



The effect of social approval on perceptions following social media message sharing applied to fake news

Joseph B. Walther ^{1,2,*}, Zijian Lew ^{1,3}, America L. Edwards ¹, Justice Quick ¹

¹Department of Communication, University of California, Santa Barbara, USA

²Center for Information Technology and Society, University of California, Santa Barbara, USA

³Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore

*Corresponding author: Joseph B. Walther, Department of Communication, University of California, Santa Barbara 93106-4020, USA. email: jwalther@ucsb.edu

Abstract

A field experiment examined social approval in the form of Twitter “Likes” on individuals’ perceptions after retweeting a fictitious news story about a politician. The study incorporated research about feedback effects on self-perception online, partisan bias, and negativity principles. Participants read or retweeted a (verifiably false) news story via social media, and researchers appended systematic increments of Likes to the retweets. A baseline hypothesis test found no effect on perceptions due to retweeting versus simply reading a news story. Results supported a predicted three-way interaction effect between positive versus negative news story, political congruence with participants’ political party identification, and the reception of 0–22 Likes on perceptions of the politician. More Likes magnified negative perceptions of politicians, from fictitious news, when news stories were negative and focused on politicians from one’s opposite political party.

Keywords: social approval, social media, Likes, identity shift, fake news

Research has approached the question of how social media affect attitudes and perceptions through several perspectives and in a variety of contexts (see [Flanagin, 2017](#), for review). A great deal of this research has focused on the cues that affect a reader’s perceptions of information, recommendations, and news. The cues may involve qualities of the messages, characteristics of the message sources, and endorsements that the messages receive either tacitly (through popularity cues such as approval ratings) or explicitly (through user-generated comments) (see [Walther & Jang, 2012](#), for review). Less research has focused on how social media message senders might influence themselves as their messages do or do not resonate in the social media ecosystem. However, there is good reason to suspect that “self-effects” and social effects interact to influence social media authors’ own attitudes with considerable potency (see [Valkenburg, 2017](#), for review).

Whether dealing with actual news, innocuous urban legends, or even falsehoods, it is not the mere consumption of information that concerns us, but that people who believe it also propagate it ever so easily to others through social media’s affordances ([Metzger et al., 2021](#)). This study explores a hypothetical reversal of that causal order: First, people propagate information to others, and then, in turn, come to believe it. Their beliefs may change all the more strongly if they garner social approval for their shared message. People share information online to pursue various goals, among which are to garner recognition and social approval ([Crockett, 2017](#)) even when the veracity of the information is dubious ([Duffy et al., 2020](#)). And the provision of social approval symbols in social media is simple to encode, easy to transmit, and well-understood, through prolific “paralinguistic digital affordances,” such as thumbs up graphics, hearts, and other such one-click symbols ([Hayes et al., 2016](#)). By these means, the positive social responses that social media message senders

subsequently receive may magnify changes in their own perceptions to align with what they post.

Whether this process takes place when information is verifiably correct is an interesting question. Yet another context in which the study of social media’s influence on perceptions has become particularly vital pertains to the flow of fake news. Whether message sharing and social approval wed individuals’ beliefs to things that are untrue provides a relatively more challenging and potentially important research context. For that reason, this research explored whether the receipt of social approval indicators from other social media users after sharing news—in this case, verifiably false news—affects individuals’ own (false) perceptions, under circumstances that commonly resemble the transmission of fake news. It found that when individuals share news via social media, and others provide social approval merely by “Liking” their posting, it intensifies the individuals’ negative perceptions of the target of a news story, even though the events in the news story never happened. A field experiment examined what happened when individuals chose one of the two web-based online news stories they originally saw linked to Twitter tweets, and “quote tweeted” (i.e., shared it in a retweet with their own original comment) encouraging others to read it, too (or they simply read it, without sharing, in a control condition). Unbeknownst to participants, the politicians and events described in the stories did not actually exist. Researchers then covertly appended social approval indicators—heart-shaped “Likes” on Twitter—on participants’ retweets in systematically different increments of five, from 0 to 20. Posttest scores indicated a significant influence of these accumulated social approval markers on participants’ evaluations of the news stories’ targets, in an interaction effect with influences of negative versus positive portrayals of the targets and whether the target represented the same or opposing political party. While

the findings align with previous research about the spread of fake news, they also extend our understanding by considering previously unexamined influences of sharing news stories followed by the accumulation of social approval cues.

The research drew on several theories and lines of research in computer-mediated communication (CMC). It includes aspects of the hyperpersonal model of CMC and its extension, the identity shift paradigm, that focus on changes to one's perceptions due to one's own utterances when interacting with others. It recognizes negativity effects and the partisan bias principle from political communication research. It investigates the value of a paralinguistic cue in social media—the accumulation of “Likes”—as a form of feedback that affects changes in perception. Results speak to the importance of feedback and social approval online, and illuminate a novel approach to the effects of news sharing: how it affects the attitudes of the people who propagate it, not because of their consumption of content alone, but rather, in part, due to the social responses they accrue through their propagation of it.

Social media, feedback effects, and perception change

Research has approached the question of how social media affect attitudes and perceptions through several perspectives and contexts (see Flanagan, 2017, for review). One perspective that focuses on changes in perceptions in online communication is the hyperpersonal model of CMC (Walther, 1996). The model contends that social media afford malleability of online message creation and encourage selective self-presentation. Feedback from other online communicators reinforces these self-presentations, causing changes in individuals' perceptions of self, partner, and relationships (see Walther & Whitty, 2021, for review). Related research also shows that attitudes toward the objects that conversation partners discuss online can be affected by self-presentation and feedback processes. For instance, chatting online about a product, with a goal of attracting (or repelling) one's chat partner, leads to changes in individuals' own perceptions about the product they are discussing (Walther et al., 2010). The feedback component of the hyperpersonal model distinguishes it from intrapersonal models of perception change. For instance, a “self-effect” may occur when an individual prepares and expresses opinions about media content to another person, aligning perceptions with what the individual uttered, regardless of whether a recipient responds (Pingree, 2007; cf. Valkenburg, 2017). In contrast, Walther and Whitty (2021) suggested, when individuals garner feedback via Likes for having shared information online, they may perceive its content to be more authentic, leading them to internalize the beliefs and attitudes they implied online, rather than harbor doubts that they pandered something fake.

Although it initially did not include feedback processes, an extension of the hyperpersonal model known as identity shift further explores how individuals' online self-presentation leads to changes in perceptions of self. Drawing on psychological research on the making of public commitments to a position (see Schlenker et al., 1994, for review), Gonzales and Hancock's (2008) experiment requested participants, through random assignment, to write anecdotes about themselves, either in a public blog or a private document, that would lead readers to infer that a participant was introverted, or extraverted, respectively. Post-test

self-assessments of personality showed that participants who wrote in the public blogs (but not private documents) had come to perceive themselves as significantly more introverted or extroverted, respectively, as a result (see also Cho et al., 2018, for similar internalization of one's political expressions). Altogether, this research shows that merely sharing positions online has the capacity to change individuals' self-perception in the direction that the shared information implies. In light of this theoretical and empirical background, H1 predicts a similar effect on one's perceptions following one's public sharing of online news:

H1: Sharing a news story online leads to perceptual changes in the direction of that story more strongly than merely consuming the story.

Additional research yet again extended the identity shift phenomenon, returning it more squarely in line with the hyperpersonal model's specification of sender/receiver/channel/feedback effects. Beyond the effects of online public self-presentation changing self-perceptions, additional research demonstrated that feedback confirming one's online self-presentation further magnifies perception change. For instance, Walther et al. (2011) asked participants to describe themselves online as if introverted or as if extraverted, after which half the participants in these two conditions received feedback reflecting their respective trait, and the other half did not. Feedback significantly magnified the participants' perceptions of themselves as introverted or extraverted, compared to no feedback. Further replication explicitly tied these effects to social media: Carr and Foreman (2016) had individuals write self-descriptions using Facebook status updates, and feedback, in the form of others' comments, caused individuals to experience greater self-perception changes. Additional research extended the self-presentation/feedback effect beyond perceptions of themselves, to include other attitude objects. For instance, Carr and Hayes (2017) asked participants to present themselves online as either strongly loyal or weakly loyal to a product brand, to which they received either confirming or disconfirming feedback. No effect emerged due to the strong versus weak loyalty affirmations, but participants' brand identification was significantly greater after they received confirming rather than disconfirming feedback.

In sum, in a variety of online settings, individuals who publicly describe themselves in a certain way or participate in advocacy internalize their performance and shift their perceptions and attitudes in the direction of that performance. When they receive feedback providing social approval, perceptions change even more. These notions are consistent with offline research, in which social corroboration of individuals' attitudes leads them to adopt even more extreme opinions (Baron et al., 1996).

