

No Comments (from You): Understanding the Interpersonal and Professional Consequences of Disabling Social Media Comments

Michelle E. Daniels and Freeman Wu

Abstract

Presumably to reduce cyberbullying and promote mental health, online influencers often limit viewers' ability to post social media comments. In this research, the authors find that influencers incur significant interpersonal and professional repercussions for doing so. Across seven studies, including a Twitter dataset and six experiments utilizing both consequential and hypothetical dependent measures, the authors find that consumers form more negative impressions of and are less persuaded by influencers who disable social media comments. These outcomes are driven by the perception that the influencer is less receptive to consumer voice (e.g., consumers' thoughts, opinions, and suggestions) and, thus, less sincere. However, the authors find that this effect is mitigated in situations where consumers believe that it is reasonable for influencers to prioritize self-protection.

Keywords

social media, online comments, influencer marketing, consumer voice, sincerity, mental health, parasocial

Online supplement: <https://doi.org/10.1177/00222429241252842>

Submitted December 9, 2022

Online influencers—individuals who have built large social media followings by sharing both personal stories and branded content (De Veirman, Cauberghe, and Hudders 2017)—continue to grow in popularity among consumers and brands alike (Lamberton and Stephen 2016; Leung, Gu, and Palmatier 2022). To appeal to a broad audience, influencers are expected to be approachable and engage with followers in a relatively intimate manner (Schouten, Janssen, and Verspaget 2020), but this approachability can come at a personal cost. As consumers become accustomed to influencers' accessibility, they may feel emboldened to share their feedback—both supportive and critical—online. This constant deluge of user feedback can, at best, feel overwhelming and, at worst, be severely detrimental to influencers' mental health. In fact, a number of high-profile celebrity and influencer suicides around the world have been directly linked to cyberbullying and incessant hateful online comments (Kim 2020; McCurry 2022; Stewart 2022).

One way for influencers to cope with this level of user feedback is to disable their social media comments (Gallucci 2019). Indeed, a pretest with social media users revealed a strong lay

belief that social media users, especially influencers, can improve their mental health by turning off their comments. Further, disabling comments is a prevalent practice; nearly 80% of social media users surveyed can recall encountering public figures (e.g., influencers and celebrities) who did so (for pretest details, see Web Appendix A). But is disabling comments an effective solution for these individuals?

Despite recognizing that disabling comments can potentially improve influencers' mental health, we argue that people nonetheless penalize influencers for utilizing this feature. In the present research, we find that when influencers turn off their comments, they signal that they are less receptive to consumers' thoughts, opinions, and suggestions—a construct that we define as “consumer voice,” which we adapt from the employee voice literature (e.g., Morrison 2014; Van Dyne and LePine 1998). Consequently, such influencers are judged as less sincere and

Michelle E. Daniels is Assistant Professor of Marketing, Culverhouse College of Business, University of Alabama, USA (email: medaniels2@ua.edu). Freeman Wu is Assistant Professor of Marketing, Owen Graduate School of Management, Vanderbilt University, USA (email: freeman.wu@vanderbilt.edu).

incur both interpersonal and professional consequences. However, this effect is attenuated in situations where consumers believe that it is reasonable for influencers to prioritize self-protection.

We make several important theoretical and substantive contributions. First, we add to the burgeoning influencer marketing literature (e.g., Cascio Rizzo et al. 2023; Leung et al. 2022) by documenting a pervasive yet previously unexplored online behavior, disabling social media comments, that can undermine influencer effectiveness. Second, drawing from research on employee voice (Morrison 2014), we introduce the notion of consumer voice to marketing and explicate the role of voice receptiveness perceptions in shaping consumer–influencer interactions. Relatedly, extant research has primarily focused on the upsides of parasocial interactions for influencer effectiveness (Jin and Ryu 2020). Yet our work argues that because parasocial interactions can increase expectations for voice receptiveness, they may also unintentionally create more hostile environments for influencers when they try to establish boundaries with the public. More broadly, given that social interactions are increasingly taking place online, we add to a growing body of work on consumer inference-making in digital environments (e.g., Appel et al. 2020; Li, Chan, and Kim 2019). Substantively, this research informs marketing practitioners by shedding light on the factors that can undermine the effectiveness of influencer–brand partnerships, while offering a tangible and actionable solution to minimize this backlash.

Conceptual Framework

Consumer–Influencer Interactions

One of online influencers' primary strengths is their ability to interact with their followers in a relatively intimate and informal manner (De Veirman, Cauberghe, and Hudders 2017), which makes them appear sincere and approachable (Ault 2014; Ferchaud et al. 2018). These positive judgments are made possible by the various ways influencers strategically engage with viewers, including directly addressing viewers in their posts and sharing seemingly exclusive “insider information” (Abidin 2015; Reinikainen et al. 2020; Tolson 2010). Such conversational intimacy increases the likelihood of experiencing a parasocial interaction, an illusory sense of mutual awareness in which observers feel as though they are engaging in give-and-take dialogues with a media personality (Horton and Wohl 1956; Labrecque 2014). Yet, in reality, these interactions are often one-sided (Kowert and Daniel 2021).

Although the intimacy that characterizes parasocial interactions can make influencers seem more approachable and accessible, in turn enhancing their persuasiveness (Schouten, Janssen, and Verspaget 2020), it can also shape what consumers come to expect from these individuals. The seemingly personal and two-sided nature of consumer–influencer interactions may lead consumers to develop heightened expectations regarding access to influencers' attention and feel entitled to such access (Grubbs and Exline 2016), especially once influencers have established a norm of approachability. This entitlement is

further strengthened by the pivotal role that everyday consumers play in determining an influencer's overall worth and value. Because influencers' professional standing is largely determined by the size of their online following (De Veirman, Cauberghe, and Hudders 2017), as their platform grows, consumers may feel as though they contributed to the influencer's professional success, a belief that is reinforced by the gratitude that influencers often express for their followers when they reach certain follower count milestones (e.g., celebrating 100,000 followers) (Leung et al. 2022).

Notably, although traditional celebrities first achieved success in formal institutions (e.g., acting, music, sports), those who are active on social media often assume many of the same roles as influencers and utilize their online clout for professional gain. For example, Cristiano Ronaldo, one of the most popular soccer players of all time, earns an estimated \$3.2 million per sponsored post, while Selena Gomez, a famed singer and actress, commands \$2.6 million (Geyser 2024a). Their ability to demand such figures is driven, at least in part, by the size of their respective follower bases. As such, although not all celebrities are influencers, those who maintain an active social media presence should likewise be expected to engage with followers in a similar manner. In sum, we contend that consumers not only feel entitled to share their thoughts, opinions, and suggestions with influencers but also expect these individuals to be receptive to their feedback. Next, we conceptualize a relevant marketing construct that captures these expectations and discuss how they impact reactions to influencers who disable social media comments.

Consumer Voice Receptiveness

To better understand consumers' inclination to share feedback with influencers and their expectations for how this feedback should be received, we draw from organizational theory on employee voice, which refers to an employee's informal and discretionary communication of work-related concerns, ideas, and suggestions (Morrison 2014). According to this literature, employees voice their opinions to managers, typically with the intention of solving workplace issues and improving company functioning (Van Dyne and LePine 1998). Likewise, consumers appreciate the opportunity to express themselves in the marketplace, whether in person or online (Lovett, Peres, and Shachar 2013). Accordingly, we define “consumer voice” as a consumer's informal and discretionary communication of thoughts, opinions, and suggestions to another marketplace entity.

Importantly, consumer voice expression is ubiquitous on social media. As a recent instantiation of such voice expression, Tieghan Gerard, an acclaimed food influencer who runs the blog *Half Baked Harvest*, shared a recipe that many social media users deemed as misrepresentative of Vietnamese cuisine, sparking a barrage of comments (Nguyen 2023). For example, one user noted on Gerard's (2023) post, “I really wish you'd delete this and start over. We all make mistakes, but you have such a large audience and owe it to them to do the most basic research if you're going to present reworked dishes from

other cultures.” Notably, although this comment provided constructive feedback, not all voices are intended to offer suggestions for improvement. Indeed, within the employee voice literature, certain voices are destructive (e.g., expressing hurtful, critical, or debasing opinions), while others are supportive (e.g., expressing support for or speaking out in defense of something) (Maynes and Podsakoff 2014). In the same vein, we argue that consumer voice can also be destructive, as evidenced by the deluge of harassment observed in social media comments (Hassan et al. 2018), or supportive, as evidenced by the legions of loyal fans who readily come to an influencer’s defense when their idol is being criticized (e.g., *Just Jared* 2023). For example, to belittle Gerard, one user commented, “It makes sense why your account is half-baked—it shows in your ideas,” whereas to support her, another user commented, “This recipe is SO good!! Thank you for sharing! Obsessed.”

