the general public [17,18]. For example, the term “infodemic” rose to prominence during the 2020 pandemic as a way to describe viral misinformation [19]. Beyond mere metaphor, researchers have also used epidemiological models to model information spread, suggesting that, in some ways, information does spread like a virus [20–22]. Despite its popularity and utility, this metaphor (like all metaphors) is imperfect and fails to capture some nuances of information spread. See **Box 1** for a discussion of the strengths and limitations of this metaphor.

What Goes Viral on Social Media?

Social media and the internet has made it much easier to study the psychology of online virality. Indeed, the wealth of text data on the internet launched the field of computational social science [23], which led to numerous large computational studies using “big data” from the internet. Many of these studies have focused on the study of information **diffusion** (or “virality”) [24], exploring how certain properties of information (e.g., emotional language) relate to measures of engagement (e.g., social media shares). Note that these studies typically do not aim to predict extreme virality (as the term “go viral” may suggest), likely because viral outliers are rare and thus difficult to forecast. Instead, they tend to focus on factors that are correlated with engagement.

Studies on the spread of information on social media have focused on a variety of different psychological constructs. For instance, research has found that the following variables predict information spread on social media: out-group animosity [6,25–28], incivility [29], negativity [13,27,30,31], moral outrage [7,8,32], anger [33], misinformation [34,35], indignant disagreement³, and high-arousal information [36,37]. However, it is likely that several of these

psychological constructs overlap. For example, moral outrage is negative, high-arousal, hostile, frequently about out-groups, and correlates with misinformation [32]. Thus, rather than focusing on a number of separate constructs, a more parsimonious framework is needed to organize the literature on social information spread.

The **circumplex model of affect** can help us understand the literature on information spread on social media [38]. The circumplex model is a psychological theory that suggests that all emotions can be represented as points along two dimensions of affect: **valence** and **arousal**. For instance, anger is a high-arousal (or intense) negative emotion, sadness is a low-arousal negative emotion, excitement is a high-arousal positive emotion, and calm is a low-arousal positive emotion. This model is popular in the emotion literature (though there are many alternate models [39]), and is considered by many to be better than focusing on “basic” emotions, which vary widely across cultures [40]. We map several of the key findings from the literature of social media virality on the dimensions of the emotion circumplex. See *Figure 1* for a map of several key studies in the virality literature placed on the circumplex model of affect.

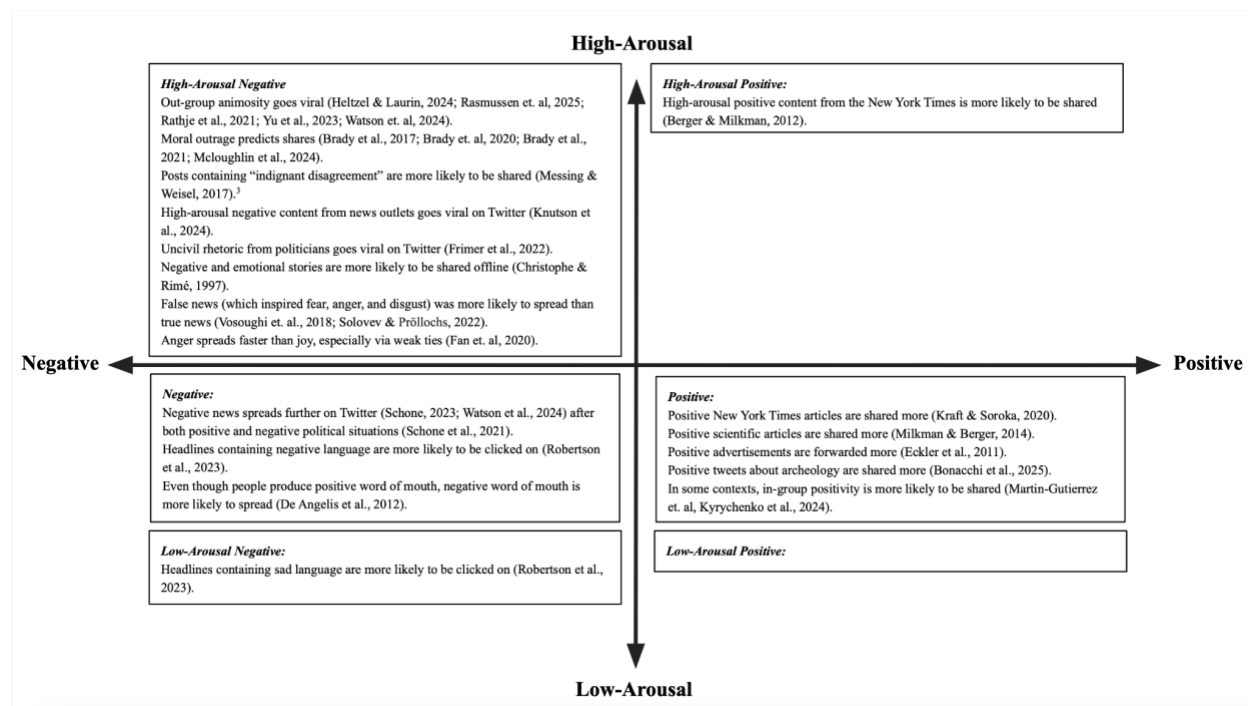


Figure 1. High-arousal and negative content goes viral. We mapped the key studies cited in this article [7,8,13,25–37,41–48] according to the emotional circumplex, which represents various emotions according to two dimensions: arousal and valence. As shown, most studies find that high-arousal negative emotions are more likely to spread or go viral. However, there are a number of exceptions, with a few studies finding that positivity or low-arousal negativity are likely to spread in some contexts. Notably, we found no studies in which low-arousal positive

information is most likely to spread. This organization of key studies in the virality literature illustrates some striking patterns in the type of information that is most likely to spread across contexts—likely due to common psychological processes driving information spread. Note: this figure just shows a subset of key studies in the literature cited throughout this article rather than a comprehensive review of all studies on information spread.

As shown in *Figure 1*, there is a clear pattern in the literature on virality. Most studies suggest that high-arousal negative emotional expressions (such as hostility, moral outrage, out-group animosity, and indignant disagreement) go “viral.” There are no studies (to our knowledge) demonstrating that low-arousal positive content is most likely to spread. Thus, there appears to be some consistency in the wide range of studies in terms of what goes viral, which can be attributed to similar psychological processes driving information spread across numerous contexts.

There are, however, a few exceptions to this pattern. For instance, studies have found that positive New York Times articles [44], advertisements [45], science findings [49] and tweets about archeology [46] are more likely to be shared online. Notably these studies look at non-political contexts, suggesting that there may be important moderators at play (e.g., whether a message is about politics). Other work has found that negativity [31] and toxicity [50] achieve more virality among public figures than regular users, further demonstrating the role of contextual moderators. Many of these conflicting findings can be explained by structural features of environments in which information is shared (including norms, networks, algorithms, incentive structures, etc.). We aim to reconcile these conflicting findings and explain how psychological processes and structural features interact to shape information spread.