In this respect, H1, above, takes on added importance. In order to study whether getting feedback after publicly sharing information affects one's perceptions, research must establish whether simply sharing (without feedback) is sufficient to change participants' perceptions about the topic to which their shared information pertains. If so, sharing could be confounded with feedback, because one must share in order to get feedback. Therefore, H1 is not only an application of identity shift principles. It is an important baseline question that can help establish whether sharing and feedback are potentially independent, additive, or interactive effects.

In online studies, feedback generally takes the form of verbal statements (see Carr et al., 2021, for review). Yet social media have other ways to deliver affirming feedback and signals of social approval, using graphical symbols that are generically referred to as Likes. These “markers of approval” (Tokunaga & Quick, 2018, p. 158) provide, according to Facebook’s first president, “a social-validation feedback loop” (Solon, 2017, n.p.).

Likes as social approval feedback

Likes refer to graphical symbols of affirmation appended by social media users to other users’ message postings. Across platforms, social media users can click a button to make these kinds of symbols appear publicly. The “thumbs up” symbol on Facebook is one form of a Like; the equivalent symbol on Twitter is heart-shaped. Reddit uses symbols denoting upvotes. The symbols are phatic: they have no semantic value, but they convey social approval, affirmation, and sociability (Carr et al., 2016; Hayes et al., 2016).¹

Considerable research has identified the impacts of observing Likes, the number of Likes, and other popularity cues as they appear on social media postings (see Haim et al., 2018, for review). When they appear on others’ postings, as Flanagan (2017) summarized, “online news articles with greater Facebook ‘likes’ have been shown to be clicked on more frequently, selected earlier, and read for longer than articles with a low number of likes” (pp. 454–455; Winter et al., 2016).

The present study focuses, however, on seeing Likes appended by others to one’s own postings, consistent with the work by Rosenthal-von der Pütten et al. (2019) conceiving of “‘Likes’ as social rewards,” that constitute a “numerical representation of social acceptance,” and act as reinforcers (p. 76). Several studies provide indirect support for the contention that receiving Likes connotes social approval and affirming feedback. Hayes et al. (2016) reported three main gratifications associated with receiving Likes: emotional gratifications (feelings of happiness and self-worth), status gratifications (improved social standing), and social gratifications (enhancing relationships). Because Likes can accumulate, the number of Likes one receives can also connote different amounts of approval. Scissors et al. (2016) investigated the number of Likes that Facebook users considered sufficient to indicate success of their posts, and how important that recognition is to them. Between 8 and 10 Likes were considered a sign of success, although the range was considerable and the distribution was skewed: 25% reported that one or two Likes was sufficient to know their message was read.

Some research indicated that Likes have greater impact when people receive them from friends than from strangers (Carr et al., 2018; Scissors et al., 2016). Others found that the source is tangential compared to the sheer volume of Likes, in terms of their effects on one’s feelings of belongingness and self-esteem (Reich et al., 2018). An experiment using a simulated social media platform found that receiving even a moderate number of Likes from strangers induced reactions similar to those of overcoming ostracism: greater self-esteem, better mood, more friendliness, more meaningful existence, and less sadness and anger (Wolf et al., 2015).

If Likes connote these affirming properties, then the more Likes users see appended to their own social media users’ postings, the more it may lead them to internalize the perceptions and attitudes conveyed in those postings, even if they

had little original commitment to that message. In the context of fake news, if retweeting a news story with a comment connotes tacit agreement, through selective self-presentation, and if receiving Likes through social media provides feedback to reinforce that expression of attitude, can this process persuade an individual to magnify beliefs even if they have no basis in fact?

Fake news as context

The social influence dynamics that this research explores may occur when people share information that is verifiably true, but it may be equally likely to explain what happens when people share fake news, as well. In their seminal investigation of the 2016 US Presidential election campaigns, Allcott and Gentzkow (2017, p. 213) defined fake news as “news articles that are intentionally and verifiably false, and could mislead readers. We focus on fake news articles that have political implications . . . Our definition includes intentionally fabricated news articles” (for other definitions, see Tandoc et al., 2018). Fake news has the potential to disrupt democratic processes, public health, and other critical societal functions (Allcott et al., 2020). The practice is global, and a 2020 survey of media users in Africa, Eastern and Western Europe, South America, and Asia indicated that 40%–70% of adults use social media as a daily source of news, even though they do not trust it (Watson, 2020). One survey suggested that 23% of the US adults have shared fake news, knowingly or unknowingly, with friends and others (Barthel et al., 2016).

Fake news has a typical topography of presentational and transmission characteristics. It typically resides on websites made to resemble authentic news sites in terms of their domain names and layout (see PolitiFact, 2017). Its circulation relies primarily on the transmission of links embedded in social media messages (Guess et al., 2018). Political topics are among the fake news most frequently circulated through Twitter, according to an analysis of 126,000 stories tweeted by 3 million people between 2006 and 2017 (Vosoughi et al., 2018). Replicating these characteristics allows us to test various hypotheses about message sharing and social approval in the context of fake news.

One might hypothesize that, the more Likes one receives when sharing a (fake) news story via social media, the more one’s perceptions align with the depiction in the story. The direction of the effect of Likes, of course, depends on the positive versus negative valence of a (fake) news story. That is, if a story is positive, the more Likes one received after sharing the story, the more positively one feels toward the subject of the story. Alternatively, if a story is negative, the more Likes one receives after sharing the story, the more negatively one feels toward the subject of the story.

Valence effects

Two additional aspects of the processing of positive versus negative information, however, suggest that the aforementioned prediction may be too simple. First, research finds that people generally pay more attention to information that focuses on negative rather than positive things (Baumeister et al., 2001). The mere consumption of negative (fake) news stories may induce stronger effects on readers’ perceptions than do positive news stories, regardless of whether or not they share a story and/or whether or not they receive signals of social approval. Second, negative stories may also be more

likely to trigger the experience and the expression of moral outrage, and thereby prime outraged individuals for confirming social responses they then receive for publicly sharing that very outrage. According to [Crockett \(2017\)](#),

... outrage expression provides reputational rewards. People are not necessarily conscious of these rewards when they express outrage. But ... a concern for reputation at least implicitly whets our appetite for moral outrage. Of course, online social networks massively amplify the reputational benefits of outrage expression. (p. 770)

Thus, individuals may receive reinforcement in terms of enhanced reputation and status when they share a news story via social media that depicts moral violations. If the research indicating that the valence of news matters, in the context of (fake) news sharing, then perceptions should be expected to shift more strongly following the accumulation of Likes when social media users share negatively focused stories than when they share positively focused ones. In addition, negative news might be more likely to arouse anger, and stories that arouse readers' anger make them especially susceptible to party-consistent misinformation ([Weeks, 2015](#)).

Partisan congruence

The effects of positive versus negative depictions of a character in a news story might also be influenced, or overridden, by partisan congruence. In other words, partisan congruence may moderate the influence of Likes on perception changes following an online self-presentation. People are more influenced by false information that is congruent or congenial to their existing political ideologies ([Guess et al., 2020](#); [Neyazi & Muhtadi, 2021](#)), just as they are with political news generally. Fake news tends to be shared and interpreted along political partisan lines (even when the fake news is not overtly politically-focused, such as that pertaining to coronavirus disease (COVID-19) remedies; [Allcott et al., 2020](#)). Partisan congruence exists when self-professed Republican participants read stories favoring Republican politicians or denigrating Democratic politicians, or when Democratic participants read stories favoring Democratic politicians or denigrating Republican ones. Partisan incongruence exists when Republicans read negative Republican or positive Democrat stories, and when Democrats read negative Democrat or positive Republican stories.²

Altogether, it seems plausible that negative/other-party news may be the most influential form. In light of potential for negative political ads and stories to have stronger effects than positive ones do, and that the sharing of moral violations (that are negatively valenced) leave social media users especially susceptible to social approval, social media users may be especially sensitive to the number of Likes they receive when sharing fake news that is partisan congruent and depicts (negative) moral violations. As such, hypothetical expectations about the potential for selective self-presentation of fake news and garnering social approval to change one's perception to align with shared news are likely moderated by news story valence and partisan congruence, as follows:

H2: The more Likes one receives for sharing negative (rather than positive) news about out-party (rather than in-party) actors, the more negative one's perceptions of the target.

There is a three-way interaction among the number of Likes one receives, story valence, and partisan congruence.

The hypothesis suggests that the impact of Likes takes the form of cleaved, divergent moderation effect, in the framework advanced by [Holbert and Park \(2020\)](#). That is, different levels of Likes change the relationships of the valence \times congruence effect on perceptions, magnifying the relationship in opposing directions. A graphic depiction of the hypothesis appears in [Figure 1](#).

Method

The main experiment involved 622 participants, 129 of whom served in control groups. The other 493 read, selected one of two news stories that had been fabricated for this research, and shared the news story on their personal Twitter accounts via a retweet. Over the next several days, their retweets received a specific number of Likes from researchers using Twitter accounts with minimal identifiers. At the end of five to seven days, both control and experimental participants completed scales measuring their perception of the politicians who were the respective foci of the news stories. Analyses reveal partial support for the study's hypotheses. Several pilot tests informed the creation of stimuli prior to the main experiment.