Given that social media readily facilitates consumer voice expression through the ability to leave and exchange comments, the accessibility of the influencer’s comments section should play a crucial role in shaping perceptions of their receptiveness to consumer voice. When comments are accessible, consumers can reasonably infer that the influencer might read, and perhaps even engage with, user feedback. However, when an influencer disables access to their comments, they make the one-sided nature of their communications salient and disrupt the illusion of conversational intimacy that typically characterizes consumer–influencer interactions. In doing so, we assert that consumers will infer that an influencer who disables comments is less receptive to their voice, which they feel entitled to share. Importantly, based on our theorizing, consumer voice receptiveness captures an influencer’s overall receptivity to consumers’ open communication of thoughts, opinions, and suggestions, whether intended to be constructive, destructive, or supportive in nature. As such, disabling comments elicits the perception that the influencer is less receptive to the consumer’s voice, an inference that should, in turn, negatively impact dispositional judgments. We predict that the impaired perceptions of consumer voice receptiveness stemming from disabling comments will cause the influencer to be judged as less sincere, a relationship we explore in the following section.

Sincerity

As detailed in the previous section, influencers cultivate their following by developing a relatable and approachable online persona, which raises the expectation that they will be receptive to consumer voice. That is, people infer, through the influencer’s approachable persona, that they are genuinely interested in what others have to say. Accordingly, we define sincerity as the degree to which one’s outward behaviors honestly convey their own internal beliefs to others (Gillath et al. 2010; Trilling 1973). We focus on sincerity judgments, as opposed to the related, yet distinct, concept of authenticity, because sincerity specifically pertains to “honesty when conveying thoughts and feelings to others” (Gillath et al. 2010, p. 841), which is more germane to the present research, whereas authenticity is more about being honest to oneself (Erickson 1995; Gillath et al. 2010; Trilling 1973).

When an influencer signals lower receptiveness to consumer voice by disabling comments, such inferences violate what is expected of influencers and undermine their perceived approachability. This violation will lead consumers to infer that the influencer’s online persona may have simply been a facade engineered for professional gain (Campbell and Kirmani 2000), rather than an honest reflection of their disposition. As a result, consumers will judge influencers who appear less receptive to their voice as less sincere.

We expect influencers to incur both interpersonal and professional consequences due to these negative judgments of sincerity, given that sincerity perceptions play a critical role in shaping both interpersonal liking (Liden and Mitchell 1988; Nguyen, Seer, and Hartman 2008) and endorsement effectiveness (Filieri et al. 2023; Ohanian 1990). Indeed, people who are perceived as less sincere are generally less effective at persuading others (Barasch, Berman, and Small 2016). Consequently, disabling social media comments should reduce the influencer’s perceived voice receptiveness and sincerity, in turn undermining overall impressions and the persuasiveness of their endorsements. More formally:

H₁: Consumers form more negative impressions of and are less persuaded by online influencers who disable (vs. enable) social media comments.

H₂: The relationship between comment status (i.e., disabled vs. enabled) and impressions, as well as between comment status and persuasiveness, is sequentially mediated by perceived consumer voice receptiveness and sincerity.

The Role of Comment Valence

Notably, the fact that a comment becomes publicly accessible raises the intriguing question of whether it would be more reputationally damaging to enable comments and allow negative information about oneself to be visible or disable comments and block access to such information altogether. On the one hand, because consumers are prone to overweighting negative information (Ito et al. 1998), it is conceivable that influencers would suffer just as many, if not more, reputational consequences for enabling comments and disclosing negative information about themselves. On the other hand, by allowing such information to remain publicly visible, influencers should be considered more receptive to consumer voice, and thus more sincere, than those who disable access to their comments. Therefore, despite the possibility for negative information to taint consumer judgments, we predict that influencers who disable comments will be evaluated more negatively than those who leave negative comments visible. For the same reason, we expect consumers to respond more negatively to influencers who disable comments than those who keep positive or mixed-valence comments publicly visible.

When Prioritizing Self-Protection Is Reasonable

Our conceptualization inspires a theoretically grounded yet substantively impactful moderator. When an influencer disables

comments, we argue that they signal lower receptiveness to consumer voice and are thus judged as less sincere. Notably, a critical assumption of our theorizing is that these reduced perceptions of voice receptiveness are, in fact, believed to be indicative of their disposition (i.e., sincerity). However, based on correspondence bias, the link between an individual's observable behavior and ensuing dispositional attributions can be disrupted by relevant contextual cues (Gilbert and Malone 1995; Jones and Nisbett 1987). Thus, we anticipate that in certain situations, consumers may intuit that the influencer's perceived lack of voice receptiveness from disabling comments is driven by situational forces (Fein, Hilton, and Miller 1990; Kelley 1973) and therefore less reflective of their underlying sincerity.

What types of contexts might cause consumers to rely less heavily on the influencer's behavior to form dispositional judgments, attenuating the relationship between perceived voice receptiveness and sincerity? In the present research, we focus on situations where consumers believe that it would be reasonable for influencers to prioritize self-protection, or the goal to protect oneself against social pain. Prior research notes that people often balance the goal of seeking social connectedness against the competing goal of protecting themselves from social pain, such that those who are feeling particularly vulnerable to the latter are more likely to withdraw from others (Murray et al. 2008; Murray, Holmes, and Collins 2006). Accordingly, we argue that consumers can recognize when acts of self-protection are more reasonable and thus less reflective of the influencer's dispositional sincerity. For example, when someone is grieving a personal loss, their decision to disable comments may be considered a reasonable form of self-protection given their state of vulnerability, despite its negative impact on their perceived voice receptiveness. Likewise, when an influencer publicly discloses their struggles with mental health, observers should recognize that the influencer's desire to protect their own well-being outweighs their apparent lack of voice receptiveness and thus judge them less harshly, suggesting a potential intervention to mitigate the backlash against disabling comments. Critically, an influencer cannot justifiably disable comments in response to all personally difficult situations; only those where consumers believe it is reasonable for the influencer to prioritize self-protection will attenuate this relationship. For example, although an influencer may wish to prioritize self-protection after a transgression by disabling comments, doing so would prevent them from learning from and correcting their mistakes (Baker and McNulty 2011); thus, consumers are unlikely to view their decision to disable comments as reasonable in this case. More formally:

H₃: The relationship between consumer voice receptiveness and sincerity perceptions is attenuated in situations where consumers believe that it is reasonable for influencers to prioritize self-protection.

Overview of Current Studies

We test our predictions across seven studies, including four preregistered experiments (<https://researchbox.org/2658>). Using real-world

quasi-experimental data from Twitter (now known as "X"), Study 1 documents the relationship between turning off comments and negative online word of mouth (WOM), while Study 2 gauges spontaneous reactions to influencers who disable social media comments. Next, using consequential measures of engagement, Studies 3a and 3b test H₁ by assessing the impact of disabling comments on consumer interest in learning about an endorsed brand and their choice of promotional discount codes, two major sources of income for influencers. Importantly, Study 3b also tests our full conceptual model via mediation (H₂). Next, Study 4 compares, in the context of an apology following a transgression, the reputational costs of disabling comments to enabling negative, positive, and mixed-valence comments. Finally, our last two studies examine the moderating role of situations where consumers believe that it would be reasonable for influencers to prioritize self-protection (H₃). Study 5 demonstrates that the negative effect of disabling comments is attenuated when an influencer is grieving a personal loss, while Study 6 shows that the backlash is eliminated when the influencer discloses social media's toll on their mental health. In doing so, Study 6 also offers a practical and actionable intervention for influencers who decide to turn off comments. Across studies, unless otherwise noted, we aimed to recruit a minimum of 200 participants per cell for studies with continuous dependent variables and 350 participants per cell for studies with binary outcomes. All data, questionnaires, and syntax are available on OSF (<https://osf.io/zpj2x/>). See Figure 1 for our conceptual framework and study overview.

Study 1

A central tenet of our theorizing is that people react negatively to the disabling of social media comments. To offer initial support for this hypothesis in an ecologically valid setting, Study 1 analyzes tweets scraped from Twitter—a platform commonly used by individuals to express their opinions of others' behaviors (Xiong and Liu 2014). We examine tweets discussing reactions to social media users who turn off their comments and compare them with a random sample of tweets posted in the same time frame. In doing so, we explore whether decisions to disable comments can spark negative WOM.

Method

To explore WOM regarding the disabling of social media comments, we conducted a quasi-experiment on Twitter. Through Twitter's application programming interface, researchers can scrape all public tweets from the past nine days.¹ Utilizing the R package *rtweet* (Kearney 2019), we first collected tweets that mentioned the terms "[he/she/they] turned off comments" or "turned off [her/his/their] comments." Importantly, the words did not have to appear in this exact order; the scraping process would extract both "she turned comments off" and "she turned off comments," enabling us to capture as many

¹ As of May 2023, Twitter's application programming interface no longer provides free access for academic researchers.