Psychological Processes Underlying Virality

In a crowded information environment in which many different pieces of information are competing for our attention (such as a social media news feed), one would expect that the most attention-grabbing information would be the most likely to be shared. Similarly, in the offline world, one would expect the most memorable pieces of information to be spontaneously brought up in in-person conversations. Indeed, research suggests that negative and high-arousal content is not only more likely to be shared, but is also more likely to be attended to and remembered [51,52], which may help explain the viral spread of this type of content online and offline [41].

Our attraction to high-arousal negative information likely has deep evolutionary origins [53]. It is often adaptive for people to pay attention to threatening information, since it is likely to hurt our chances of survival [54]. Social media algorithms appear to take advantage of our attraction to threat in order to attract our attention and keep us engaged [55]. However, our evolutionary attraction to threat appears to be mismatched with the affordances of the online world [56], because it is likely not useful to be made constantly aware of distant threats that are constantly happening all over the globe.

Beyond attention, other psychological processes, such as motivation [57], likely drive the spread of information as well [58]. Researchers have explored how various motivations, such as identity-based motivations [59], status motivations [60], reputational motivations [61], affiliation-based motivations [62], certainty motivations [63], and even a motivation to create chaos [64] all drive the sharing of certain types of (mis)information. Similarly, classic social psychological work from the 1940s (long before social media) suggested that status-based and uncertainty-reduction motivations drive the spread of offline rumors [65].

Sharing is a fundamentally social activity, and what people share online will likely reflect social and reputational motives. Indeed, neuroscience studies have found that brain areas implicated in social processing are activated in response to viral content [66,67]. There is also evidence that more positive content is more likely to be shared than privately viewed, perhaps in part due to self-conscious self-presentation motives to present oneself in a positive light [44]. Further, people attend to a more politically diverse set of information than they share, suggesting that social identity motives can constrain the set of information people are likely to share [68].

People's social motivations may at times conflict with accuracy motivations and lead to sharing of misinformation [59]. For instance, making social motivations (such as our motivation to conform with our in-group) salient by making people think about whether a news story would be liked by members of their political party decreased people's ability to correctly discern the accuracy of that headline [69]. By contrast, interventions that make accuracy motivations salient lead people to share more accurate information [69–72].

A final reason people may share information may be that sharing intrinsically *feels good*—potentially as an evolutionary byproduct of our need to communicate important information to others. Neuroscience research indicates that people find sharing information to be rewarding [73]. This desire to share may stem, in part, from our need to feel a sense of **shared**

reality (or perceived commonality of inner states) with other people [74]. One perspective suggests that people’s personal need to disclose negative information turns people into “news broadcasters” whose stories inform the community of important events [75].

The Paradox of Virality

The modern spread of information is facilitated by algorithms that show people content they are likely to engage with. Facebook has argued that their news feed ranking algorithm shows people what they “want to see”⁴. Does the fact that people are more likely to engage with high-arousal negative content imply that they like this type of content? No. A number of studies suggest that people report strongly disliking this content. One study found that, even though people are aware that negative and high-arousal content goes “viral” on social media, the vast majority of people across the political spectrum reported that they do not want this to be the case. Instead, they would prefer that more positive, nuanced, and accurate content goes viral [24] (see *Figure 2*). Other studies have illustrated similar findings: for instance, most people do not approve of divisive tweets [28] or uncivil politicians from any side of the political spectrum [76].



Figure 2. People know divisive content goes viral, but do not want it to. This figure shows the results of a survey of a representative sample of 511 Americans. Participants answered questions about categories of content they think do (vs. should) go viral on a scale of 1 “strongly disagree” to 7 “strongly agree.” Even though participants thought that various forms of high-arousal negative content spread widely on social media, they did not want this to be the case. Instead, they think that positive, accurate, and nuanced content should spread on social media. This is called

the “paradox of virality,” or the idea that widely shared content is often not widely liked. Figure 2 adapted from Rathje et al. (2023) [24].

This gap between people’s stated preferences and behavior is called the **paradox of virality**, or the idea that widely shared content is not necessarily widely liked⁵. This gap constitutes a paradox because it appears to be contradictory that people would contribute to the spread of information they do not think should spread widely. There are a few potential explanations for this paradox, which we describe below.

A revealed preference for division. One explanation might be that people are either unaware of or unwilling to report their actual preferences. In other words, people’s social media behavior may reflect their revealed (as opposed to stated) preferences [77]. Many economists have suggested that people’s “revealed” preferences are more meaningful than their stated preferences [78]. And indeed, in designing the news feed ranking algorithm, social media platforms like Facebook seem to prioritize people’s revealed preferences or behavior rather than their stated preferences.

Yet, people’s stated preferences may be worth listening to. For instance, 70% of smokers say they would like to quit smoking [77]. The reason many smokers continue to smoke is because of addiction—not a “revealed preference” for smoking. Similarly, experiments have suggested that a substantial portion of social media usage can be attributed to self-control failures (akin to “addiction”) [79]. In the context of addictive products (such as social media, junk food, or cigarettes), it is important to pay attention to stated preferences, since people may not have the ability to align their behavior with their true preferences.

Social media companies can prioritize people’s stated preferences to design better algorithms [55]. One series of large-scale field experiments found that unfollowing hyperpartisan influencers reduced out-party animosity and made people feel better about their social media feeds, without reducing social media engagement [80]. This suggests that aligning people’s social media feeds with their stated preferences has positive consequences.

Incentive Structures. Another potential explanation for the paradox of virality is that the incentive structure of many social media platforms encourages people to share information they do not personally like to see. People may share content for social approval [8]. They may also dwell on content because they feel they cannot look away—just as people might feel like they cannot look away from a car crash on the side of the road. Since social media algorithms amplify

attention-grabbing content, content that people cannot look away from (but do not like) may be more likely to go “viral” on social media.

Superspreaders. A third potential explanation for why widely shared content is not widely liked is that a small number of passionate “superspreaders” may be disproportionately responsible for sharing it. One estimate suggests that about 0.1% of Twitter users share about 80% of misinformation on Twitter [81,82]. A similar study found that about 3% of toxic Reddit users share about 33% of content on the platform [83]. While the vast majority of people disapprove of divisive tweets, the small fraction of people who frequently interact with politicians on Twitter (and thus share the majority of political posts) approve of divisive posts [28].

Thus, the preferences of these “superspreaders” who frequently share social media content may be very different than the vast majority of people. Indeed, a survey of online participants found that most people prefer to share information that is positive about their in-party than negative about their out-party [84], despite out-party negativity going viral [6,25–28]. One reason for this discrepancy between survey results and social media data may be that a small number of people are contributing to the majority of online sharing behavior [85].