Number of Likes

A small pilot study examined what the typical number of Likes individuals' posts receive on Twitter. The study recruited 25 Amazon Mechanical Turk (MTurk) workers who resided in USA, were 18+ years of age, had an active Twitter account (according to MTurk selection criteria), and who self-reported generating at least five tweets or retweets in the last year. Respondents completed a 5-min survey for US\$1. Their average age was 39.5 ($SD = 12.79$); 50% identified as female and 50% as male. The survey asked them to look up their three most recent tweets or retweets, and to indicate the date of each. After describing how to find the information, it asked participants to indicate the number of impressions, engagements, Likes, retweets, and follows each post generated. Two users reported over a hundred Likes, but they appeared to be business accounts and were eliminated from further analysis.

The modal number of Likes participants reported was zero ($M = 1.83$, $SD = 5.70$). Although there was a range of 0–37 Likes, 83% of scores were 0 or 1, with two scores of 30 or more. Despite the higher numbers in previous Facebook research, the low numbers in this pilot study led to the adoption of experimental manipulations involving Likes in equal intervals from 0, 5, 10, 15, and 20. For most people, even 5 Likes would be notable, and 20 would be high.

Fake news stimuli

The researchers fabricated two positive and two negative news stories, following the journalistic style of news stories that have been linked to various fact-checking sites. Each story focused on a California State Senator or State Assembly member. The names of the people and places in the stories were fictitious. The photographs appearing in the stories obscured individuals' identification. The politicians' names featured a parenthetical indication of their political party and a location they appeared to represent (e.g., "State Senator Fred

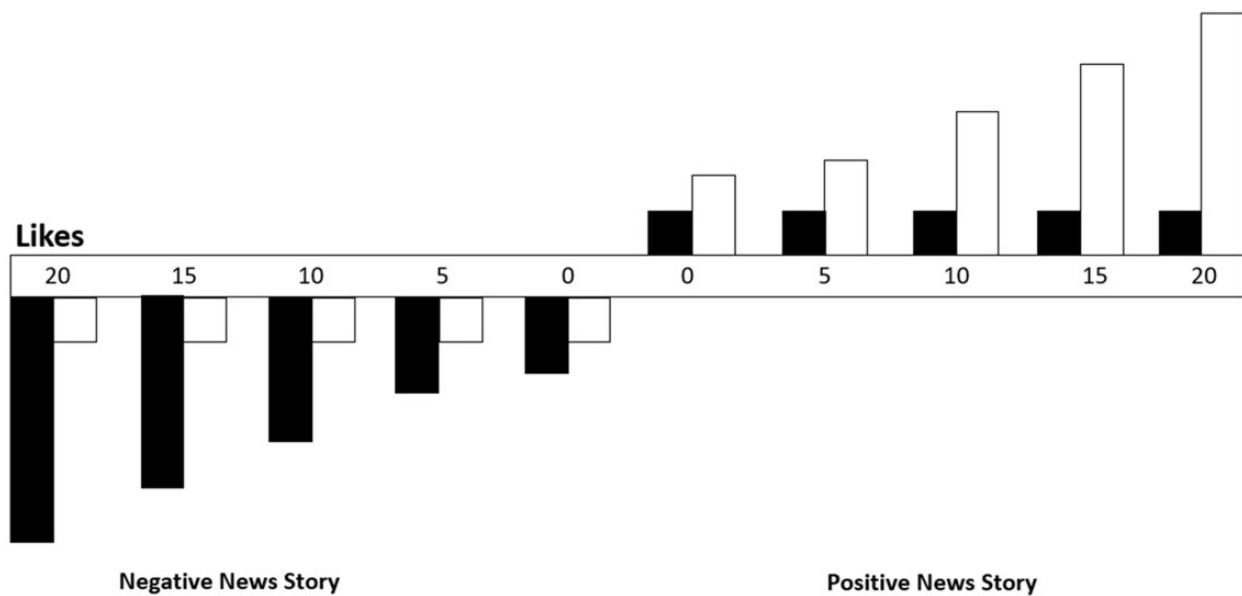


Figure 1. Hypothesized effect of likes on perception of out-party and in-party politician due to positive versus negative news story.

Stevens (D., Bakersfield). . .”). The letter indicating party affiliation was rotated between D and R for Democrat and Republican.

A pilot study tested initial drafts of the two positive versus two negative stories to assess whether the intended differences in valence affected perceptions of the politicians in the story appropriately. The pilot test presented different tweets with embedded links connecting to news stories, appended with 0, 5, 10, 15, or 20 Likes in order to assess both baseline perceptions attributable to valence, as well as potential popularity cue effects due to the number of Likes, and valence-by-Likes. Recruiting 424 participants from MTurk, selection criteria included being 18+ years old, residing in USA, having an active Twitter account (according to MTurk), and self-reporting that they tweeted or retweeted at least five times in the last year. Pilot test respondents had an average age of 39.2 ($SD = 1.46$); 55% identified themselves as female, 43% as male, with 1% declining or indicating another answer. After reading a randomly assigned news story, they completed Stroud et al.’s (2005–2006) “candidate evaluation questionnaire” to measure their perceptions of the news stories’ targets. In previous research, scores on the scale’s nine interval Likert-type items “emotional and competence-related words” (i.e., credible, reliable, appealing, dignified, competent, honest, flaky, interesting, irrational) indicated sensitivity to partisan differences in perceptions of candidates. Two additional items related to credibility were added, “expert” and “trustworthy”; Cronbach’s $\alpha = .98$; $M = 4.80$; $SD = 32.7$. The pilot study took participants 5 min to complete and provided US\$1 in compensation.

A factorial analysis of variance (ANOVA) with positive versus negative valence as a fixed factor and Likes as a random factor showed a significant main effect of valence on perception of politicians, $F(4,414) = 1092.40$, $p < .001$, with no significant effect of Likes, $F(4,414) = 0.467$, $p = .76$, and no interaction effect. Post hoc tests revealed no differences among the two positive stories or among the two negative stories, respectively.

The onset of the COVID-19 pandemic during the development of this study required changes to the stories to make

them plausible considering social distancing and face-covering requirements. Among the final versions of the fake stories, one negative story described a politician who appeared inebriated at a hospital visit and made fun of the pediatric patients’ handmade face coverings. The other negative story portrayed an official shaking hands with constituents outside of a grocery store in violation of social distance requirements, entering the store without a mask, and declaring that the pandemic was “d-o-n-e.” Positive stories either depicted a State Senator’s gift to a pediatric hospital to fund clean rooms where families could visit during the pandemic, or an Assembly member’s US\$1 million donation to provide medical-grade personal protective equipment for hospital workers who would otherwise face shortages. Stories were between 251 and 258 words long and featured the same news organization logo, date, and similar bylines. Each had one main photo, with three smaller photos with links to “related stories” at the bottom of the page (see stimuli in [Supplemental Materials](#)).

The stories themselves appeared on web server space that was leased specifically for this research. Consistent with many actual fake news sites, it featured a news-like, commercial domain name (see [PolitiFact, 2017](#)), “cmcnewsnow.com,” and appeared to be copyrighted to Consolidated Media Corporation. Although fake news can be promoted on any social media platform, this experiment used Twitter tweets to link to the stories. According to [Allcott et al. \(2019\)](#), while exposure to “false content” on Facebook has fallen since 2016, it continued rising on Twitter. The Twitter account created for this study featured the handle, @CMCNewsNow.

For the main study, researchers recruited participants by placing advertisements on Facebook, targeted toward seeking users over the age of 18 years, residing in USA, with Twitter experience. The advertisement encouraged volunteers to sign up for a two-part university-sponsored research project in exchange for US\$10. Individuals who clicked on the advertisement were directed to a more detailed sign-up web page, which stated additional qualifications: Participants were required to have used Twitter at least five times in the last year,

agree to utilize a laptop/desktop computer during the study (due to the necessity of multiple windows being open), be willing to provide their Facebook Messenger address for communication and payment through Facebook Pay, and to have a public (not private) Twitter account (to enable the experimental manipulation). This last requirement would enable researchers to view personal identifiers, the instructions explained, although they were of no interest to the research and would not be accessed. Nevertheless, to maximize participants' confidence in their privacy in light of such incidents as the Cambridge Analytica research exploitation scandal of 2016, researchers pledged that participants' data would not be used for any purpose other than this particular study or provided to any other researchers. Participants learned that the two parts of the study would occur five to seven days apart; the former would take 7 min and the latter, 15. Researchers used participants' Messenger addresses, about which participants were also informed, to track participation (including assignments to conditions and the merger of part 1 and part 2 data). Instructions also indicated that participants would be asked to retweet a message which would be visible to their followers and others, and "there is a risk that people might respond to your retweet with comments (positive or negative) that are beyond anyone's control." When prospective participants consented to these terms, researchers randomly assigned them to different conditions and sent a link to participants by Messenger to begin the first part of the study.