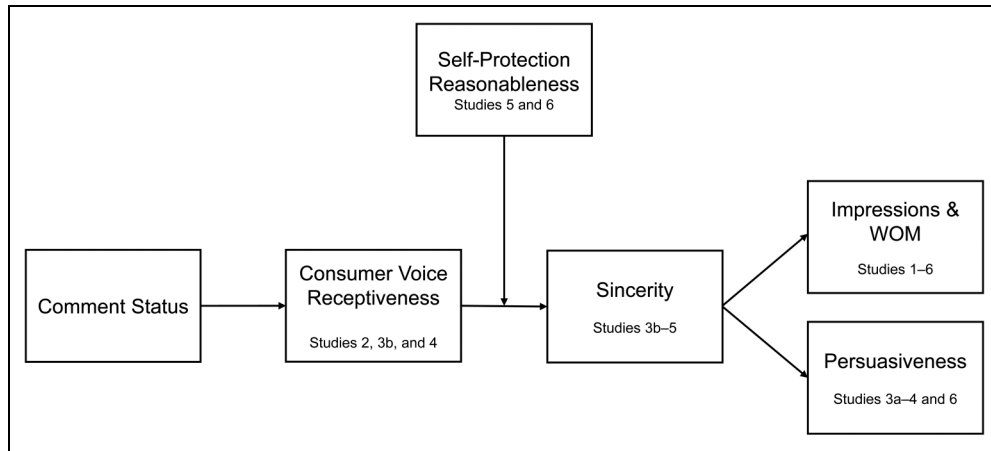


Figure 1. Conceptual Framework and Study Overview.

tweets as possible containing these key terms. For the control group, we collected a sample of tweets that were posted in the same time frame and contained at least one of the following pronouns: he, she, they, his, her, or their. We only collected control tweets containing pronouns so that both sets of tweets would reference a third party, increasing comparability with the treatment group. Finally, we removed any duplicate tweets (i.e., separate tweets that contained identical text) and only retained the first tweet from any given account, so that no single user would be overrepresented in our sample. Our final dataset included 1,222 tweets in the treatment group and 1,582 tweets in the control group, for a total of 2,804 unique tweets (dummy coded, treatment = 1, control = 0).

As our focal dependent variable, we analyzed tweet sentiment using TextAnalyzer (Berger et al. 2020). We were primarily interested in the relative proportion of negative sentiment expressed in tweets that referenced turning off comments versus tweets in the control group, which serves as an indicator of negative WOM. To this end, we focused our TextAnalyzer analysis on the negative sentiment lexicon developed by Kiritchenko, Zhu, and Mohammad (2014), which detects the presence of negative sentiment in short, informal text-based messages (e.g., tweets). In addition to analyzing tweet sentiment, following Valsesia, Proserpio, and Nunes (2020), we collected several control variables. At the tweet level, we recorded tweet length, the number of likes and retweets received, the number of other Twitter users mentioned, the number of hashtags used, and whether the tweet contained a URL, image, or video. At the user level, we recorded the age of the account at the time of the tweet, the length of the user's bio, the user's number of followers, the number of accounts the user follows, the number of tweets the user had posted, the number of tweets the user had liked, whether the account was verified, whether the user's bio contained a URL, and whether the profile used the default Twitter image.

Results

An initial examination indicated that none of the tweets in our dataset included a video, so we did not include this control variable

in our model. We retained all the other control variables mentioned previously. As predicted, an ordinary least squares regression analysis revealed that tweets in the treatment group contained more negative sentiment than those in the control group ($\beta = -.05$, $t(2,786) = -8.75$, $p < .001$), an effect that also held without the control variables. A follow-up examination of the means confirmed that tweets from the treatment group contained more negative sentiment than those from the control group ($M_{\text{treatment}} = -.47$, $SD = .13$ vs. $M_{\text{control}} = -.41$, $SD = .16$; range: -1.92 to $-.05$). A table summarizing effect sizes, significance levels, and descriptive statistics for the control variables is available in Web Appendix B.

Discussion

Using real reactions scraped from Twitter, Study 1 documents the negative WOM that is generated about people who turn off their social media comments. This initial finding suggests that by attempting to limit voice expression through disabling comments, such individuals may instead be inviting even more criticism. This notion is consistent with work showing that censorship attempts often backfire and further increase criticism of the target (Jansen and Martin 2015). Importantly, although the use of Twitter data enabled us to capture overall sentiment toward disabling comments in a naturalistic and relevant setting, we acknowledge several limitations of this quasi-experiment. First, because the scraping process utilized specific key terms, we have incomplete information regarding the subject of each tweet; thus it is possible that at least some of these tweets did not reference online influencers, the focus of our investigation. As such, we cannot disentangle consumer reactions to influencers from other social media users. Further, as with most secondary data, there may be omitted variables that could bias our results, such as the impetus for why the referenced individuals disabled their comments in the first place. We address these limitations in the following experiments.

Study 2

The goal of Study 2 is to test H_1 by examining, using a free-writing task, people's spontaneous reactions to an online

influencer who either allows or does not allow other users to comment on her social media posts. To bolster these open-ended responses, we also gauged overall impressions of the influencer via self-reported survey measures.

Method

We recruited Amazon Mechanical Turk (MTurk) panelists via CloudResearch to first complete a prescreening measure to ensure that they were actual users of Instagram, the focal context in Study 2. Those who failed this measure were precluded from participating. We utilized a similar screening procedure across all studies.

A total of 200 MTurk panelists passed this screener and completed a two-cell (comment status: disabled vs. enabled) between-subjects experiment for payment.² Two independent coders, unaware of the study hypothesis, identified 13 gibberish or irrelevant responses to the free-writing task, leaving a final sample of 187 participants ($M_{\text{age}} = 38.8$ years; 51.9% male, 47.1% female, .5% nonbinary; 77.0% White, 8.6% Black; 1 did not report gender). All responses deemed gibberish are reported in Web Appendix C.

Participants first viewed an Instagram post from Aileen—an online influencer who was promoting Reel, a brand of eco-friendly paper towels—who was described as having her comments either disabled or enabled. Both posts were identical aside from the comment icon, which was absent in the comments disabled condition, consistent with Instagram's layout. Next, participants described their thoughts and feelings about the influencer in a free-writing task using at least 50 words. After completing this writing task, participants responded to three items capturing their overall impressions of Aileen (1 = "Very unfavorable"/"Very unlikable"/"Very negative," and 9 = "Very favorable"/"Very likable"/"Very positive"; $\alpha = .99$). Last, participants completed manipulation checks and demographic measures. Manipulation checks confirmed that the manipulation was effective. For brevity, in this and the following studies, we include all manipulation check measures and results in Web Appendix D (along with study stimuli in Web Appendix E and study questionnaire links in Web Appendix F).

To complement self-reported influencer impressions, we also assessed impressions by quantifying participants' written responses in two other ways. First, as in Study 1, we analyzed response sentiment using TextAnalyzer (Berger et al. 2020; Kiritchenko, Zhu, and Mohammad 2014). Second, the same independent coders rated the overall valence of each response (0 = Did not mention Aileen, 1 = Extremely negative, 10 = Extremely positive, $r = .85$, $p < .001$). Initially, coders determined whether each response referenced Aileen's receptiveness to consumer voice (0 = Not receptive to consumer voice, 1 = Receptive to consumer voice, 2 = Did not mention). However, because they did not identify any responses that described

Aileen as receptive to consumer voice, responses were recoded into a dichotomous dependent variable (0 = Did not mention, 1 = Not receptive to consumer voice). Moreover, to examine whether additional judgments of Aileen spontaneously emerged upon learning that she disabled comments, coders rated whether responses referenced Aileen's potential transgressions (e.g., did something wrong or said something offensive) or the belief that she was trying to protect her mental health (both coded as 0 = Did not mention, 1 = Mentioned). Disagreements were resolved via discussion prior to analysis. Finally, coders provided several nonfocal ratings (e.g., Aileen's perceived egotism; see Web Appendix G).