Further illustrating this “superspreader” phenomenon, one study found that hostile tweets received more “likes” but a lower “likes-per-retweet” ratio [29]. This lower likes-per-retweet ratio suggests that hostile tweets were relatively less liked when accounting for how much they were retweeted and seen. A lower likes-per-retweet ratio was correlated with greater disapproval of the tweet from human judges—indicating that it is a proxy of audience approval. Thus, a high “likes” count may give the false impression that a tweet is widely liked, when in reality only a small fraction of those who see it like it.

Structural Features Shaping Information Spread

Is social media and the internet unique in its tendency to amplify the high-arousal negative content, or does this content also go “viral” in the offline world? Research has found that negative [42], emotionally arousing [75], and moral gossip [12] is likely to spread widely in the offline world. Positive gossip is exceedingly rare [10]. Furthermore, the word of mouth marketing literature suggests that, even though people are more likely to produce positive word-

of-mouth marketing, negative word-of-mouth marketing is more likely to spread [43]. This observation that people produce positivity but share negativity is found in the online world as well: multiple studies suggest that politicians talk about the in-group more, even though out-group animosity spreads more [6,25]. And while the average social media post is surprisingly positive [86], negativity is more likely to spread [30,31]. Overall, these studies reveal that there are striking similarities in the type of information that is likely to spread online and offline.

Studying how information spreads offline, or studying how information spread historically (see **Box 2**), can help us uncover the **structural features** that shape information spread. Returning to the information as a virus metaphor, psychological factors—such as the negativity or arousal of information—can make information more contagious. However, just as diseases are more likely to spread in certain contexts (such as a crowded room), structural features of online or offline environments can also shape information spread. We briefly review some structural factors that may shape information spread below.

Face-to-face interaction (or lack thereof). A common assumption is that people’s behavior becomes more hostile behind the anonymity of the internet. However, one study suggests that this assumption may be overstated: people who are hostile online are also hostile offline, as opposed to online anonymity suddenly making non-hostile individuals hostile [87]. Social media may instead feel more hostile because hostility is more visible, potentially because of algorithms [6], the size of online networks enabling “superspreader” dynamics [81], and other factors. This study supports the idea that the offline and online world are similar in underappreciated ways.

Despite this, the lack of face-to-face contact in online spaces may still have important consequences. For instance, one study found that conversations about controversial political issues were less productive when they happened over text (rather than face-to-face or via video call) [94] suggesting that certain features of online discourse might be more hostile.

Algorithms. Online environments, unlike offline environments, are mediated by algorithms. This can have both positive and negative consequences: while some work suggests that algorithms can filter out uncivil language and low-quality news [89], the fact that algorithms amplify attention-grabbing content means that they can also amplify high-arousal negativity and other forms of content people do not want to see but is nonetheless engaging [24]. Indeed, a

recent audit of the Twitter/X algorithm found it amplified political content that did not align with people's stated preferences, such as emotionally-charged content and out-group negativity [55].

Design Features. Other design features besides algorithms alone may further impact information spread. As one example, some have hypothesized that the ease with which information can be reshared without friction facilitates the spread of polarizing and false content [90]. Indeed, a large field experiment conducted in collaboration with Meta found that eliminating the “reshare” feature reduced the amount of false news and political content that spread [91]. Thus, decreasing the ability for content to go “viral” quickly may be a solution for reducing the spread of misinformation or extreme content. Indeed, WhatsApp has implemented limits on how much a message can be forwarded to try to slow the spread of misinformation [92]. Further, providing some other way for people to signal their disapproval of posts, such as via a “dislike” or “distrust” button [70,93], would reduce the virality of widely disliked content.

Norms. For decades, social psychologists have noted the power of social norms in shaping behavior [94]. Norms also powerfully shape online behavior. For instance, one study found that norm learning shapes the sharing of moral outrage online [8]. One also might expect that the norms of various social media platforms might differ. Indeed, recent research suggests that, on more liberal platforms (such as BlueSky), more liberal leaning news sources go viral, and vice versa for conservative platforms (such as Truth Social) [95]. The norms within a platform can also rapidly change. For instance, after Elon Musk's purchase of X (formerly Twitter) and implementation of numerous design changes, people began posting more low-quality news on social media—potentially because of a change in norms (as well as design features, algorithms, and other structural features) [80].

Network Size & Structure. Another reason why online information spread is different from offline information spread is that the online world has much larger social networks. The sheer size of online networks allows for “superspreader” dynamics that could not exist in the offline world. For instance, while an individual can spread a harmful rumor to dozens of their offline connections, the spread of this rumor offline could not compare to the spread facilitated by an account with millions of followers. One study found that, while some people spread more gossip than others (implying the existence of gossip superspreaders), those who gossiped frequently had fewer friends—suggesting that gossip superspreaders might have limited reach

offline [96]. This contrasts with the online world, where superspreaders with the most extreme voices typically have the broadest reach [97].

Beyond size, other network properties may impact virality. For instance, studies have found that weak social ties are more responsible for the spread of anger [33] and novel information [98]. The online world has many more weak ties than the offline world, which may substantially influence the amount of angry and novel information people are exposed to online. Social media **echo chambers** may also impact the diffusion of information [99]. For example, moral and emotional language predicts the sharing of information *within* (but not between) political echo chambers [7]. Though it should be noted that echo chambers exist offline as well, and there is debate over the extent to which online versus offline echo chambers differ [100].

A Global Perspective on Virality

One limitation to most literature on the psychology of virality is that most work is based on English-language data from Western contexts. While there may be some universal aspects of information that attract attention and lead people to spread information, there are important cross-cultural differences. With the advent of **large language models**, it has become easier to analyze data from multiple different languages and cultures around the world and explore cross-cultural differences in the spread of information [101,102].

Some prior studies have observed cross-cultural differences in information spread. For instance, while high-arousal negativity tends to be more “contagious” in a Western, individualist context (i.e., the United States), high-arousal positivity tends to be more contagious in an Eastern, collectivist context (i.e., Japan) [103]. Additionally, while out-group animosity (or more broadly, identity-related content) has repeatedly been shown to go viral in the United States [6,25–27,104], this may be partially due to the fact that the United States is a unique political context, marked by particularly high levels of political sectarianism [105]. Indeed, in-party love spreads more efficiently on Twitter in Spain than out-party hate [47]. Another study, looking at the Russia-Ukraine war, found that out-group animosity appeared to be a stronger predictor of virality in Ukraine before the Russian invasion, but in-group solidarity surged as the strongest predictor of virality *after* the Russian invasion [48]. This suggests that external threats may be a structural feature shaping information spread. While most studies on virality have focused on one

culture at once, **Box 3** suggests that future work should analyze and compare many countries and cultures simultaneously to examine the key variables driving cross-cultural differences.