Assignment to conditions

In all conditions, the first part of the study stated, "Research has shown that many people pass along news headlines via social media without actually reading the stories they are linked to even though they sometimes comment on them," and continued, "Today we'll ask you to look at two news Tweets and the news stories they're linked to on the web." Subsequent instructions depended on participants' random assignment to control or experimental conditions following what is best described as a $2 \times 2 \times 5$ design with offset 2×2 Control conditions. The first two-level factor in both Experimental and Control conditions was whether participants would read two positively or two negatively valenced stories. The second two-level factor was whether the stories they read depicted Republicans or Democrats. In the five-level experimental condition, participants would retweet and comment on one of the two stories they saw, and surreptitiously receive 0, 5, 10, 15, or 20 Likes to their retweet and comment. The initial conditions and the variables they incorporate are depicted in Table 1.

Control group participants read two tweets and the two web-based stories to which the tweets linked, reflecting one of the positive/negative, Democrat/Republican combinations described above. Instructions asked them to select one story as the "most important for people to read," and to record the headline of the story in our questionnaire. They then read that they would be contacted for the second part of the study within a week.

Experimental condition participants also saw two tweets linked to stories, in the same combinations as Control participants. Their instructions, however, directed them to retweet the story they thought was more important, and to "add a comment... that would encourage other people" to read it, using three to five words (known as a "quote tweet"). Instructions then described how to find the "link to tweet" for their quote tweet and record it in our questionnaire. A "link to tweet" is a web-based hyperlink that points directly

to a specific tweet. Participants read that they would be contacted for part 2 within a week.

Many prospective participants ($n = 346$) did not complete part 1 of the study. Most dropped out without a response. Some contacted the researchers to withdraw explicitly. Two expressed confidence that a news story was fake. Other participants were removed by the researchers if their Twitter account was not public (preventing Liking of their retweet), who provided the "link to tweet" from the original tweet and not from their retweet, or whose online comments stated that it was part of an experiment. Although the level of attrition may confine generalizability of results to individuals who do not suspect or who do not resist fake news, this constraint retains a large and important population of interest. The final sample consisted of 622 participants, with ages between 18 and 70 years ($M = 29$; $SD = 10.18$). About 61% of the participants identified as female, 38% as male, and 1.4% declined to answer. The sample was just over half Caucasian (53%), 24% were Asian or Pacific Islander, 7% Latinx, 6% Black, 0.2% Native American, and 1% not listed; 8% indicated several racial/ethnic backgrounds.

Liking

When participants finished part 1 of the study, the researchers randomly assigned cases to be given a certain number of Likes (i.e., 0, 5, 10, 15, 20). Twenty research assistants stripped their Twitter profiles of their academic affiliation and location, and did not tweet for a week prior to this portion of the experiment. Then, each link-to-tweet was assigned to a number of "Likers" who applied a Like to each retweet. If a retweet was randomly assigned to receive five Likes, then five Likers received the link-to-tweet. As researchers verified completion of the Likers' assignments, it became apparent that the Likes appeared over a number of hours or days, rather than all at once. As to be expected, retweets occasionally garnered Likes from genuine followers and not the researchers.

Post-test dependent measures

By the end of a 5-day interval, both control and experimental participants received messages with web links directing them to the second questionnaire. Experimental participants received instructions showing how to find the readership analytics Twitter generates for each tweet, and to record the number of impressions, engagements, Likes, and retweets for both the retweet they generated as part of this study, and for their tweet or retweet that immediately preceded it. Only the number of Likes for the experimental retweet had value for this study. Control participants, alternatively, were asked to identify the story they read in part 1 out of a multiple-choice menu of six one-line story synopsis (only two of whom were incorrect). Both control and experimental participants completed an expanded Stroud et al. (2005–2006) candidate evaluation questionnaire measuring their perceptions of the politician featured in their chosen story, as described previously; $\alpha = .97$ ($M = 4.54$; $SD = 2.42$). A single item also asked participants whether they had fact-checked the story, and participants completed a three-item, nine-interval scale assessing how authentic they found the news story to be (Appelman & Sundar, 2016), $\alpha = .91$ ($M = 5.65$, $SD = 2.20$). Then, they provided demographic information, including their political party (Democrat, Republican, Independent, Other party, or No party).³ Following the questionnaire, participants read an extensive debriefing. The debriefing revealed

Table 1. Initial 2 × 2 × 5 experimental with offset 2 × 2 control research design with story valence (positive/negative) × protagonist party (Democrat/Republican) by five levels of likes after retweeting a story, versus no-retweet control condition

Story Valence	Hypothesis 1		Hypothesis 2			
	CONTROL	EXPERIMENTAL (Retweet)				
	No Retweet	0 ♥	5 ♥	10 ♥	15 ♥	20 ♥
Positive	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.
Negative	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.	Dem. / Rep.
<i>N</i>	129	97	103	93	105	95

that the stories and Likes were fake, and explained the purpose of the deception and the research. It encouraged participants to delete their retweets, and suggested other steps to mitigate any potential risk. The entire debriefing statement is available upon request.

Partisan congruence

Partisan congruence was coded based on combinations of participants' political party responses in light of whether the stories they read featured a Democrat or Republican politician and whether the stories were positive or negative with regard to the politician. The congruent condition comprised: (a) Democrat participants who shared a positive story about a Democrat politician, (b) Democrat participants who shared a negative story about a Republican politician, (c) Republican participants who shared a positive story about a Republican politician, and (d) Republican participants who shared a negative story about a Democrat politician. The incongruent condition comprised: (a) Democrat participants who shared a positive story about a Republican politician, (b) Democrat participants who shared a negative story about a Democrat politician, (c) Republican participants who shared a positive story about a Democrat politician, and (d) Republican participants who shared a negative story about a Republican politician.

An additional nonpartisan group comprised participants who identified their political affiliation as independent or explicitly as "no party," and therefore were neither congruent nor incongruent with the valence and the party affiliation associated with the politician in the respective news stories.

Exploratory analysis of fact-checking effects

Initial analysis examined whether fact checking affected participants' perceptions of the news stories' authenticity or their perceptions of the politicians described in those stories. When social media users fact-check and find a story to be false, it can immunize their perceptions regardless of other popularity metrics that may accompany a news story (Chung & Kim, 2021). On the other hand, when people fact-check and do not find a news story to have been discredited, they tend to believe the truth of the story more strongly (Pennycook et al., 2020). The stories in this study did not appear on any fact checking sites. In this study's experimental conditions, 194 participants reported that they attempted to fact-check the fake news story

and 299 participants reported that they did not; 28 control conditions participants did, and 101 did not attempt to fact-check. An ANOVA involving all experimental and control participants revealed a main effect of fact-checking attempts on perceptions of news stories' authenticity, $F(1, 588) = 148.75, p < .001$. Those who tried to fact-check experienced the news story as less authentic ($M = 4.90; SD = 2.51$) than those who did not attempt to fact-check ($M = 6.04; SD = 1.90$). No interaction effects involving fact-checking and other factors in this research affected perceptions of the news stories' authenticity. It is impossible to assess whether individuals who initially thought the stories were less authentic went to fact check them, or whether people who fact-checked them decided they were less authentic. Nevertheless, whether someone attempted to fact-check produced no main or higher-order interaction effects on their perceptions of the politician in the news story, when fact-checking was included in the model alongside other predictors (i.e., valence, congruence, and number of Likes) in the experimental condition.

Initial inspection of the number of Likes participants received identified one outlier who reported receiving 173 Likes. After removing this individual's data from further analysis, the number of Likes participants received ranged from 0 to 22 ($M = 9.92; SD = 7.14$).

Results

Hypothesis 1: Simply reading versus sharing

The analysis of H1 examined whether merely retweeting a (fake) news story affects people's perception of politicians in the stories more than merely reading them. The comparison involved the control groups (who read and selected, but did not retweet a story) and a subset of the experimental participants who retweeted but received no Likes; see the specification of these groups marked Hypothesis 1 in Table 1. Despite unequal cell sizes ($n = 97$ for those who retweeted but had no Likes, and $n = 129$ for those who merely read), Levene's test indicated homogeneity of variance among these conditions, $p = .82$. A conservative analysis required inclusion of other potentially relevant factors, since a main effect of reading versus retweeting could plausibly be moderated by other factors in this study, such as seeing positive versus negative stories, as well as partisan congruence/incongruence (i.e., the match or

mismatch between the stories' depiction of Republicans versus Democrats and participants' own identification as Republican, Democrat, or Independent). An omnibus ANOVA on perception of the politician therefore involved three factors: condition (comparing participants in the experimental group who retweeted but received no Likes, versus control group participants who were asked only to read and select one among two news stories) \times story valence \times partisan congruence/incongruence.