Results

Impressions. One-way ANOVA analyses on all three measures of impressions revealed significant effects of comment status. The sentiment analysis ($M_{\text{disabled}} = -.38$, $SD = .08$ vs. $M_{\text{enabled}} = -.34$, $SD = .07$; range: $-.62$ to $-.17$; $F(1, 185) = 10.55$, $p = .001$, $\eta_p^2 = .05$), independent coder ratings ($M_{\text{disabled}} = 4.83$, $SD = 2.58$ vs. $M_{\text{enabled}} = 5.90$, $SD = 2.85$; $F(1, 176) = 6.84$, $p = .010$, $\eta_p^2 = .04$), and self-reported measures ($M_{\text{disabled}} = 5.07$, $SD = 2.45$ vs. $M_{\text{enabled}} = 6.23$, $SD = 2.44$; $F(1, 185) = 10.65$, $p = .001$, $\eta_p^2 = .05$) all followed the same pattern: participants formed more negative impressions of Aileen when she disabled (vs. enabled) comments.³

Consumer voice receptiveness. An initial logistic regression analysis indicated separation of the data, such that there were no mentions of a lack of voice receptiveness in the comments enabled condition. To address this concern, we applied a Firth correction (Firth 1993). As expected, more participants referenced Aileen's lack of receptiveness to consumer voice when she disabled (vs. enabled) comments (50.55% vs. 0%, $\chi^2(1) = 13.46$, $p < .001$).

Beliefs of mental health protection. As with consumer voice receptiveness, we applied a Firth correction because an initial logistic regression indicated data separation. Results indicated that marginally more participants referenced Aileen's attempts to protect her mental health when she disabled (vs. enabled) comments (7.69% vs. 0%, $\chi^2(1) = 3.71$, $p = .054$).

Perceived transgressions. A logistic regression revealed that references to Aileen's potential transgressions did not significantly differ regardless of whether she disabled (vs. enabled) comments (5.49% vs. 1.04%, $\chi^2(1) = 2.39$, $p = .122$). Importantly, we did not apply a Firth correction to this analysis because the model did not indicate data separation.

² Due to the labor-intensive nature of Study 2's coding task, we recruited a smaller, yet still well-powered, sample.

³ Coders identified nine participants who did not mention Aileen in their written responses and were thus removed from the independent coder ratings analysis, which is why the degrees of freedom for this analysis differed.

Discussion

Study 2 builds on Study 1 by examining consumers' open-ended reactions to influencers who turn off social media comments, while doing so in a controlled experimental setting that offers enhanced internal validity. Across sentiment analysis, independent coder ratings, and self-reported measures, we found convergent support for the prediction that consumers form more negative impressions of influencers who disable their comments (H_1). Further, we provide preliminary evidence that influencers who disable comments are spontaneously judged as lacking receptiveness to consumer voice and trying to protect their mental health, although the former inference occurred more readily. Moreover, beliefs that the influencer may have committed a transgression did not differ significantly between conditions, suggesting that disabling comments can cause negative reactions even in the absence of perceived wrongdoing. In Studies 3a and 3b, we assess, in a consequential manner, the negative effect of disabling social media comments on two major sources of income for influencers: consumer interest in their sponsored content and promotional discount codes.

Study 3a

The primary goal of Study 3a is to further test H_1 by investigating the impact of disabling comments on influencer persuasiveness in a behaviorally consequential fashion. Influencers typically generate a substantial portion of their income through affiliate marketing, earning a commission when consumers engage with their sponsored content through personalized affiliate links or discount codes (Christison 2022). In Study 3a, we examine whether disabling comments can shape consumer interest in engaging with the influencer's sponsored content, thereby impacting both the influencer and the endorsed brand.

Method

A total of 700 Connect by CloudResearch panelists ($M_{\text{age}} = 38.1$ years; 48.3% male, 50.1% female, 1.0% nonbinary; 71.9% White, 10.1% Black; 4 did not report gender) completed a two-cell (comment status: disabled vs. enabled) between-subjects study for payment.

Participants first read some background information about an online influencer named Jennifer, including her age and the focus of her posts, before viewing her Instagram profile, which included basic account information (e.g., follower and following count) and her nine latest posts. Participants were then shown Jennifer's most recent post, which depicted her enjoying springtime. They imagined clicking on this post's comment icon, which brought them to a separate screen that displayed either (1) a message informing users that comments on the post have been limited or (2) a text box enabling users to leave a comment. Importantly, the presented stimuli are identical to what Instagram users actually see when they click the comment icon, depending on their respective conditions.

Next, participants viewed Jennifer's second most recent post, which depicted her promoting the same brand of ecofriendly paper towels as Study 2. Participants imagined clicking on the post's comment icon, which, depending on their assigned conditions, again brought them to a screen that featured either a message informing them about the limiting of comments or a text box for users to leave a comment.

Next, participants rated their overall impressions of Jennifer using the same items as Study 2 ($\alpha = .97$). To assess consumer interest in the influencer's sponsored content in a behaviorally consequential manner, participants were then asked whether they wanted to learn more about Reel Paper, the brand endorsed by Jennifer, or complete an unrelated activity (adapted from Valsesia, Proserpio, and Nunes 2020). Participants were informed that Jennifer would earn a small commission if they decided to learn more about Reel Paper. Those who selected this option were presented with additional information regarding Reel Paper, whereas those who selected the other option were presented with information about a different brand. Finally, participants completed manipulation checks and demographic measures.

Results

Impressions. A one-way ANOVA on impressions revealed that participants formed more negative impressions of Jennifer when she disabled comments ($M_{\text{disabled}} = 5.40$, $SD = 2.09$ vs. $M_{\text{enabled}} = 6.08$, $SD = 2.01$; $F(1, 698) = 18.95$, $p < .001$, $\eta_p^2 = .03$).

Persuasiveness. A logistic regression on participants' activity selection revealed that fewer participants decided to learn more about the influencer's endorsed brand when she disabled (vs. enabled) comments (39.43% vs. 47.71%; $\chi^2(1) = 4.87$, $p = .027$, $\phi = .08$).

Discussion

Study 3a builds on and extends the previous two studies by showing that consumers not only form more negative impressions of influencers who disable comments but also find them less persuasive, supporting H_1 . Notably, given that the influencer would have earned a commission had participants decided to learn more about the brand, by choosing not to engage with her sponsored content, they were actively withholding their support for both the influencer and the endorsed brand. In Study 3b, we examine another consequential dependent measure while testing the full conceptual framework.

Study 3b

The goal of Study 3b (preregistered; <https://aspredicted.org/v43sf.pdf>) is to conceptually replicate Study 3a with another outcome variable capturing consumers' willingness to support the influencer and provide evidence for our proposed conceptual

model by assessing perceptions of consumer voice receptiveness and sincerity (H_2).

Method

A total of 705 Connect by CloudResearch panelists ($M_{\text{age}} = 38.3$ years; 49.0% male, 49.7% female, .6% nonbinary; 72.9% White, 9.4% Black; 4 did not report age, 8 did not report gender, 4 did not report race) completed a two-cell (comment status: disabled vs. enabled) between-subjects study for payment.

Study 3b followed a similar method as Study 3a. Participants read some background information about an online influencer named Caleb, viewed his Instagram profile, and were shown his two most recent posts, where he either disabled or enabled comments. Caleb's second post promoted novelty socks from a company called Sockologie. In this post, Caleb encouraged readers to utilize his discount code to order a pair of Sockologie socks. This discount code was linked in his profile biography, where influencers often store affiliate links.

Participants first rated their overall impressions of Caleb using the same items as previous studies ($\alpha = .97$). Next, participants learned that to thank them for their participation they would receive a 15% discount code that could be used at Sockologie. Participants could choose between Caleb's discount code, in which case he would receive a small commission, or a discount code for the same amount offered by Sockologie (1 = Caleb's discount code, 0 = Sockologie's discount code). After participants made their selection, we assessed perceptions of Caleb's receptiveness to consumer voice using four items based on the employee voice literature (e.g., "Caleb Abramson values other Instagram users' suggestions"; 1 = "Completely disagree," and 9 = "Completely agree"; $\alpha = .97$; Morrison 2014) and his perceived sincerity using three items based on prior work (e.g., "Caleb Abramson is an honest person"; 1 = "Completely disagree," and 9 = "Completely agree"; $\alpha = .97$; Campbell and Kirmani 2000; Tuk et al. 2009). Consistent with theorizing, a confirmatory factor analysis, which indicated a good fit with the data (GFI = .92; CFI = .98; SRMR = .02), supported treating impressions, consumer voice receptiveness, and sincerity perceptions as separate factors (Hu and Bentler 1999). The factor loadings and standardized regression coefficients for this and the following studies are available in Web Appendix H. Subsequently, we assessed the belief that Caleb is trying to protect his own mental health. For brevity, we report these measures and results in Web Appendix I. Finally, participants completed manipulation check and demographic measures. Upon study completion, participants received a 15% discount code redeemable at Sockologie.

Results

Impressions. A one-way ANOVA on impressions revealed that participants formed more negative impressions of Caleb when he disabled comments ($M_{\text{disabled}} = 5.13$, $SD = 2.02$ vs. $M_{\text{enabled}} = 5.72$, $SD = 1.98$; $F(1, 703) = 14.99$, $p < .001$, $\eta_p^2 = .02$).

Willingness to support influencer. A logistic regression on participants' discount code selection revealed that fewer participants

selected Caleb's discount code when he disabled (vs. enabled) comments (48.72% vs. 59.60%; $\chi^2(1) = 8.38$, $p = .004$, $\phi = .11$).