Concluding Remarks

We integrated diverse literatures—including psychology, neuroscience, communications, political science, computer science, network science, marketing, and more—on the psychology of virality. While this topic has recently been explored extensively in the context of social media, our review finds that similar types of information tend to spread across contexts—online, offline, and historically. We outlined how cognitive and social processes (such as attention, motivation, etc.) interact with structural features (such as algorithms, incentives, network structures, etc.) to shape information spread. To better understand the interplay of these factors in shaping virality, future work should explore what goes viral across language, cultures, and contexts (see “Outstanding Questions”). Understanding the psychology of virality is crucial, because what goes viral dictates what people consume in their information diets, and evolving technologies (such as social media, artificial intelligence, etc.) are rapidly changing how information spreads.

Box 1: A Viral Metaphor. While phrases such as “going viral” became popular in the digital era, the spread of ideas has been described as a virus long before the internet. In the 1970s, Richard Dawkins famously likened the spread of **memes** (or units of cultural transmission, such as ideas, behaviors, or styles) to the spread of viruses [106]. There are many ways in which information spreads like a virus online and offline. For example, A) just as some viruses are more contagious than others, some forms of information spread more easily than others [24]. B) just as diseases have symptoms, information exposure has symptoms, or psychological and behavioral consequences [107]. C) just as viruses can mutate over time, and some mutations of a virus may become more (or less) contagious, information also mutates over time, with some mutations making the information more (or less) contagious [65]. D) just as **superspreaders** disproportionately contribute to epidemic spread, superspreaders also disproportionately contribute to the spread of certain types of information (e.g., misinformation on social media) [81]. E) just as a crowded room is conducive to the spread of a virus, some environments (e.g., some social media platforms) might also be conducive to the spread of certain types of information [24]. Finally, F) Just as interventions (masks, vaccines, etc.) are used to stem the spread of disease, interventions can also be used to create immunity against misleading or harmful information [108].

However, some scholars argue that this metaphor is misleading. For example, information, unlike viruses, often spreads purposefully [109], spreads faster within (but not across) identity groups [110], and spreads through less connected individuals [111]. Others argue the virus metaphor ignores human agency, since people do not passively consume information, but interpret, transform, or ignore information [112]. While this may be thought of as a weakness of the metaphor, others suggest that our tendency to ignore and resist information is a feature of our “mental immune systems,” expanding on the metaphor further [113]. Thus, there are pros and cons to this metaphor. Because metaphors shape how we think [114], they can be useful for communicating to a broad audience, theorizing and generating new hypotheses, but also risk misleading us. Rather than endorsing or rejecting the metaphor of “virality,” we acknowledge its significant influence in academic and popular discourse. Since the metaphor itself is “viral,” it is difficult to ignore in discussions of information spread.

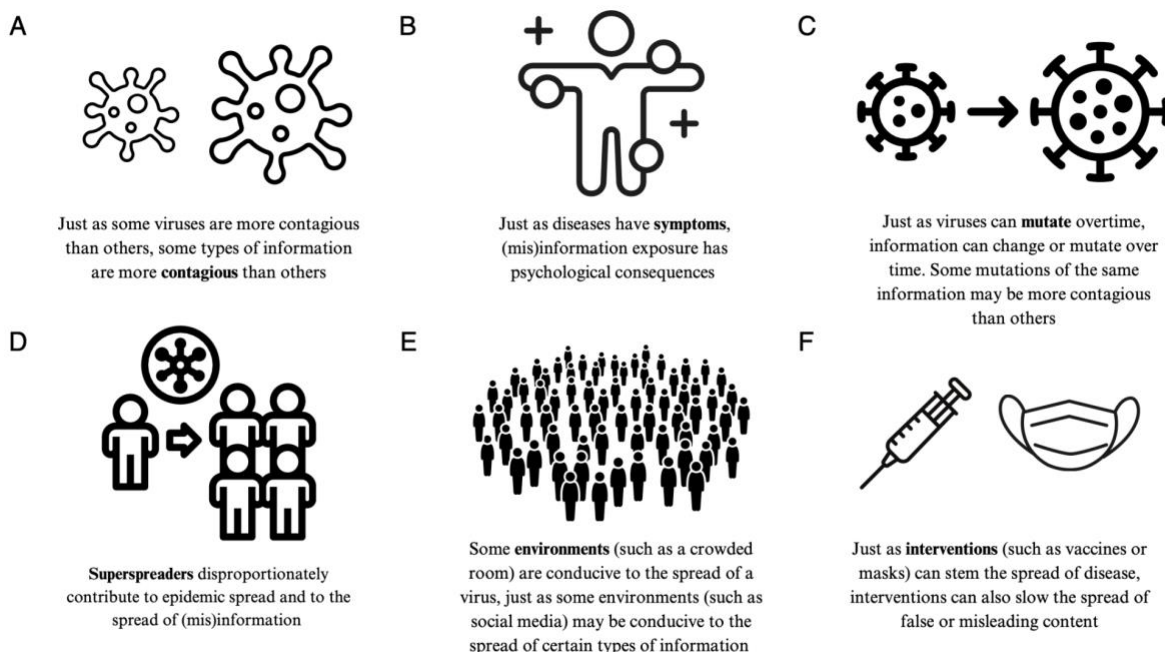


Figure I. Expanding on the virus metaphor. The metaphor of information as a virus can be used as an organizing framework for understanding the psychology of information spread (or “virality”). This metaphor can be extended at length. For instance, like viruses, *A*) some types information are more “contagious” than others, *B*) (mis)information has “symptoms,” or psychological consequences, *C*) information can mutate over time, *D*) “superspreaders” disproportionately contribute (mis)information spread, *E*) some environments are more conducive to the spread of certain types of information than others, and *F*) interventions can be employed to stem the spread of harmful information. We use the information-as-virus metaphor as an organizing framework for the literature on virality.

Box 2. The Historical Spread of Information. *Did high-arousal negativity go “viral” long before the creation of social media and the internet? If so, it would suggest something about human psychology rather than the online world per se. Indeed, historical accounts seem to suggest that high-arousal negativity always went “viral.” For instance, the phrase “if it bleeds, it leads” is believed to have been first used in the 1980s to describe the practice of leading with negative news to attract attention¹. The term “yellow journalism” was used in the 1800s to refer to the practice of creating sensationalistic stories to attract attention [115].*

Historical psychology can help us explore this question empirically [116]. Recent papers have aimed to analyze historical data (e.g., digitized books and headlines) to examine changes over time. Similar methodology could be used to see whether high-arousal negative language has increased over time, potentially in response to the incentive structure of modern technology.

Some articles have used these methods to explore how the spread or prevalence of content has changed throughout history. For instance, one paper found news headlines have expressed more high-arousal negative emotion (specifically, anger, fear, disgust and sadness) over the period 2000-2019, which coincides with the rise of social media [117]. Similarly, another study found that politicians have been using more incivil language on Twitter from 2009 onwards, and this increase in incivility was found to be partially driven by reinforcement learning in which politicians expressed more incivility in response to more positive feedback (e.g., “likes” and “retweets”) [29]. The same study found that this increase in incivility was not found in offline political debates or congressional speeches, suggesting a mismatch between the online in offline world. However, there has also been a general increase in offline feelings of political sectarianism in representative surveys of United States citizens [105].