The ANOVA did not reveal any difference in evaluation of politicians due to whether participants merely read news about the politician, compared to participants who read, selected, and publicly retweeted with a comment a news story. A significant main effect emerged for valence, $F(1, 211) = 292.77$; $p < .001$; $\eta_p^2 = 0.56$, such that those who read or retweeted a positive story perceived the politician in the story significantly more favorably ($M = 6.35$; $SD = 1.55$), than those who read or retweeted a negative story ($M = 2.77$; $SD = 1.66$). There was also an effect for political congruence regarding positive perceptions of the target, $F(2, 211) = 3.30$; $p = .039$; $\eta_p^2 = 0.02$: Politically congruent stories yielded a mean of 4.70 ($SD = 2.55$), compared to incongruent stories' mean of 4.23 ($SD = 2.30$); Independents or participants with no party affiliation, for whom there is no in/congruence, were most favorable ($M = 4.96$; $SD = 2.37$), although these findings are difficult to interpret outside of a valence-by-congruence interaction. There were, however, no significant two-way or the three-way interaction effects to disambiguate the congruence effect.

Given that no interaction effects were significant, a reduced model ANOVA reanalyzed the data, suppressing interaction effects. Again, no significant effect on perception of the politician emerged due to reading versus retweeting, $F(1, 223) = 0.623$, $p = .431$. The first hypothesis was not supported. Contrary to previous findings in the study of identity shift, merely making a public, online selective self-presentation of one's position did not generate differences in perceptions that aligned with that presentation. Whether the additional impetus of feedback would yield a stronger effect to those who enacted such self-presentations was the focus of H2.

Hypothesis 2: Effect of likes, story valence, and party congruence

To investigate H2, hierarchical regression analysis incorporated the data from all participants in the experimental conditions, that is, participants who retweeted a news story; see the specification of these groups marked Hypothesis 2 in Table 1. Because some participants received organic Likes from Twitter users aside from those provided by researchers, the analysis employed the actual number of Likes that participants reported their retweet having received, taking the form of a continuous variable. People's perception of the politician was regressed on, first, the number of Likes the retweets received (continuous), the valence of news story (0 = negative valence; 1 = positive valence), partisan congruence (0 = congruent, with two new dummy variables created to represent the incongruent and nonpartisan conditions, both coded as 1). The second step of the regression included all two-way interaction effects among these predictors, and the third step included the hypothesized three-way interaction terms. The number of Likes was mean-centered prior to creating the interaction terms.

The final model was significant, $F(11, 461) = 70.58$, adjusted $R^2 = 0.63$, $p < .001$. Results confirmed the hypothesized three-way interactions between number of Likes, story valence, and partisan congruence. The more Likes that participants' retweets accumulated, the more negative their perceptions of the politicians in the stories were, when the news stories were negative, and when the news described an out-party politician. In fact, only when participants shared a news story that reported negatively about a politician from a rival party, as the number of Likes they received increased, the more undesirable they perceived the rival party politician to be. In contrast, when participants shared a news story that reported negatively about a politician from their own party, or when participants considered themselves independent, or when the news story was positive (no matter which party's politician it favored), an increase in the number of Likes they received did not influence their perception of the politician. See Table 2 for the full results.

Statistically, the regression coefficients (see Table 2) and the regression plot (see Figure 2) demonstrate that the three-way interaction on perception of the politician was driven mainly by the number of Likes in the negative valence/politically congruent stories condition (i.e., about an out-party politician), whereas Likes had significantly less impact on perceptions of in-party politicians from a negative news story, $b = -0.12$, $SE = 0.05$, $t(461) = -2.56$, $p = .01$, or from a negative news story read by political nonpartisans, $b = -0.10$, $SE = 0.05$, $t(461) = -2.04$, $p = .04$, $\eta_p^2 = .02$.

A simple slopes analysis also confirms the nature of the interaction effect. As seen in Table 3, only the one combination of story valence (negative), congruence, and the trend representing an increasing number of Likes yields a slope that is significantly different than zero (i.e., has a coefficient for which 0 is not within the 95% confidence intervals). The slope is negative (-0.07), indicating that the greater the number of Likes in this condition, the more negative the perception of the politician.

Results also showed a significant simple effect for valence. Participants perceived politicians more favorably from positive than from negative news stories, contingent on viewing a congruent news story and receiving the mean number of Likes. The overall influence of valence on people's perception of the politician can also be observed in Figure 2.

Beyond the three-way interaction and the main effect of valence, there were no other simple or interaction effects (see Table 2). In sum, H2 was supported with respect to the effect of getting Likes in response to retweeting negative fake news stories that denigrated a politician of the opposite political party. There was no effect of Likes in the opposite direction; that is, Likes did not enhance favorable perceptions about in-party politicians depicted in positive news stories. The type of moderation implied by the results differs somewhat from that which was hypothesized. Originally a cleaved, divergent moderation hypothesis, the results, instead, present a contingent, negatively divergent effect of Likes (see Holbert & Park, 2020). Responses to positive stories and incongruent stories were unaffected by Likes; Likes moderated valence \times congruence only at one level (negative congruent) of the other aspect of the interaction.

Discussion

This research investigated a social process that impacted people's acceptance of novel information, the information, in this case, being false. In this study, when a news story negatively

Table 2. Hierarchical regression effects of number of likes, story valence, and partisan congruence on perception of politician

Predictors	R ²	b	Standard error (SE)	t	p-value
Block 1					
Intercept		2.81	0.14	20.22	<.01
Number of Likes		-0.02	0.01	-1.59	.11
Valence		3.77	0.14	27.25	<.01
Incongruent		-0.41	0.16	-2.49	.01
Nonpartisan		-0.06	0.17	-0.32	.75
Incremental R ²	0.618				<.01
Block 2					
Number of Likes × valence		-0.01	0.02	-0.40	
Number of Likes × Incongruent		0.04	0.02	1.83	
Number of Likes × nonpartisan		0.03	0.02	1.12	
Valence × incongruent		-0.32	0.33	-0.96	
Valence × nonpartisan		-0.13	0.35	-0.36	
R ² change	0.004				.45
Incremental R ²	0.622				<.01
Block 3					
Number of Likes × valence × incongruent		-0.12	0.05	-2.56	.01
Number of Likes × valence × nonpartisan		-0.10	0.05	-2.04	.04
R ² change	0.006				.03
Total R ²	0.627				<.01

Note. Coefficients represent upon-entry estimates for each model. For valence, negative valence = 0 and positive valence = 1; for congruence, congruent = 0 and dummy codes were created for incongruent (1) and nonpartisan (1). Number of Likes was mean centered.

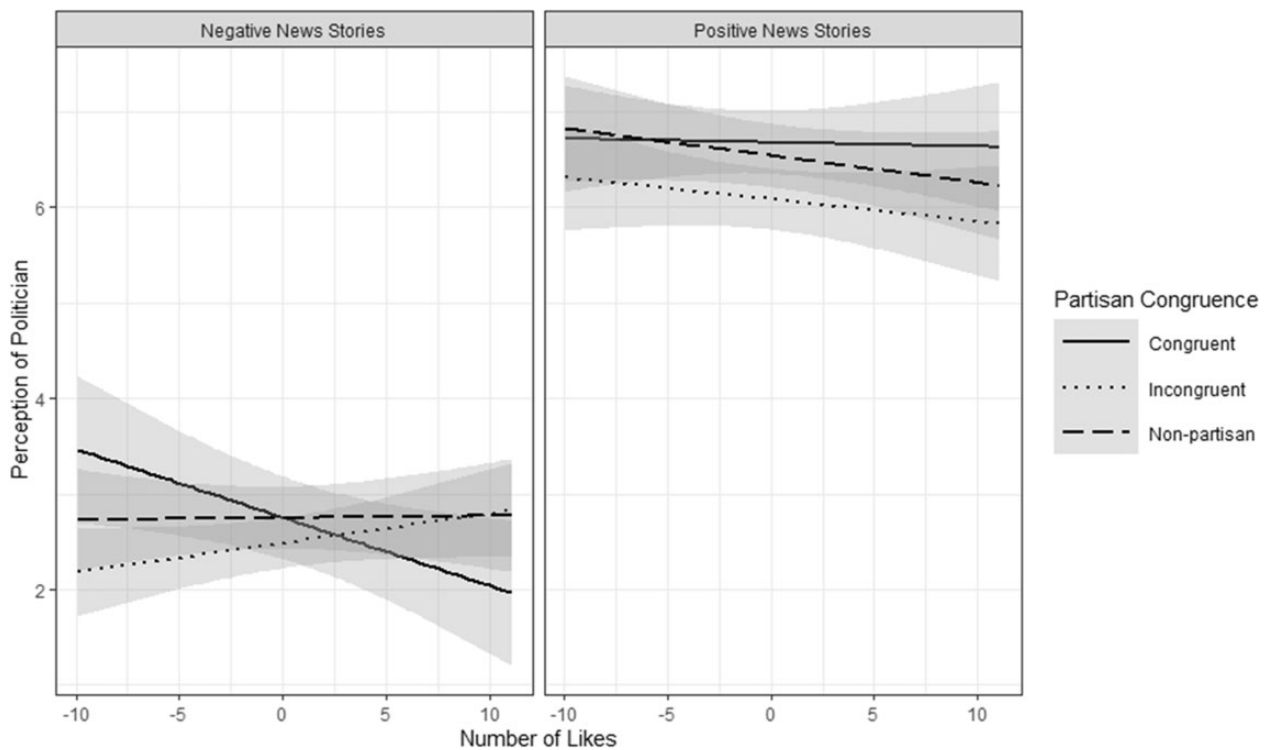


Figure 2. Effect of number of likes, story valence, and partisan congruence on perception of politician.