Consumer voice receptiveness. A one-way ANOVA on consumer voice receptiveness perceptions revealed that Caleb was perceived as less receptive to consumer voice when he disabled comments ($M_{\text{disabled}} = 3.47$, $SD = 2.05$ vs. $M_{\text{enabled}} = 5.34$, $SD = 1.97$; $F(1, 703) = 152.78$, $p < .001$, $\eta_p^2 = .18$).

Sincerity. The same analysis on sincerity perceptions revealed that Caleb was perceived as less sincere when he disabled comments ($M_{\text{disabled}} = 5.06$, $SD = 1.91$ vs. $M_{\text{enabled}} = 5.59$, $SD = 1.82$; $F(1, 703) = 14.13$, $p < .001$, $\eta_p^2 = .02$).

Mediation. We conducted two serial mediation analyses (Model 6, Hayes 2017) to test our proposed mediation path, whereby consumer voice receptiveness and sincerity perceptions sequentially drive the effects of comment status on impressions and willingness to support the influencer. Results revealed that the indirect effect of comment status on impressions through consumer voice receptiveness and sincerity was significant ($b = .81$, 95% CI: [.63, 1.01]). The pattern of mediation results for the other dependent measure, discount code selection, was also significant in the same direction ($b = .77$, 95% CI: [.57, 1.03]). Full details about individual mediation pathways for this and the following studies are available in Web Appendix J.

Discussion

Study 3b again shows that consumers form more negative impressions of and are less willing to support an influencer who disables social media comments, underscoring the interpersonal and professional repercussions of this behavior. Further, Study 3b provides direct support for H_2 by explicating the critical driving roles of consumer voice receptiveness and sincerity perceptions.

Up to this point, to ensure a parallel test between conditions, we have been comparing consumer responses to influencers who disable versus enable comments without displaying specific comments. However, this design leaves open the question of how the content of such comments might affect consumer reactions when they are visible on a post. As alluded to previously, people often express their voice on social media and generate comments that can be constructive, disparaging, or supportive in nature. Thus, in Study 4, we compare consumer responses to an influencer who disables comments versus enables comments that vary in their overall valence, and we focus on a context particularly conducive to generating user feedback.

Study 4

Study 4 (preregistered; <https://aspredicted.org/u6g7x.pdf>)⁴ has several primary objectives. First, we test the robustness of our

⁴ Perceived hypocrisy was originally treated as an alternative mechanism, but factor analysis revealed that it loaded with perceived sincerity. Including

theorizing by directly comparing the reputational costs of disabling comments to enabling negative, positive, and mixed-valence comments. Since consumers readily rely on others to inform their own judgments (Lamberton, Naylor, and Haws 2013) and weight negative information heavily (Ito et al. 1998), it is conceivable that keeping negative comments visible would be just as reputationally damaging as, if not more so than, preventing such comments from being shown. Despite this possibility, we contend that influencers are still worse off disabling comments (vs. leaving negative comments visible) due to the impaired voice receptiveness perceptions that this behavior engenders, which in turn undermines their perceived sincerity. For the same reason, we also expect more negative reactions to an influencer who disables comments than one who keeps positive or mixed-valence comments publicly visible. We examine these predictions in the context of an apology posted in response to a personal transgression, a situation particularly prone to evoking feedback from both supporters and detractors (e.g., Images 2023; Reslen 2023).

Relatedly, because previous studies did not offer plausible explanations for why comments were disabled in the first place, it is possible that the negative reactions to disabling comments could have simply been driven by the absence of a rational justification for the influencer's behavior. By utilizing an apology post following a transgression, Study 4 provides a salient rationale for the influencer's decision to disable comments (i.e., to avoid further criticism). However, considering that the desire to avoid criticism conveys lower perceptions of voice receptiveness, we expect this situation to continue to generate consumer backlash.

Additionally, we focus on a real-life celebrity with an active social media presence, Drew Barrymore. Today, online influencers include both digital content creators—the focus of previous studies—and celebrities like Drew Barrymore, who leverage their large social media followings to influence consumer decision-making. As such, based on our theorizing, we expect the negative impact of disabling comments to generalize to such celebrities. Moreover, by using a well-known celebrity like Drew Barrymore, we can assess the influence of engagement with that celebrity to determine if our proposed effects emerge across different levels of engagement. Finally, Study 4 examines whether the negative reactions to disabling comments could be explained by reactance, a negative motivational state that emerges when people experience threats to their freedom (Brehm 1966).

Method

We recruited a total of 798 MTurk panelists ($M_{\text{age}} = 40.7$ years; 50.1% male, 48.3% female, 1.0% nonbinary; 73.8% White, 12.2% Black; 2 did not report age, 6 did not report gender, 1 did not report race) via CloudResearch to complete a four-cell

(comment status: disabled vs. enabled-negative vs. enabled-positive vs. enabled-mixed) between-subjects study for payment.

Prior to completing the main study, participants indicated their overall level of engagement with Drew Barrymore by rating their agreement with the item: "I feel as if I have a close personal connection with the celebrity Drew Barrymore" (1 = "Strongly disagree," and 9 = "Strongly agree"; adapted from Sprott, Czellar, and Spangenberg [2009]).

Next, participants were shown an Instagram post by Drew Barrymore that was inspired by her response to the widespread criticism over her decision to resume her talk show amid the 2023 Hollywood writers' strike (Abad-Santos 2023). After receiving this criticism, Drew Barrymore issued a statement on Instagram expressing her deepest apologies to anyone her decision hurt. Then, participants were shown the comments section. While the disabled condition mirrored that of previous studies, participants in the enabled-negative condition encountered six negative user comments (e.g., "Shame on you! You're a terrible role model" and "I am so disappointed in you @drewbarrymore"), those in the enabled-positive condition encountered six positive user comments (e.g., "Way to go! You're an amazing role model" and "I have so much respect for you @drewbarrymore"), and those in the enabled-mixed condition encountered six user comments, three of which were negative (e.g., "I am so disappointed in you @drewbarrymore"), and three of which were positive (e.g., "Way to go! You're an amazing role model").

After viewing the post, participants rated their overall impressions of Drew Barrymore ($\alpha = .98$), her perceived receptiveness to consumer voice ($\alpha = .96$), and her perceived sincerity ($\alpha = .97$) using the same items as Study 3b. Additionally, we assessed her persuasiveness using three items (e.g., "How likely would you be to look at Drew Barrymore's recommendations on Instagram?"; 1 = "Not at all likely," and 9 = "Very likely"; $\alpha = .97$). To measure reactance as a potential alternative mechanism, participants completed three items assessing reactance (e.g., "To what extent does Drew Barrymore's post trigger a sense of reactance in you?"; adapted from Abendroth and Diehl [2006]; $\alpha = .86$). Last, we collected exploratory, manipulation check, and demographic measures.

Results

Impressions. A one-way ANOVA on impressions revealed a significant effect of comment status ($F(3, 794) = 13.05, p < .001, \eta_p^2 = .05$; see Figure 2, Panel A). Participants formed more negative impressions of Drew Barrymore when she disabled comments than when she enabled comments and they were negative ($M_{\text{disabled}} = 5.04, SD = 2.10$ vs. $M_{\text{enabled, negative}} = 5.69, SD = 2.12$; $F(1, 794) = 10.01, p = .002, \eta_p^2 = .01$), positive ($M_{\text{enabled, positive}} = 6.33, SD = 2.04$; $F(1, 794) = 38.95, p < .001, \eta_p^2 = .05$), or mixed ($M_{\text{enabled, mixed}} = 5.76, SD = 2.07$; $F(1, 794) = 12.10, p < .001, \eta_p^2 = .02$). Compared with when Drew Barrymore enabled comments that were positive, participants formed more negative impressions of her when she enabled comments that were negative ($F(1, 794) = 9.44, p = .002, \eta_p^2 =$

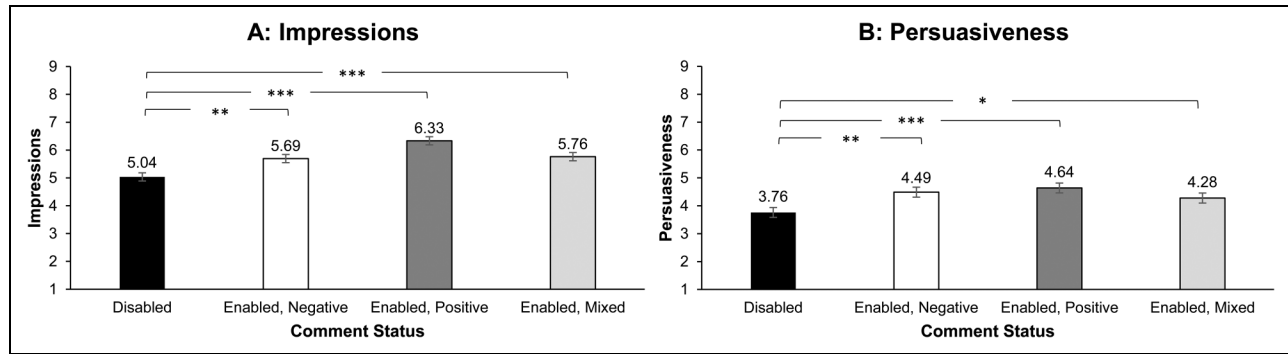


Figure 2. Impressions and Persuasiveness as a Function of Comment Status.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Notes: Error bars = ± 1 SE.