However, another recent study found persistent patterns of toxicity across 34 years of social media data, suggesting that while social media has evolved tremendously, conversation patterns have remained remarkably consistent [118]. Other work has found that belief in conspiracy theories has not increased over time and that the advent of social media has not coincided with an increase in conspiracy belief [119]. Altogether, these longitudinal results leave us with conflicting findings: while they show us that certain aspects of information spread are remarkably consistent (likely due to similar underlying psychological processes), they also demonstrate how structural features of the modern information environment may have altered information spread.

Box 3: The Future of Virality. *Much prior work on the psychology of virality has relied on dictionary methods (e.g., counting the number of negative or positive words in a social media post). However, large language models have transformed psychological text analysis and can allow for a much more granular understanding of what goes viral. Large language models such as GPT have been found to be much more accurate at detecting human-annotated psychological sentiment than dictionary methods [101], see **Figure I**. They also accurately classify text across languages—including several lesser-spoken and under-resourced languages [101]. Another way large language models can improve the study of virality is by allowing for more sophisticated agent-based models that can model how information spreads through online and offline networks [120].*

These methodological advances allow us to study the type of information that spreads across many more languages and cultures, which could potentially make broader, cross-cultural analyses of virality much easier, given the multilingual effectiveness of modern large language models [101]. However, these cross-cultural studies are still dependent upon data from social media platforms, which is becoming increasingly less available [121]. A few cross-cultural studies have looked at several countries at once, finding some cross-cultural similarities in online interaction patterns [50,122]. For instance, one paper found that right-leaning political parties were amplified by Twitter's algorithm across several political systems [122]. However, even larger studies can look more systematically at how information spread is moderated by important variables (e.g., country-level individualism-collectivism, income-level, inequality, etc.), which would help construct a more global theoretical account of the psychological and structural features that shape information spread. As the online and offline world is becoming increasingly intertwined, future work should also explore how online information spread impacts offline information spread and behavior (and vice versa). As an example of this, some scholars have linked online rhetoric to voting behavior⁶ and protest behavior [123].

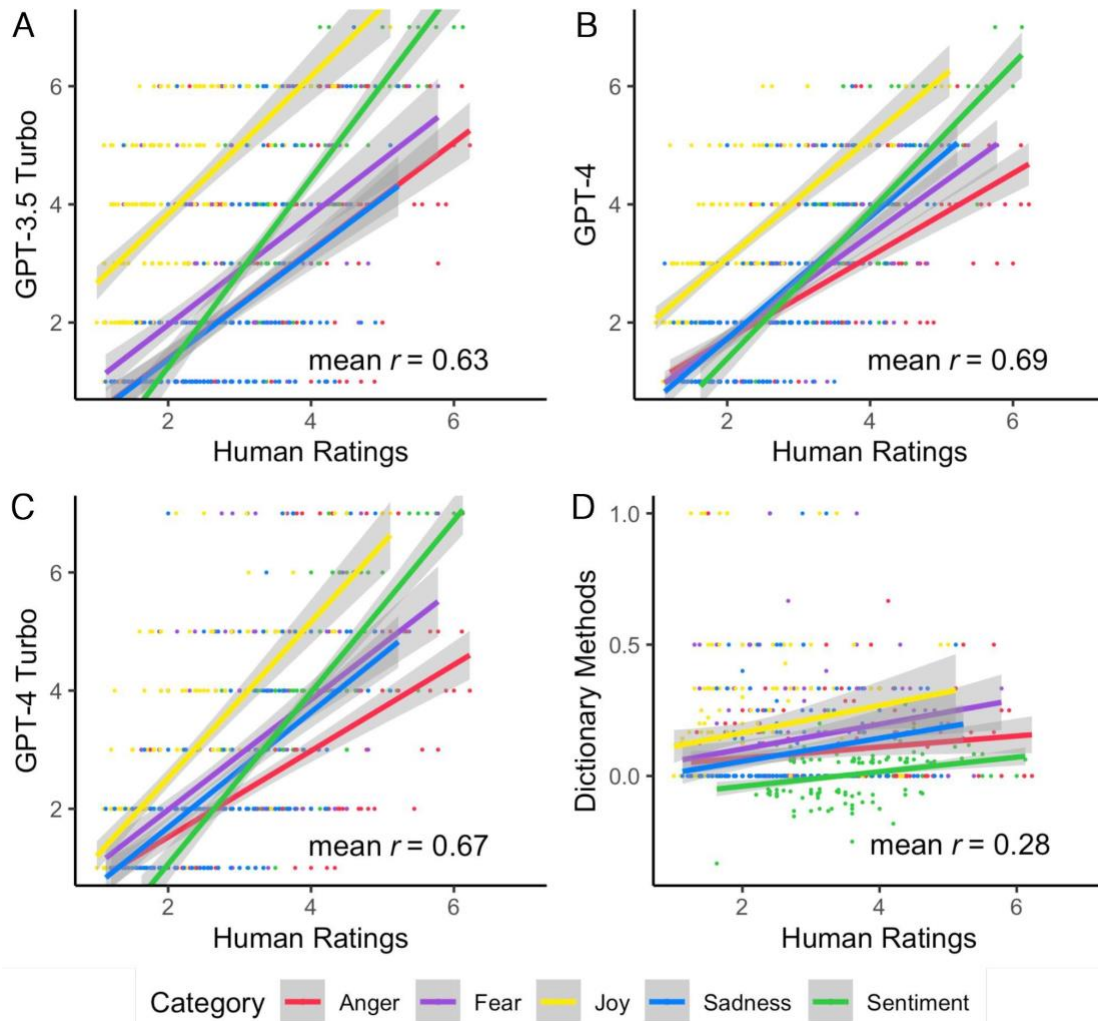


Figure I. Large language models can help advance the psychology of virality. Various versions of the popular large language model GPT (including GPT-3.5, GPT-4, and GPT-4 Turbo) were much more accurate at detecting manually annotated sentiment and discrete emotions (with mean correlations ranging between 0.63-0.67) as compared to dictionary-based text analysis (with a mean correlation of 0.28). GPT also worked well across languages, which can help facilitate more cross-cultural and multilingual research. Further, open-source large language models are rapidly improving, which can help facilitate inexpensive, open, and reproducible research. Figure adapted from Rathje et. al. (2024) [101].