Note. Gray space surrounding lines represents 95% confidence interval bands.

depicted a politician of the opposite political party, the more social approval one received in response to retweeting the story, in the form of Likes, the more one’s perceptions aligned with those implied in the story. This pattern did not obtain for positive news stories: When people retweeted positive stories, they liked the politician better regardless of the politician’s, or the social media user’s, political affiliation.

This research suggests a renewed application of the hyperpersonal model of CMC, and a variation to the identity shift

paradigm that grew from it. Although the hyperpersonal model arose before the advent of social media, Walther and Whitty (2021) discussed its continued applicability to contemporary social network sites. Three particularly important processes that apply in this context include, first, the potential for selective online self-presentation—from self-description and interaction to sharing artifacts—that primes an individual for engagement and transformation of perceptions. Second, the tendency for small messages—in this case,

Table 3. Simple slopes analysis of the trends reflecting increasingly greater likes in each combination of news story valence \times partisan congruence, with corresponding subsample sizes

Valence	Congruence	<i>n</i>	Number of Likes trend	Standard error (SE)	95% confidence interval
Negative	Congruent	77	-0.072	0.024	-0.118 to -0.025
Positive	Congruent	83	-0.005	0.025	-0.053 to 0.045
Negative	Incongruent	86	0.031	0.023	-0.014 to 0.076
Positive	Incongruent	87	-0.023	0.023	-0.068 to 0.023
Negative	Nonpartisan	77	0.002	0.023	-0.043 to 0.047
Positive	Nonpartisan	63	-0.028	0.024	-0.076 to 0.020

Note. *df* = 461.

a Likes symbol and an accompanying integer—to garner potent symbolic and suasive value. Third is the impact of feedback to transform individuals' perceptions of themselves and their attitudes.

The research also raises interesting issues for the identity shift paradigm. Most previous studies showing that one's public online assertions shape one's perceptions have involved (a) assertions about oneself, that is, a self-description; rarely has this research examined individuals' explicit advocacy of some externality, and even then, (b) statements that have been primarily narrative and relatively explicit (see Carr et al., 2021, for review). In contrast, in this study, individuals' self-presentation of their initial perceptions or affiliations appeared in a very abbreviated form: A retweet that included a short comment. In this study, unlike many others, the mere public presentation of a tacit position appeared not to be sufficient to change one's perceptions, without the reinforcing effect of feedback: As seen in the analysis of H1, retweeting had no greater effect on perceptions than just reading a news story. Consistent with other replications and extensions to the identity shift paradigm, though, when affirming feedback accrued, the more one's perceptions did change, at least under the circumstances predicted by negativity and partisan bias principles.

Because there was no not-fake news condition in this study (as is also the case in many experimental studies of fake news), it cannot be said that the fakeness of news caused anything. There is no fake news effect in this study. Additional research is needed to know if these processes apply better or worse within or without the context of fake news. Within the context of fake news, however, some important and original points emerge. Research on fake news, for the most part, has focused on how individuals process fake news messages they encounter, primarily the extent to which they accept, ignore, reject, or fact-check stories they see in social media (Tandoc et al., 2020). With some important recent exceptions,⁴ it is as though fake news has a hypodermic effect once it passes thresholds of credulity and partisanship. Research on fact-checking and other warnings are consistent with this view, as inoculations against hypodermic effects (see Maertens et al., 2021). What makes social media social, however, is interactions among the readers and sharers and others beyond the point of news consumption. Factors such as those are nevertheless seldom considered among the causes of people's acceptance or rejection of news that traverses social media. The results of H1 initially seem to support such a view, having found no differences in perceptions of the news stories' focus whether one passively consumed a story or retweeted it. As the findings for H2 revealed, differences due to active retransmission depended on several other factors.

The present study asks about social influences on one's perception that go beyond mere consumption and whatever motivations participants had for willingness to share a news story. It verified the active, potentially polarizing effect of incrementally greater feedback following the sharing of fake news, that lead people to adopt increasingly more negative attitudes about out-party politicians. The present research recasts the influence of negative fake news not just as a matter of receiving it, or as a byproduct of a story's apparent popularity, but as a culmination of social process, the effects of which depend on the social sharing of destructive news consistent with one's partisan bias and receiving incrementally greater social reinforcement for having shared it.

Limitations

An important question to be asked is whether this study really applies to fake news, for which the answers appear to be both no and yes. No, because the principles in this study are not about fake news. They present a potentially broader model of how individuals accrue stronger beliefs in something that is novel—that they did not know before—that is consistent with their worldview and emotions when their public presentation of it garners incrementally greater social reinforcement. The hypotheses do not specify fake news; the methodology does. In that respect, and given its connection to identity shift dynamics, the model promises to have a more generalizable range. Ongoing research is exploring how social approval may exacerbate beliefs in other negative, even abhorrent messaging in social media. Another reason to say no is that here are numerous complex issues arising in the rapidly mounting scientific research on misinformation, disinformation, and online deception that includes “rumors, deliberately factually incorrect information, inadvertently factually incorrect information, politically slanted information, and ‘hyperpartisan’ news” (Tucker et al., 2018, p. 2), that the present study has not incorporated.

The question of whether this research pertains to fake news can also be answered, yes. The stimuli in this study were not simply fictitious experimental contrivances. They met Allcott and Gentzkow's (2017) definition, and others' criteria, quite clearly: fabricated news articles that are intentionally and verifiably false, and could mislead readers, with political implications; hosted on a website, promoted by a social media feed. The only contestable criterion is how easy they could be verified as false. Because major fact-checking organizations tend to focus on widely read news sites, the false stories on *cmcnewsnow.com* never got noticed, apparently; they did not appear on any fact-checking sites of which the researchers are aware. Fact checking is certainly one form of verifiability,

although it is far from perfect. When a false news story does not appear on a fact-checking site, its mere absence can inflate inferences about a story's veracity (see Chung & Kim, 2021, for review). As mentioned, whether participants in this study employed fact-checking sites made no significant difference in results of this study. But fact-checking sites are not the only means of verification whether a story is true or not. As Sundar et al. (2021) point out, "misinformation attributed to a public figure . . . could be verified as false upon scrutiny" (p. 303). In that respect the "checkability" (see Merpert et al., 2018) of the stimuli in this study was exceptionally strong. A search engine query of California elected officials readily dispels the existence of the State Assembly members and State Senators described by name and specific districts in the stories. Therefore, whether or not the theoretical issues and the results of this study may bear on a wider set of phenomena, they apply to the issue of belief in verifiably fake news very well.

Relatedly, there may be concerns about the use of artificial, original (fake) fake news stories about fictitious individuals rather than false stories about actual politicians. In this respect we are ethically constrained to avoid promulgating falsehoods about real people, potentially contaminating and polarizing the actual political ecology (a task better left to real politicians). Numerous other misinformation studies have also employed fictitious rather than authentic news stimuli (e.g., Dai et al., 2021; Ecker et al., 2010). One advantage to using artificially invented events and persons, however, is enhanced confidence that no participant had pre-existing biases for or against the specific characters in the stories, that could otherwise contaminate results. Moreover, according to Maertens et al. (2021, p. 5), "Using fictional items also maximizes experimental control over isolating the manipulation techniques and avoids familiarity confounds with real fake news."

The present study also faces questions about limits to internal and external validity due to the manufactured Likes, and the answers to those questions are not clear-cut. One constraint of the experiment was that the Likes appended to participants' retweets originated with strangers rather than participants' friends or followers. Carr et al. (2018) showed that feedback from friends has bigger impacts on self-concept change than does feedback from strangers. On that basis one can surmise that this research underestimates the effect that takes place in spontaneous social media use when friends and followers Like a message, if and when Likes appear.

Two competing concerns arise with regard to the representativeness of the final sample. On the one hand, it is possible that the experiment impelled participants to share fake news who otherwise might not have. In that regard, the attrition of 346 prospective participants is notable. Altay et al. (2022) found that most people do not share fake news, for fear of damaging their reputation. Perhaps those who left the study are of this ilk, leaving more naïve or more extreme partisans in the experimental sample.