.01) or mixed ($F(1, 794) = 7.48, p = .007, \eta_p^2 = .01$). The negative and mixed comments conditions did not differ ($F < 1$).

Persuasiveness. The same analysis on persuasiveness also revealed a significant effect of comment status ($F(3, 794) = 4.65, p = .003, \eta_p^2 = .02$; see Figure 2, Panel B). Drew Barrymore was less persuasive when she disabled comments than when she enabled comments and they were negative ($M_{\text{disabled}} = 3.76, SD = 2.36$ vs. $M_{\text{enabled, negative}} = 4.49, SD = 2.54; F(1, 794) = 8.36, p = .004, \eta_p^2 = .01$), positive ($M_{\text{enabled, positive}} = 4.64, SD = 2.59; F(1, 794) = 12.19, p < .001, \eta_p^2 = .02$), or mixed ($M_{\text{enabled, mixed}} = 4.28, SD = 2.59; F(1, 794) = 4.19, p = .041, \eta_p^2 = .01$). None of the other conditions differed significantly ($p > .153$).

Consumer voice receptiveness. The same analysis on consumer voice receptiveness perceptions revealed a significant effect of comment status ($F(3, 794) = 33.36, p < .001, \eta_p^2 = .11$). Drew Barrymore was perceived as less receptive to consumer voice when she disabled comments than when she enabled comments and they were negative ($M_{\text{disabled}} = 3.97, SD = 2.16$ vs. $M_{\text{enabled, negative}} = 5.48, SD = 2.14; F(1, 794) = 50.54, p < .001, \eta_p^2 = .06$), positive ($M_{\text{enabled, positive}} = 6.01, SD = 2.01; F(1, 794) = 93.16, p < .001, \eta_p^2 = .11$), or mixed ($M_{\text{enabled, mixed}} = 5.17, SD = 2.18; F(1, 794) = 32.15, p < .001, \eta_p^2 = .04$). Compared with when Drew Barrymore enabled comments that were positive, she was perceived as less receptive to consumer voice when she enabled comments that were negative ($F(1, 794) = 6.42, p = .012, \eta_p^2 = .01$) or mixed ($F(1, 794) = 15.51, p < .001, \eta_p^2 = .02$). The negative and mixed comments conditions did not differ ($p > .158$).

Sincerity. The same analysis on sincerity perceptions revealed a significant effect of comment status ($F(3, 794) = 8.45, p < .001, \eta_p^2 = .03$). Drew Barrymore was perceived as less sincere when she disabled comments than when she enabled comments and they were negative ($M_{\text{disabled}} = 5.39, SD = 2.23$ vs. $M_{\text{enabled, negative}} = 6.10, SD = 2.07; F(1, 794) = 11.24, p < .001,$

$\eta_p^2 = .01$), positive ($M_{\text{enabled, positive}} = 6.43, SD = 2.07; F(1, 794) = 24.11, p < .001, \eta_p^2 = .03$), or mixed ($M_{\text{enabled, mixed}} = 6.05, SD = 2.09; F(1, 794) = 9.61, p = .002, \eta_p^2 = .01$). Drew Barrymore was perceived as marginally less sincere when she enabled comments that were mixed as opposed to positive ($F(1, 794) = 3.20, p = .074, \eta_p^2 = .004$). None of the other conditions differed significantly ($p > .121$).

Mediation. Based on our theorizing, the negative impressions and lowered persuasiveness of an influencer who disables social media comments should be driven sequentially by consumer voice receptiveness and sincerity perceptions, respectively. To examine this proposed mediation path, we conducted two multicategorical serial mediation analyses with 10,000 bootstrap samples (Model 6, Hayes 2017). In the first analysis, comment status was entered as a multicategorical independent variable (dummy coded with disabled comments as the comparison group), impressions was the dependent variable, and consumer voice receptiveness and sincerity perceptions were the respective mediators. In the second analysis, persuasiveness replaced impressions as the dependent variable. As expected, the indirect effects comparing disabled comments to enabled negative, positive, and mixed-valence comments were significant for both dependent variables (CIs contained 0). For more details on each pathway, see Table 1.

Alternative explanations. Because comment status impacted reactance ($F(3, 794) = 8.61, p < .001, \eta_p^2 = .03$), we compared our proposed process against this alternative mechanism in driving our focal dependent variables. To do so, we conducted a parallel multicategorical serial mediation analysis (Model 81, Hayes 2017) for each dependent variable, with sincerity and reactance serving as competing second mediators. As expected, the serial indirect effects through sincerity on both outcome variables remained significant ($b > .47$, all CIs exclude 0) and were substantially stronger than those through reactance. For brevity, we report all pathways in Web Appendix J.

Table 1. Study 4 Mediation Details.

Dependent Variable	Comparison (Comments Disabled vs.)	Coefficients			Serial Indirect Effect
		Comment Status to Consumer Voice Receptiveness	Consumer Voice Receptiveness to Sincerity	Sincerity to Impressions/ Persuasiveness	
Impressions	Enabled, negative	1.51***	.66***	.76***	b = .76, 95% CI: [.53, 1.00]
	Enabled, positive	2.05***			b = 1.03, 95% CI: [.80, 1.28]
	Enabled, mixed	1.21***			b = .61, 95% CI: [.38, .85]
Persuasiveness	Enabled, negative	1.51***	.66***	.56***	b = .56, 95% CI: [.38, .75]
	Enabled, positive	2.05***			b = .76, 95% CI: [.57, .96]
	Enabled, mixed	1.21***			b = .45, 95% CI: [.27, .64]

As of May 2023, Twitter's application programming interface no longer provides free access for academic researchers.

Engagement. We also examined the role of engagement in shaping our hypothesized effects. Overall, the level of engagement with Drew Barrymore fell below the scale midpoint of 5 ($M = 2.87$, $SD = 2.35$; $t(798) = -25.65$, $p < .001$). Importantly, the direct effects of comment status on our dependent variables remain significant when controlling for participants' level of engagement with Drew Barrymore (all $p \leq .001$) and do not differ depending on how engaged participants were (all two-way interactions: $p > .24$).

Discussion

Study 4 further bolsters support for H_1 and H_2 in several critical ways. First, we show that it can be more reputationally costly to disable comments than to enable them, regardless of whether these comments are negative, positive, or mixed. Interestingly, these effects emerged despite the possibility for negative comments to further taint judgments of the individual following a transgression, attesting to the robustness of this phenomenon. However, we note that the observed pattern of results is likely contingent on various contextual factors, a point we revisit in the "General Discussion" section.

Second, we find that the negative impact of disabling comments generalizes to a real-life celebrity with an active social media presence and does not appear to hinge on the consumer's level of engagement with that celebrity. Additionally, our proposed pathway via consumer voice receptiveness and sincerity remained the strongest even when we included reactance as a parallel mediator. We offer additional support for the strength of sincerity relative to reactance in a supplementary study ($N = 180$; see Web Appendix L), although we acknowledge that this phenomenon is multiply determined.

Notably, the backlash against disabling comments held even when there was a rational justification for this behavior (i.e., to avoid criticism), ruling out the possibility that these effects were driven by the absence of a plausible explanation for disabling comments. In our final two studies, we turn our attention to situations where consumers believe that it would be reasonable for influencers to prioritize self-protection, and we examine how such contexts can alter the extent to which consumers rely on voice receptiveness inferences to inform the influencer's perceived sincerity (H_3).

Study 5

In Study 5 (preregistered; <https://aspredicted.org/vh6v3.pdf>), we test H_3 by investigating whether the negative judgments of sincerity from disabling comments are attenuated when consumers believe that it is reasonable for an influencer to prioritize self-protection, such as during a difficult personal loss. As a result, consumers should be more likely to perceive an influencer's lack of voice receptiveness as reasonable for the given context (i.e., the grief-inducing situation) rather than indicative of their internal disposition (i.e., the influencer's sincerity), a possibility we test in Study 5. Further, we broaden the generalizability of our investigation by examining reactions to a post on Twitter, which also enables users to turn off comments or "replies" to their tweets.