Resources

¹<https://about.fb.com/news/2018/01/news-feed-fyi-bringing-people-closer-together/>

²https://books.google.com/books?id=_OcCAAAMBAJ&printsec=frontcover#v=onepage&q&f=false

³<https://www.pewresearch.org/politics/2017/02/23/partisan-conflict-and-congressional-outreach/>

⁴<https://transparency.meta.com/features/explaining-ranking/>

⁵<https://www.bostonglobe.com/2023/12/07/opinion/rathje-van-bavel-paradox-of-internet-virality/>

⁶<https://www.nytimes.com/2022/10/22/us/politics/republican-election-objectors-rhetoric.html>

Outstanding Questions

- What psychological processes drive information sharing and consumption?
- Do similar factors predict information spread in offline (as opposed to online) contexts?
- Does the information-as-virus metaphor accurately describe the spread of information?
Does using the metaphor impact how people think about the spread of (mis)information?
- Did similar factors predict information spread historically? Would analyzing data from early newspapers and other early text find that similar factors predicted virality long before the advent of social media?
- How do structural features of modern technologies (e.g., algorithms, platform design, network size, etc.) influence the spread of information? How do these structural features interact with psychological processes to shape information spread?
- Are there cross-cultural differences in information spread, and what cultural variables influence the spread of information?
- Can advances in AI (such as text analysis or agent-based modeling via large language models) help us better understand how information spreads?
- How does online behavior impact offline behavior, and vice versa? How do online and offline networks interact?

Glossary

Arousal. The intensity of an emotion, regardless of its negativity or positivity. Anger and excitement are both high-arousal emotions, and sadness and calm are low-arousal emotions.

Circumplex model of affect. A model that proposes that emotions can be visualized according to two core dimensions of affect: valence (the negativity or positivity of an emotion) and arousal (the intensity of an emotion).

Diffusion. Another common term for information spread.

Dictionary methods. A way of classifying or analyzing text data by using predefined dictionaries (e.g., counting the number of negative words in a social media post using a predefined “negativity” dictionary).

Echo Chambers. Environments in which individuals are primarily exposed to information that reinforces their existing beliefs, while rarely encountering opposing viewpoints. Echo chambers can occur both online (e.g., on social media) and offline (e.g., in ideologically segregated neighborhoods).

Memes. Units of cultural transmission—such as ideas, behaviors, or styles—that are spread from person to person.

Paradox of virality. A term used to refer to the idea that widely shared content is often not widely liked.

Psychological factors. Psychological processes that may shape the spread of information (e.g., our preferential attention toward negative and high-arousal information).

Shared Reality. Based on shared reality theory, this term refers to a perceived commonality of inner states with others.

Structural features. Aspects of an environment (such as networks, norms, or algorithms) that may influence information spread.

Superspreaders. Individuals, influencers, or social media accounts that play a disproportionate role in spreading certain types of information.

Valence. The positivity or negativity of an emotion. Sadness and anger are negatively valenced emotions.

Virality. Information spread in online and offline contexts.

References

1. Dunbar, R.I.M. (1996) *Grooming, gossip, and the evolution of language*, Harvard University Press
2. Harari, Y.N. (2024) *Nexus: A brief history of information networks from the Stone Age to AI*, Signal
3. Fisher, M. (2022) *The Chaos Machine: The Inside Story of How Social Media Rewired Our Minds and Our World*, Hachette UK
4. Groh, M. *et al.* (2022) Deepfake detection by human crowds, machines, and machine-informed crowds. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2110013119
5. Van Bavel, J.J. *et al.* (2021) Political psychology in the digital (mis) information age: A model of news belief and sharing. *Social Issues and Policy Review* 15, 84–113
6. Rathje, S. *et al.* (2021) Out-group animosity drives engagement on social media. *Proc Natl Acad Sci USA* 118, e2024292118
7. Brady, W.J. *et al.* (2017) Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* 114, 7313–7318
8. Brady, W.J. *et al.* (2021) How social learning amplifies moral outrage expression in online social networks. *Sci. Adv.* 7, eabe5641
9. Dunbar, R.I. *et al.* (1997) Human conversational behavior. *Human nature* 8, 231–246
10. Grayson, S. *et al.* (2025) Ironic effects of prosocial gossip in driving inaccurate social perceptions. *Journal of Experimental Social Psychology* 116, 104682
11. Baumeister, R.F. *et al.* (2004) Gossip as Cultural Learning. *Review of General Psychology* 8, 111–121
12. Dores Cruz, T.D. *et al.* (2021) An Integrative Definition and Framework to Study Gossip. *Group & Organization Management* 46, 252–285
13. Robertson, C.E. *et al.* (2023) Negativity drives online news consumption. *Nature human behaviour* 7, 812–822
14. Nielsen, R.K. *et al.* (2023) Reuters institute digital news report 2020 Oxford University. <https://www.digitalnewsreport.org/survey/2020/>.
15. Neihouser, M. *et al.* (2022) Political information consumption and electoral turnout during COVID: the case of the 2020 municipal elections. *French Politics* 20, 182
16. Goldberg, M.H. *et al.* (2019) Discussing global warming leads to greater acceptance of climate science. *Proc. Natl. Acad. Sci. U.S.A.* 116, 14804–14805
17. Dawkins, R. (1993) Viruses of the mind. *Dennett and his critics: Demystifying mind* 13, e27
18. Gladwell, M. (2006) *The tipping point: How little things can make a big difference*, Little, Brown
19. Zarocostas, J. (2020) How to fight an infodemic. *The lancet* 395, 676
20. Abdullah, S. and Wu, X. (2011) An epidemic model for news spreading on twitter. in *2011 IEEE 23rd international conference on tools with artificial intelligence*, pp. 163–169
21. Bettencourt, L.M. *et al.* (2006) The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications* 364, 513–536
22. Kauk, J. *et al.* (2021) Understanding and countering the spread of conspiracy theories in social networks: Evidence from epidemiological models of Twitter data. *Plos one* 16, e0256179
23. Lazer, D. *et al.* (2009) Computational social science. *Science* 323, 721–723
24. Rathje, S. *et al.* (2024) People Think That Social Media Platforms Do (but Should Not)