This does not negate the possibility that the remaining sample's responses are more extreme than those who share fake news outside experimental settings. Yet the pattern of negativity and partisan congruence seem consistent with other accounts of normal political behavior online. Barberá's (2015) analysis of 15 million political tweets preceding the US 2012 presidential election between Democratic candidate Obama and Republican candidate Romney found that

conservatives tweeted about Obama more frequently than did liberals, and liberals tweeted more often about Romney than did conservatives. Barberá deduced "that most tweets sent during this period were negative" (p. 87), and that politically moderate users tweeted less often than those who had more extreme political orientations.

As a field experiment, this study was subject to naturally occurring volatility that may or may not be present in other studies or replications. The research took place within the months immediately preceding the 2020 US presidential election, during which political news stories were frequent and highly charged. Additionally, there was great controversy over federal, state, and local governments', as well as individuals' responses to COVID-19, in terms of its legitimate danger and the efficacy of various behavioral and biological measures to combat it. News often combined politics and COVID-19 issues (see Hart et al., 2020), and "there was serious concern among the U.S. public about the potential effect of made-up news and information during the 2020 presidential election" (Watson, 2021, n.p.).

Notwithstanding these limitations, the field experiment offers several advantages that are relatively uncommon in other experimental or field research on social media. By recruiting participants on Facebook who were also active on Twitter, the study sampled a fairly authentic population. They composed their own retweets' comments, and posted the messages in the actual social media platform that anyone in the world could see. The design and analysis, moreover, was able to isolate the effects of reading versus retweeting fake news, and retweeting alone versus getting Likes, each comparison providing a check for the relative contributions of each plausible cause in the acceptance of false, negative, out-party perceptions that transpire through the selective self-presentation of fake news and its reinforcement through signals of social approval.

Supplementary material

Supplementary material is available at *Journal of Communication* online.

Data availability statement

The data for this article cannot be shared publicly due to commitments to participants in the study made in order to facilitate responses that could be used to reveal individual identifiers.

Funding

This research was funded by donations to the UCSB Center for Information Technology and Society.

Acknowledgments

The authors express sincere gratitude to a number of research assistants who helped in every stage of this project: Jake Quercia, Julia Trites, Chrissy Berry, Elizabeth Siy, Lea Hassakorzian, Paula Wang, Rachel Wu, Erin Valenzuela, Myah Mashhadialireza, Kashara Ryno, and Marina Alvarez. The authors also thank the editor and reviewers for their insightful suggestions. A previous version of this research was

presented at the 2022 conference of the International Communication Association.

Conflict of interest statement: None declared.

Notes

1. The term “paralinguistic digital affordances” (Hayes et al., 2016) includes Likes and other icons that individuals can append to others’ social media messages to connote love, shock, anger, tears, astonishment, etc. This study utilized only Twitter, mindful of Hayes et al.’s conclusion that people deploy and interpret Facebook’s Likes and Twitter’s somewhat differently: Facebook’s Likes often reflect mindless reactions to friends, whereas Twitter’s are more deliberate, less frequent, and more definitively connote approval.
2. These combinations of party by valence were the basis for computation of the partisan congruence/incongruence variable in the data analysis that follows. Additionally, a third level was computed when participants indicated that they were political Independents, in which case they were neither congruent nor incongruent, but neutral, with respect to the political party and valence of the stimuli.
3. Although there are a number of ways to assess individuals’ political orientation, the relevant classifications for this research focused on a limited identification of major party affiliation or declaration of partisan independence. Among participants in experimental conditions, 20 individuals indicated affiliations with other parties—neither Democrat, Republican, registered Independent, or no party whatsoever—whose positions were vague with respect to their relationship to Democrat or Republicans. These 20 participants were excluded from all analyses that involved partisan congruence. Other approaches to assessing political identity include strength of party identification and/or political ideology (see, e.g., Bailenson et al., 2008, that also employed fictitious political candidates). Such approaches, however, provide little advantage over party affiliation for the purposes of assessing party congruence with the Democrat or Republican politicians featured in the news stories, compared to research about fake news that takes into theoretical consideration finer-grained variations in political ideology (e.g., Peacock et al., 2021).
4. For instance, Metzger et al. (2021) found that partisans often append comments to fake news explicitly discrediting the content, warning others not to believe the story, sharing fake news in order to ridicule the opposition’s dishonesty. Following Metzger et al.’s (2021) findings, an informal inspection commenced of the user-generated comments accompanying the retweets in this study. The comments showed little evidence of undermining the story. The vast majority of comments ranged from disparaging the particular politician and cynicism about politicians in the negative stories, to expressing admiration for the politicians in the positive stories, urging support for the fighting COVID-19 generally, and urging others simply to read the story. Only five comments used the word “fake” or expressed overt skepticism about the truth of the stories. In four of those cases, the participant was a Democrat and the story depicted a Democrat negatively, consistent with the strategy articulated by Metzger et al. and the general findings of this study.

References

- Allcott, H., Boxell, L., Conway, J. C., Gentzkow, M., Thaler, M., & Yang, D. Y. (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus epidemic. *Journal of Public Economics*, 191, 104254. <https://doi.org/10.1016/j.jpubeco.2020.104254>
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, 6(2), 205316801984855–84858. <https://doi.org/10.1177/2053168019848554>
- Altay, S., Hacquin, A.-S., & Mercier, H. (2022). Why do so few people share fake news? It hurts their reputation. *New Media & Society*, 24(6), 1303–1324. <https://doi.org/10.1177/1461444820969893>
- Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, 93(1), 59–79. <https://doi.org/10.1177/1077699015606057>
- Bailenson, J. N., Iyengar, S., Yee, N., & Collins, N. A. (2008). Facial similarity between voters and candidates causes influence. *Public Opinion Quarterly*, 72(5), 935–961. <https://doi.org/10.1093/poq/nfn064>
- Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, 23(1), 76–91. <https://doi.org/10.1093/pan/mpu011>
- Baron, R. S., Hoppe, S. I., Kao, C. F., Brunzman, B., Linneweh, B., & Rogers, D. (1996). Social corroboration and opinion extremity. *Journal of Experimental Social Psychology*, 32(6), 537–560. <https://doi.org/10.1006/jesp.1996.0024>
- Barthel, M., Mitchell, A., & Holcomb, J. (2016, December 16). *Many Americans believe fake news is sowing confusion*. Pew Research Center. <http://web.archive.org/web/20210511204650/https://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/>
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370. <https://doi.org/10.1037/1089-2680.5.4.323>
- Carr, C. T., & Foreman, A. C. (2016). Identity shift III: Effects of publicness of feedback and relational closeness in computer-mediated communication. *Media Psychology*, 19(2), 334–358. <https://doi.org/10.1080/15213269.2015.1049276>
- Carr, C. T., & Hayes, R. A. (2019). Identity shift effects of self-presentation and confirmatory and disconfirmatory feedback on self-perceptions of brand identification. *Media Psychology*, 22(3), 418–444. <https://doi.org/10.1080/15213269.2017.1396228>
- Carr, C. T., Hayes, R. A., & Sumner, E. M. (2018). Predicting a threshold of perceived Facebook post success via Likes and reactions: A test of explanatory mechanisms. *Communication Research Reports*, 35(2), 141–151. <https://doi.org/10.1080/08824096.2017.1409618>
- Carr, C. T., Kim, Y., Valov, J. J., Rosenbaum, J. E., Johnson, B. K., Hancock, J. T., & Gonzales, A. L. (2021). An explication of identity shift theory. *Journal of Media Psychology*, 33(4), 202–214. <http://dx.doi.org/10.1027/1864-1105/a000314>
- Carr, C. T., Wohn, D. Y., & Hayes, R. A. (2016). Thumbs Up as social support: Relational closeness, automaticity, and interpreting social support from paralinguistic digital affordances in social media. *Computers in Human Behavior*, 62, 385–393. <https://doi.org/10.1016/j.chb.2016.03.087>
- Cho, J., Ahmed, S., Keum, H., Choi, Y. J., & Lee, J. H. (2018). Influencing myself: Self-reinforcement through online political expression. *Communication Research*, 45(1), 83–111. <https://doi.org/10.1177/0093650216644020>
- Chung, M., & Kim, N. (2021). When I learn the news is false: How fact-checking information stems the spread of fake news via third-person perception. *Human Communication Research*, 47(1), 1–24. <https://doi.org/10.1093/hcr/hqaa010>
- Crockett, M. J. (2017). Moral outrage in the digital age. *Nature Human Behaviour*, 1(11), 769–771. <https://doi.org/10.1038/s41562-017-0213-3>
- Dai, Y., Yu, W., & Shen, F. (2021). The effects of message order and debiasing information in misinformation correction. *International Journal of Communication*, 15, 1039–1059. <https://ijoc.org/index.php/ijoc/article/view/15364>
- Duffy, A., Tandoc, E., & Ling, R. (2020). Too good to be true, too good not to share: The social utility of fake news. *Information, Communication, & Society*, 23(13), 1965–1979. <https://doi.org/10.1080/1369118X.2019.1623904>