Method

We recruited a total of 801 MTurk panelists ($M_{\text{age}} = 40.4$ years; 56.2% male, 42.4% female, .8% nonbinary; 74.7% White, 10.6% Black; 7 did not report age or race, 12 did not report gender) via CloudResearch to complete a 2 (comment status: disabled vs. enabled) \times 2 (self-protection reasonableness: reasonable vs. control) between-subjects study for payment.

Participants first read some background information about an online influencer named Lindsey before viewing one of her recent tweets, for which she either disabled or enabled replies. To manipulate the reasonableness of self-protection, Lindsey's tweet announced that either she had recently gotten a new puppy and expressed how much excitement she and her family were feeling (control) or her 16-year-old dog had recently passed away and expressed how much she and her family miss their dog (reasonable). Notably, a separate posttest confirmed that it was more reasonable for Lindsey to prioritize self-protection in the reasonable (vs. control) condition ($M_{\text{reasonable}} = 7.38$, $SD = 1.73$ vs. $M_{\text{control}} = 5.81$, $SD = 1.83$; $F(1, 198) = 38.99$, $p < .001$, $\eta_p^2 = .16$). For details, see Web Appendix D.

After viewing Lindsey's tweet, participants rated their overall impressions of her ($\alpha = .97$) and her perceived sincerity ($\alpha = .95$) using the same items as prior studies. Last, they completed manipulation checks and demographic measures.

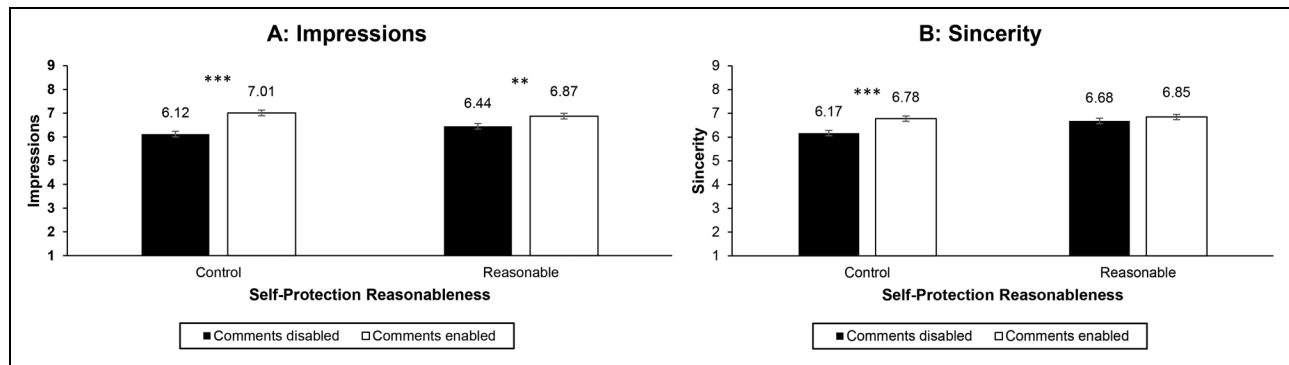


Figure 3. Impressions and Sincerity Perceptions as a Function of Comment Status and Self-Protection Reasonableness.

** $p < .01$.

*** $p < .001$.

Notes: Error bars = ± 1 SE.

Results

Impressions. A 2 (comment status) \times 2 (self-protection reasonableness) ANOVA on impressions revealed a main effect of comment status, such that participants formed more negative impressions of Lindsey when she disabled (vs. enabled) comments ($M_{\text{disabled}} = 6.28$, $SD = 1.74$ vs. $M_{\text{enabled}} = 6.94$, $SD = 1.59$; $F(1, 797) = 31.60$, $p < .001$, $\eta_p^2 = .04$). Importantly, results also revealed a two-way interaction ($F(1, 797) = 3.87$, $p < .050$, $\eta_p^2 = .005$; see Figure 3, Panel A). In the control condition, participants formed more negative impressions of Lindsey when she disabled (vs. enabled) comments, replicating prior studies ($M_{\text{disabled}} = 6.12$, $SD = 1.78$ vs. $M_{\text{enabled}} = 7.01$, $SD = 1.73$; $F(1, 797) = 28.83$, $p < .001$, $\eta_p^2 = .03$). However, this effect, while still significant, was attenuated when Lindsey was grieving a personal loss, thereby making it more reasonable for her to prioritize self-protection ($M_{\text{disabled}} = 6.44$, $SD = 1.70$ vs. $M_{\text{enabled}} = 6.87$, $SD = 1.44$; $F(1, 797) = 6.67$, $p = .010$, $\eta_p^2 = .01$).

Sincerity. The same analysis on sincerity perceptions revealed a main effect of comment status, such that Lindsey was perceived as less sincere when she disabled comments ($M_{\text{disabled}} = 6.42$, $SD = 1.69$ vs. $M_{\text{enabled}} = 6.81$, $SD = 1.51$; $F(1, 797) = 11.78$, $p < .001$, $\eta_p^2 = .01$), and a main effect of self-protection reasonableness, such that she was perceived as less sincere in the control condition ($M_{\text{control}} = 6.47$, $SD = 1.70$ vs. $M_{\text{reasonable}} = 6.76$, $SD = 1.51$; $F(1, 797) = 6.66$, $p = .010$, $\eta_p^2 = .01$). Importantly, results also revealed a two-way interaction ($F(1, 797) = 3.85$, $p = .050$, $\eta_p^2 = .005$; see Figure 3, Panel B). Replicating previous studies, Lindsey was perceived as less sincere when she disabled (vs. enabled) comments in the control condition ($M_{\text{disabled}} = 6.17$, $SD = 1.74$ vs. $M_{\text{enabled}} = 6.78$, $SD = 1.61$; $F(1, 797) = 14.57$, $p < .001$, $\eta_p^2 = .02$). However, Lindsey's perceived sincerity did not differ regardless of comment status when self-protection was reasonable ($M_{\text{disabled}} = 6.68$, $SD = 1.60$ vs. $M_{\text{enabled}} = 6.85$, $SD = 1.41$; $F(1, 797) = 1.08$, $p = .299$).

Mediation. We conducted a moderated mediation analysis (Model 8, Hayes 2017) to test our proposed pathway, whereby the relationship between comment status and impressions is mediated by sincerity perceptions in the control condition but attenuated in the reasonable condition. As predicted, the indirect effect of comment status on impressions through sincerity was significant in the control condition ($b = .49$, 95% CI: [.22, .76]), but it was not significant when self-protection was reasonable ($b = .13$, 95% CI: [−.10, .38]). Importantly, the index of moderated mediation was significant ($b = .36$, 95% CI: [.01, .71]).

Discussion

In Study 5, we demonstrate that the negative impact of disabling comments is mitigated when consumers believe that it is reasonable for influencers to prioritize self-protection, such as when they are experiencing grief over a personal loss (H_3). In our final study, we devise a practical solution to mitigate this phenomenon by focusing on another situational factor that makes it more acceptable for influencers to prioritize self-protection and has become ubiquitous on social media—mental health challenges.

Study 6

Given the extreme negativity directed toward public figures on social media (Ward 2020), the objective of Study 6 (preregistered; <https://aspredicted.org/5up4d.pdf>) is to examine how influencers could communicate their desire to prioritize self-protection when the need arises, thereby reducing the backlash against disabling comments. As revealed in the introduction, despite recognizing that social media users, including influencers, can improve their mental health by disabling comments, consumers nonetheless penalize them for doing so. Notably, some participants in Study 2 spontaneously generated mental health protection as a possible reason for disabling comments.

This finding suggests that if influencers can make mental health concerns more salient by publicly disclosing social media's toll on their mental health, perhaps consumers will recognize that the influencer's need for self-protection is reasonable and therefore be less likely to penalize them for disabling social media comments.

Method

We recruited a total of 794 MTurk panelists ($M_{\text{age}} = 40.2$ years; 50.4% male, 48.8% female, .4% nonbinary; 72.7% White, 11.4% Black; 4 did not report age, 6 did not report gender, 3 did not report race) via CloudResearch to complete a 2 (comment status: disabled vs. enabled) \times 2 (self-protection reasonableness: reasonable vs. control) between-subjects study for payment.

Study 6's procedure closely followed those of previous studies. Participants first read some background information about an online influencer named Rachel before being presented with her latest Instagram post, where she either disabled or enabled comments. To manipulate the reasonableness of self-protection, Rachel's caption either described how much she loves sharing her life with her followers (control) or in addition to this description, disclosed that the negativity on social media can become overwhelming, so she has decided to take a little break for her mental health (reasonable). Notably, a separate posttest confirmed that it was more reasonable for Rachel to prioritize self-protection in the reasonable (vs. control) condition ($M_{\text{reasonable}} = 8.07$, $SD = 1.26$ vs. $M_{\text{control}} = 7.28$, $SD = 1.48$; $F(1, 198) = 16.43$, $p < .001$, $\eta_p^2 = .08$). For details, see Web Appendix D.