- Amplify Divisive Content. *Perspect Psychol Sci* 19, 781–795
25. Yu, X. *et al.* (2023) Partisanship on Social Media: In-Party Love Among American Politicians, Greater Engagement with Out-Party Hate Among Ordinary Users. *Polit Behav* DOI: 10.1007/s11109-022-09850-x
 26. Rasmussen, J. *et al.* (2025) Negativity and Identity Language Have Additive Effects on Online News Consumption. at <https://osf.io/h3cn5_v1/download>
 27. Watson, J. *et al.* (2024) Negative online news articles are shared more to social media. *Scientific Reports* 14, 21592
 28. Heltzel, G. and Laurin, K. (2024) Why Twitter Sometimes Rewards What Most People Disapprove of: The Case of Cross-Party Political Relations. *Psychol Sci* 35, 976–994
 29. Frimer, J.A. *et al.* (2022) Incivility is rising among American politicians on Twitter. *Social Psychological and Personality Science*
 30. Schöne, J.P. *et al.* (2021) Negativity spreads more than positivity on Twitter after both positive and negative political situations. *Affective Science* 2, 379–390
 31. Schöne, J.P. *et al.* (2023) Negative expressions are shared more on Twitter for public figures than for ordinary users. *PNAS nexus* 2, pgad219
 32. McLoughlin, K.L. *et al.* (2024) Misinformation exploits outrage to spread online. *Science* 386, 991–996
 33. Fan, R. *et al.* (2020) Weak ties strengthen anger contagion in social media. *arXiv preprint arXiv:2005.01924*
 34. Vosoughi, S. *et al.* (2018) The spread of true and false news online. *Science* 359, 1146–1151
 35. Solovev, K. and Pröllochs, N. (2022) Moral emotions shape the virality of COVID-19 misinformation on social media. in *Proceedings of the ACM Web Conference 2022*, pp. 3706–3717
 36. Berger, J. and Milkman, K.L. (2012) What makes online content viral? *Journal of marketing research* 49, 192–205
 37. Knutson, B. *et al.* (2024) News source bias and sentiment on social media. *PloS one* 19, e0305148
 38. Posner, J. *et al.* (2005) The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 715–734
 39. Cowen, A. *et al.* (2019) Mapping the Passions: Toward a High-Dimensional Taxonomy of Emotional Experience and Expression. *Psychol Sci Public Interest* 20, 69–90
 40. Barrett, L.F. (2006) Are Emotions Natural Kinds? *Perspect Psychol Sci* 1, 28–58
 41. Brady, W.J. *et al.* (2020) Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General* 149, 746
 42. Christophe, V. and Rimé, B. (1997) Exposure to the social sharing of emotion: Emotional impact, listener responses and secondary social sharing. *Eur. J. Soc. Psychol.* 27, 37–54
 43. De Angelis, M. *et al.* (2012) On Braggarts and Gossips: A Self-Enhancement Account of Word-of-Mouth Generation and Transmission. *Journal of Marketing Research* 49, 551–563
 44. Kraft, P.W. *et al.* (2020) Social Media and the Changing Information Environment: Sentiment Differences in Read Versus Recirculated News Content. *Public Opinion Quarterly* 84, 195–215
 45. Eckler, P. and Bolls, P. (2011) Spreading the Virus: Emotional Tone of Viral Advertising and Its Effect on Forwarding Intentions and Attitudes. *Journal of Interactive Advertising* 11, 1–11

46. Bonacchi, C. *et al.* (2025) Positive sentiment and expertise predict the diffusion of archaeological content on social media. *Scientific Reports* 15, 2031
47. Martin-Gutierrez, S. *et al.* (2024) In-party love spreads more efficiently than out-party hate in online communities. *Scientific Reports* 14, 15700
48. Kyrychenko, Y. *et al.* (2024) Social identity correlates of social media engagement before and after the 2022 Russian invasion of Ukraine. *Nature Communications* 15, 8127
49. Milkman, K.L. and Berger, J. (2014) The science of sharing and the sharing of science. *Proceedings of the National Academy of Sciences* 111, 13642–13649
50. Falkenberg, M. *et al.* (2024) Patterns of partisan toxicity and engagement reveal the common structure of online political communication across countries. *Nature Communications* 15, 9560
51. Storbeck, J. and Clore, G.L. (2008) Affective Arousal as Information: How Affective Arousal Influences Judgments, Learning, and Memory. *Social & Personality Psych* 2, 1824–1843
52. Sutton, T.M. and Lutz, C. (2019) Attentional capture for emotional words and images: The importance of valence and arousal. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 73, 47
53. Öhman, A. and Mineka, S. (2001) Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological review* 108, 483
54. Baumeister, R.F. *et al.* (2001) Bad is stronger than good. *Review of general psychology* 5, 323–370
55. Milli, S. *et al.* (2023) Twitter’s Algorithm: Amplifying Anger, Animosity, and Affective Polarization. *arXiv preprint arXiv:2305.16941*
56. Robertson, C.E. *et al.* (2024) Morality in the anthropocene: The perversion of compassion and punishment in the online world. *PNAS nexus* 3
57. Globig, L. and Sharot, T. (2024) Considering information-sharing motives to reduce misinformation. *Current Opinion in Psychology*
58. Brady, W.J. *et al.* (2019) The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online. *Perspectives on Psychological Science*
59. Van Bavel, J.J. *et al.* (2024) Updating the identity-based model of belief: From false belief to the spread of misinformation. *Current Opinion in Psychology*
60. Petersen, M.B. *et al.* (2021) Beyond populism: The psychology of status-seeking and extreme political discontent. In *The psychology of populism*, pp. 62–80, Routledge
61. Altay, S. *et al.* (2019) Why do so few people share fake news? It hurts their reputation. *new media & society*
62. Lawson, M.A. *et al.* (2023) Tribalism and tribulations: The social costs of not sharing fake news. *Journal of Experimental Psychology: General* 152, 611
63. Vellani, V. *et al.* (2024) Three diverse motives for information sharing. *Communications Psychology* 2, 107
64. Arceneaux, K. *et al.* (2021) Some people just want to watch the world burn: the prevalence, psychology and politics of the ‘Need for Chaos.’ *Philosophical Transactions of the Royal Society B* 376, 20200147
65. Allport, G.W. and Postman, L. (1946) An analysis of rumor. *Public opinion quarterly* 10, 501–517
66. Scholz, C. *et al.* (2017) A neural model of valuation and information virality. *Proc. Natl.*

- Acad. Sci. U.S.A.* 114, 2881–2886
67. Falk, E.B. *et al.* (2013) Creating Buzz: The Neural Correlates of Effective Message Propagation. *Psychol Sci* 24, 1234–1242
 68. Christopher, C. *et al.* Information diets are more diverse in attention than in engagementOSF
 69. Rathje, S. *et al.* (2023) Accuracy and social motivations shape judgements of (mis) information. *Nature Human Behaviour* 7, 892–903
 70. Globig, L.K. *et al.* (2023) Changing the incentive structure of social media platforms to halt the spread of misinformation. *Elife* 12, e85767
 71. Kozyreva, A. *et al.* (2024) Toolbox of individual-level interventions against online misinformation. *Nature Human Behaviour*
 72. Martel, C. *et al.* (2024) On the Efficacy of Accuracy Prompts Across Partisan Lines: An Adversarial Collaboration. *Psychol Sci* 35, 435–450
 73. Tamir, D.I. and Mitchell, J.P. (2012) Disclosing information about the self is intrinsically rewarding. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8038–8043
 74. Higgins, E.T. *et al.* (2021) Shared Reality: From Sharing-Is-Believing to Merging Minds. *Curr Dir Psychol Sci* 30, 103–110
 75. Harber, K.D. and Cohen, D.J. (2005) The Emotional Broadcaster Theory of Social Sharing. *Journal of Language and Social Psychology* 24, 382–400
 76. Frimer, J.A. and Skitka, L.J. (2018) The Montagu Principle: Incivility decreases politicians' public approval, even with their political base. *Journal of Personality and Social Psychology* 115, 845
 77. Beshears, J. *et al.* (2008) How are preferences revealed? *Journal of public economics* 92, 1787–1794
 78. Richter, M.K. (1966) Revealed preference theory. *Econometrica: Journal of the Econometric Society*
 79. Allcott, H. *et al.* (2022) Digital addiction. *American Economic Review* 112, 2424–63
 80. Rathje, S. *et al.* Unfollowing hyperpartisan social media influencers durably reduces out-party animosity. *Preprint at <https://osf.io/acbwg/download> at <<https://osf.io/acbwg/download>>*
 81. Baribi-Bartov, S. *et al.* (2024) Supersharers of fake news on Twitter. *Science* 384, 979–982
 82. Grinberg, N. *et al.* (2019) Fake news on Twitter during the 2016 US presidential election. *Science* 363, 374–378
 83. Kumar, D. *et al.* (2023) Understanding the Behaviors of Toxic Accounts on Reddit. in *Proceedings of the ACM Web Conference 2023*, Austin TX USA, pp. 2797–2807
 84. Kasper, J. and Gilovich, T. (2025) Do people prefer to share political information that boosts their ingroup or derogates the outgroup? *Journal of Experimental Psychology: General* at <https://psycnet.apa.org/record/2025-72351-001?casa_token=yZS1z9nYiu8AAAAA:LqjtLej08hBy11L1bgqmsSl-xcH0Xo8UhD7yOrqdSnt-mIbsciAN8diJi52eYUk8-xLYUD_bLWB18v7glBqqP20>
 85. Robertson, C. *et al.* (2024) Inside the Funhouse Mirror Factory: How Social Media Distorts Perceptions of Norms. *Current Opinion in Psychology* at <<https://osf.io/kgcrq/download>>
 86. Dodds, P.S. *et al.* (2011) Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PloS one* 6, e26752
 87. Bor, A. and Petersen, M.B. (2022) The psychology of online political hostility: A comprehensive, cross-national test of the mismatch hypothesis. *American political science review* 116, 1–18