- Ecker, U. K., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38(8), 1087–1100. <https://doi.org/10.3758/MC.38.8.1087>
- Flanagin, A. J. (2017). Online social influence and the convergence of mass and interpersonal communication. *Human Communication Research*, 43(4), 450–463. <https://doi.org/10.1111/hcre.12116>
- Gonzales, A. L., & Hancock, J. T. (2008). Identity shift in computer-mediated environments. *Media Psychology*, 11(2), 167–185. <https://doi.org/10.1080/15213260802023433>
- Guess, A., Nyhan, B., & Reifler, J. (2018, January 9). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential election*. Unpublished manuscript, Princeton University Center for the Study of Democratic Politics. <http://web.archive.org/web/20210109150406/https://csdp.princeton.edu/publications/selective-exposure-misinformation-evidence-consumption-fake-news-during-2016-us>.
- Guess, A., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, 4(5), 472–480. <https://doi.org/10.1038/s41562-020-0833-x>
- Haim, M., Kümpel, A. S., & Brosius, H.-B. (2018). Popularity cues in online media: A review of conceptualizations, operationalizations, and general effects. *Studies in Communication and Media*, 7(2), 186–207. <https://doi.org/10.5771/2192-4007-2018-2-58>
- Hart, S. P., Chinn, S., & Soroka, S. (2020). Politicization and polarization in COVID-19 news coverage. *Science Communication*, 42(5), 679–697. <https://doi.org/10.1177/1075547020950735>
- Hayes, R. A., Carr, C. T., & Wohn, D. Y. (2016). One click, many meanings: Interpreting paralinguistic digital affordances in social media. *Journal of Broadcasting & Electronic Media*, 60(1), 171–187. <https://doi.org/10.1080/08838151.2015.1127248>
- Holbert, R. L., & Park, E. (2020). Conceptualizing, organizing, and positioning moderation in communication research. *Communication Theory*, 30(3), 227–246. <https://doi.org/10.1093/ct/qtz006>
- Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology: Applied*, 27(1), 1–16. <https://doi.org/10.1037/xap0000315>
- Merpert, A., Furman, M., Anauati, M. V., Zommer, L., & Taylor, I. (2018). Is that even checkable? An experimental study in identifying checkable statements in political discourse. *Communication Research Reports*, 35(1), 48–57. <https://doi.org/10.1080/08824096.2017.1366303>
- Metzger, M. J., Flanagin, A. J., Mena, P., Jiang, S., & Wilson, C. (2021). From dark to light: The many shades of sharing misinformation online. *Media and Communication*, 9(1), 134–143. <https://doi.org/10.17645/mac.v9i1.3409>
- Neyazi, T. B., & Muhtadi, B. (2021). Selective belief: How partisanship drives belief in misinformation. *International Journal of Communication*, 15, 1286–1308. <https://ijoc.org/index.php/ijoc/article/view/15477/3382>.
- Peacock, C., Hoewe, J., Panek, E., & Willis, G. P. (2021). Hyperpartisan news use: Relationships with partisanship and cognitive and affective involvement. *Mass Communication and Society*, 24(2), 210–232. <https://doi.org/10.1080/15205436.2020.1844902>
- Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, 66(11), 4944–4957. <https://doi.org/10.1287/mnsc.2019.3478>
- Pingree, R. J. (2007). How messages affect their senders: A more general model of message effects and implications for deliberation. *Communication Theory*, 17(4), 439–461. -2885.2007.00306.x
- PolitiFact Staff. (2017, April 20). *PolitiFact's guide to fake news websites and what they peddle*. PolitiFact. <https://web.archive.org/web/20220314150533/https://www.politifact.com/article/2017/apr/20/politifact-guide-fake-news-websites-and-what-they/>
- Reich, S., Schneider, F. M., & Heling, L. (2018). Zero Likes—symbolic interactions and need satisfaction online. *Computers in Human Behavior*, 80, 97–102. <https://doi.org/10.1016/j.chb.2017.10.043>
- Rosenthal-von der Pütten, A., Hastall, M. R., Köcher, S., Meske, C., Heinrich, T., Labrenz, F., & Ocklenburg, S. (2019). “Likes” as social rewards: Their role in online social comparison and decisions to like other people’s selfies. *Computers in Human Behavior*, 92, 76–86. <https://psycnet.apa.org/doi/10.1016/j.chb.2018.10.017>
- Schlenker, B. R., Dlugolecki, D. W., & Doherty, K. (1994). The Impact of Self-Presentations on Self-Appraisals and Behavior: The Power of Public Commitment. *Personality and Social Psychology Bulletin*, 20(1), 20–33.
- Scissors, L., Burke, M., & Wengrovitz, S. (2016). *What’s in a Like? Attitudes & behaviors around receiving Likes on Facebook*. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, San Francisco, CA, USA, 26 February–2 March 2016, pp. 1501–1510. <http://dx.doi.org/10.1145/2818048.2820066>
- Solon, O. (2017, November 9). Ex-Facebook president Sean Parker: Site made to exploit human ‘vulnerability’. *The Guardian*. <http://web.archive.org/web/20210510105453/https://www.theguardian.com/technology/2017/nov/09/facebook-sean-parker-vulnerability-brain-psychology>
- Stroud, L. R., Glaser, J., & Salovey, P. (2005). The effects of partisanship and candidate emotionality on voter preference. *Imagination, Cognition and Personality*, 25(1), 25–44. <https://doi.org/10.2190/22PC-5PK7-5KU0-DNCQ>
- Sundar, S. S., Molina, M. D., & Cho, E. (2021). Seeing is believing: Is video modality more powerful in spreading fake news via online messaging apps? *Journal of Computer-Mediated Communication*, 26(6), 301–319. <https://doi.org/10.1093/jcmc/zmab010>
- Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2018). Defining “fake news”: A typology of scholarly definitions. *Digital Journalism*, 6(2), 137–153. <https://doi.org/10.1080/21670811.2017.1360143>
- Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, 21(3), 381–398. <https://doi.org/10.1177/1464884919868325>
- Tokunaga, R. S., & Quick, J. D. (2018). Impressions on social networking sites: Examining the influence of frequency of status updates and “Likes” on judgments of observers. *Media Psychology*, 21(2), 157–181. <https://doi.org/10.1080/15213269.2017.1282874>
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., . . . Nyahn, B. (2018, March). *Social media, political polarization, and political disinformation: A review of the scientific literature*. William + Flora Hewlett Foundation. <https://www.hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf>.
- Valkenburg, P. M. (2017). Understanding self-effects in social media. *Human Communication Research*, 43(4), 477–490. <https://doi.org/10.1111/hcre.12113>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23(1), 3–43. <https://doi.org/10.1177/009365096023001001>
- Walther, J. B., & Jang, J.-W. (2012). Communication processes in participatory web sites. *Journal of Computer-Mediated Communication*, 18(1), 2–15. <https://doi.org/10.1111/j.1083-6101.2012.01592.x>
- Walther, J. B., Liang, Y. J., DeAndrea, D. C., Tong, S. T., Carr, C. T., Spottswood, E. L., & Amichai-Hamburger, Y. (2011). The effect of feedback on identity shift in computer-mediated communication. *Media Psychology*, 14(1), 1–26. <https://doi.org/10.1080/15213269.2010.547832>
- Walther, J. B., Van Der Heide, B., Tong, S. T., Carr, C. T., & Atkin, C. K. (2010). The effects of interpersonal goals on inadvertent intrapersonal influence in computer-mediated communication. *Human Communication Research*, 36(3), 323–347. <https://doi.org/10.1111/j.1468-2958.2010.01378.x>

- Walther, J. B., & Whitty, M. T. (2021). Language, psychology, and new new media: The hyperpersonal model of mediated communication at twenty-five years. *Journal of Language and Social Psychology, 40*(1), 120–135. <https://doi.org/10.1177/0261927X20967703>
- Watson, A. (2020, June 23). *Social media as a news source worldwide 2020*. Statista. <http://web.archive.org/web/20210414064702/https://www.statista.com/statistics/718019/social-media-news-source/>
- Watson, A. (2021, June 16). *Fake news in the U.S.—Statistics & facts*. Statista. <https://www.statista.com/topics/3251/fake-news/>.
- Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication, 65*(4), 699–719. <https://doi.org/10.1111/jcom.12164>
- Winter, S., Metzger, M. J., & Flanagin, A. J. (2016). Selective use of news cues: A multiple-motive perspective on information selection in social media environments. *Journal of Communication, 66*(4), 669–693. <https://doi.org/10.1111/jcom.12241>
- Wolf, W., Levordashka, A., Ruff, J. R., Kraaijeveld, S., Lueckmann, J.-M., & Williams, K. D. (2015). Ostracism online: A social media ostracism paradigm. *Behavior Research Methods, 47*(2), 361–373. <https://doi.org/10.3758/s13428-014-0475-x>