88. Schroeder, J. *et al.* (2017) The Humanizing Voice: Speech Reveals, and Text Conceals, a More Thoughtful Mind in the Midst of Disagreement. *Psychol Sci* 28, 1745–1762
89. Guess, A.M. *et al.* (2023) How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science* 381, 398–404
90. Haidt, J. (2022) Why the Past 10 Years of American Life Have Been Uniquely Stupid. *The Atlantic* 11
91. Guess, A.M. *et al.* (2023) Reshares on social media amplify political news but do not detectably affect beliefs or opinions. *Science* 381, 404–408
92. De Freitas Melo, P. *et al.* (2020) Can WhatsApp Counter Misinformation by Limiting Message Forwarding? In *Complex Networks and Their Applications VIII* 881 (Cherifi, H. *et al.*, eds), pp. 372–384, Springer International Publishing
93. Pretus, C. *et al.* (2024) The *Misleading* count: an identity-based intervention to counter partisan misinformation sharing. *Phil. Trans. R. Soc. B* 379, 20230040
94. Goldstein, N.J. *et al.* (2007) Invoking social norms: A social psychology perspective on improving hotels' linen-reuse programs. *Cornell Hotel and Restaurant Administration Quarterly* 48, 145–150
95. Mosleh, M. *et al.* Divergent patterns of engagement with partisan and low-quality news across seven social media platforms. at <<https://osf.io/9csy3/download>>
96. Ellwardt, L. *et al.* (2012) The co-evolution of gossip and friendship in workplace social networks. *Social Networks* 34, 623–633
97. DiResta, R. (2024) *Invisible rulers: the people who turn lies into reality*, Hachette UK
98. Bakshy, E. *et al.* (2012) The role of social networks in information diffusion. in *Proceedings of the 21st international conference on World Wide Web*, Lyon France, pp. 519–528
99. Rathje, S. *et al.* (2022) Social media behavior is associated with vaccine hesitancy. *PNAS Nexus*
100. Gentzkow, M. and Shapiro, J.M. (2011) Ideological segregation online and offline. *The Quarterly Journal of Economics* 126, 1799–1839
101. Rathje, S. *et al.* (2024) GPT is an effective tool for multilingual psychological text analysis. *Proc. Natl. Acad. Sci. U.S.A.* 121, e2308950121
102. Feuerriegel, S. *et al.* (2025) Using natural language processing to analyse text data in behavioural science. *Nature Reviews Psychology*
103. Hsu, T.W. *et al.* (2021) Social media users produce more affect that supports cultural values, but are more influenced by affect that violates cultural values. *Journal of Personality and Social Psychology*
104. Hopkins, D.J. *et al.* (2024) The rise of and demand for identity-oriented media coverage. *American J Political Sci* DOI: 10.1111/ajps.12875
105. Finkel, E.J. *et al.* (2020) Political sectarianism in America. *Science* 370, 533–536
106. Dawkins, R. (1981) In defence of selfish genes. *Philosophy* 56, 556–573
107. Loomba, S. *et al.* (2021) Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*
108. Roozenbeek, J. *et al.* (2022) Psychological inoculation improves resilience against misinformation on social media. *Science Advances*
109. Simon, F.M. and Camargo, C.Q. (2023) Autopsy of a metaphor: The origins, use and blind spots of the 'infodemic.' *New Media & Society* 25, 2219–2240
110. Drury, J. (2018) Beyond the contagion concept. *Psychology Review*
111. Lerman, K. (2016) Information is not a virus, and other consequences of human cognitive

- limits. *Future Internet* 8, 21
112. Jenkins, H. *et al.* (2008) If it doesn't spread, it's dead: Creating value in a spreadable marketplace. *Convergence Culture Consortium*
 113. Norman, A. *et al.* (2024) Do minds have immune systems? *Journal of Theoretical and Philosophical Psychology* at <<https://psycnet.apa.org/record/2025-55535-001>>
 114. Thibodeau, P.H. and Boroditsky, L. (2011) Metaphors we think with: The role of metaphor in reasoning. *PloS one* 6, e16782
 115. Kaplan, R.L. (2008) Yellow journalism. *The international encyclopedia of communication* 11, 5360–5371
 116. Muthukrishna, M. *et al.* (2021) Psychology as a Historical Science. *Annu. Rev. Psychol.* 72, 717–749
 117. Rozado, D. *et al.* (2022) Longitudinal analysis of sentiment and emotion in news media headlines using automated labelling with Transformer language models. *Plos one* 17, e0276367
 118. Avalle, M. *et al.* (2024) Persistent interaction patterns across social media platforms and over time. *Nature* 628, 582–589
 119. Uscinski, J. *et al.* (2022) Have beliefs in conspiracy theories increased over time? *PLoS One* 17, e0270429
 120. He, J.K. *et al.* (2024) Artificial intelligence chatbots mimic human collective behaviour. *British J of Psychology* DOI: 10.1111/bjop.12764
 121. Rathje, S. (2024) To tackle social-media harms, mandate data access for researchers. *Nature* 633, 36–36
 122. Huszár, F. *et al.* (2022) Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences* 119
 123. Mooijman, M. *et al.* (2018) Moralization in social networks and the emergence of violence during protests. *Nature human behaviour* 2, 389–396