Next, participants rated their overall impressions of Rachel using the same items as previous studies ($\alpha = .97$) and her persuasiveness using the same items as Study 4 ($\alpha = .97$). Last, participants completed manipulation checks and demographic measures.

Results

Impressions. A 2 (comment status) \times 2 (self-protection reasonableness) ANOVA on impressions revealed a main effect of comment status, such that participants formed more negative impressions of Rachel when she disabled comments ($M_{\text{disabled}} = 6.14$, $SD = 1.97$ vs. $M_{\text{enabled}} = 6.55$, $SD = 1.88$; $F(1, 790) = 9.54$, $p = .002$, $\eta_p^2 = .01$), and a main effect of self-protection reasonableness, such that participants formed more negative impressions of Rachel in the control condition ($M_{\text{control}} = 6.08$, $SD = 2.01$ vs. $M_{\text{reasonable}} = 6.61$, $SD = 1.82$; $F(1, 790) = 15.44$, $p < .001$, $\eta_p^2 = .02$). Importantly, these effects were qualified by a two-way interaction ($F(1, 790) = 20.46$, $p < .001$, $\eta_p^2 = .03$; see Figure 4, Panel A). In the control condition, participants formed more negative impressions of Rachel when she disabled comments ($M_{\text{disabled}} = 5.57$, $SD = 1.99$ vs. $M_{\text{enabled}} = 6.59$, $SD = 1.91$; $F(1, 790) = 29.04$, $p < .001$, $\eta_p^2 = .04$). Yet when Rachel discussed social media's toll on her mental health, thereby making it more reasonable for her to prioritize self-protection, impressions of her remained equally high regardless

of comment status ($M_{\text{disabled}} = 6.70$, $SD = 1.78$ vs. $M_{\text{enabled}} = 6.51$, $SD = 1.86$, $F(1, 790) = 1.03$, $p = .311$).

Persuasiveness. The same analysis on persuasiveness revealed main effects of comment status ($M_{\text{disabled}} = 5.05$, $SD = 2.37$ vs. $M_{\text{enabled}} = 5.43$, $SD = 2.43$; $F(1, 790) = 5.11$, $p = .024$, $\eta_p^2 = .01$) and self-protection reasonableness ($M_{\text{control}} = 4.97$, $SD = 2.47$ vs. $M_{\text{reasonable}} = 5.50$, $SD = 2.32$; $F(1, 790) = 9.78$, $p = .002$, $\eta_p^2 = .01$), which were qualified by a two-way interaction ($F(1, 790) = 4.33$, $p = .038$, $\eta_p^2 = .01$; see Figure 4, Panel B). In the control condition, Rachel was less persuasive when she disabled (vs. enabled) comments ($M_{\text{disabled}} = 4.61$, $SD = 2.40$ vs. $M_{\text{enabled}} = 5.34$, $SD = 2.48$; $F(1, 790) = 9.44$, $p = .002$, $\eta_p^2 = .01$), but when self-protection was reasonable, persuasiveness remained equally high regardless of comment status ($M_{\text{disabled}} = 5.49$, $SD = 2.26$ vs. $M_{\text{enabled}} = 5.52$, $SD = 2.38$; $F < 1$).

Discussion

Study 6 shows that when influencers acknowledge social media's toll on their mental health, making it more reasonable for them to prioritize self-protection, the negative effects of disabling comments on impressions and persuasiveness are eliminated. These findings are particularly relevant for online influencers who must constantly navigate the ever-changing demands of their careers while attempting to protect their mental health. Although mental health struggles may not spontaneously arise as the primary explanation for an influencer's lack of voice receptiveness when they disable comments, Study 6 suggests that, so long as influencers are transparent and make such concerns salient, consumers appear to respect their need to establish some personal boundaries, at least in the short term.

General Discussion

The present research reveals, across seven studies leveraging both consequential and hypothetical dependent measures, the interpersonal and professional consequences that online influencers incur when they disable social media comments. We first uncovered an association between turning off comments and negative WOM using real-world Twitter data (Study 1) before showing that consumers spontaneously generate negative impressions of influencers who disable comments (Study 2). Next, we documented the negative downstream consequences of disabling comments for influencer persuasiveness (Study 3a), elucidated our full conceptual model by assessing consumer voice receptiveness and sincerity perceptions (Study 3b), and showed that our conceptual model is robust to comment valence (Study 4). Consistent with our theorizing, this phenomenon is attenuated in situations where consumers believe that it is reasonable for influencers to prioritize self-protection, including grief over a personal loss (Study 5) and mental health struggles (Study 6).

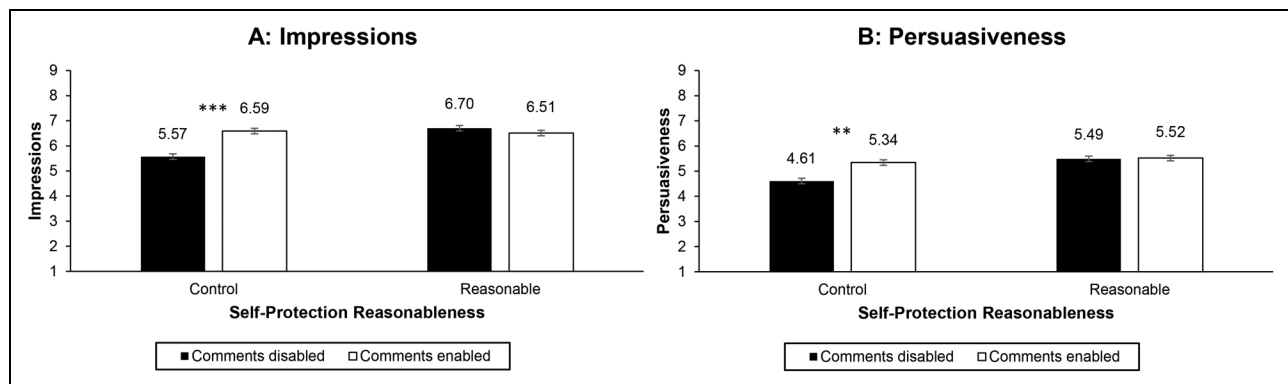


Figure 4. Impressions and Persuasiveness as a Function of Comment Status and Self-Protection Reasonableness.

** $p < .01$.

*** $p < .001$.

Notes: Error bars = ± 1 SE.

Theoretical Implications

The present research makes several important theoretical contributions. First, we contribute to the influencer marketing literature. Recent research has documented several predictors of audience engagement, including referencing close personal ties (Chung, Ding, and Kalra 2023), using sensory language (Cascio Rizzo et al. 2023), cultivating a large follower count (Leung et al. 2022; Wies, Bleier, and Edeling 2023), and not following others (Valsesia, Proserpio, and Nunes 2020). The present work adds to this body of work by identifying a prevalent yet previously unexplored online behavior, disabling social media comments, that can generate negative consumer reactions and reduce influencer effectiveness.

Second, we draw on employee voice from organizational theory (Maynes and Podsakoff 2014; Morrison 2014) to introduce to the marketing literature the consumer voice construct, defined as a consumer's informal and discretionary communication of thoughts, opinions, and suggestions to another marketplace entity. Because consumer voice expression is primarily achieved by leaving comments, consumers infer that influencers who disable comments are less receptive to their voice and thus less sincere, negatively impacting both impressions and persuasiveness. As such, disabling comments undermines a key influencer asset—their perceived receptiveness to consumer voice. Unlike earlier research on negative customer feedback (e.g., Singh 1990), consumer voice encompasses input ranging from supportive to constructive to destructive, suggesting that judgments of voice receptiveness hinge on how receptive influencers appear to be across these various types of consumer voice.

Relatedly, we enrich our understanding of how parasocial interactions can both bolster and undermine influencer engagement. Whereas prior work highlights the benefits of parasocial interactions for influencer effectiveness (e.g., Sokolova and Kefi 2020), we find that they can also promote unrealistic consumer expectations regarding appropriate levels of voice receptiveness (Grose 2023), inadvertently fostering an environment that penalizes influencers for disabling comments. In fact, a follow-up study

(preregistered; <https://aspredicted.org/x3wr6.pdf>) revealed even stronger backlash against disabling comments among those who are more likely to experience parasocial interactions (Dibble, Hartmann, and Rosaen 2016) and thus have higher expectations that influencers will be receptive to their voice (see Web Appendix M). As such, our work underscores the crucial role of consumer voice receptiveness in shaping influencer effectiveness.

More broadly, we add to a growing body of work on consumer inference-making in online environments (e.g., Appel et al. 2020; Grewal and Stephen 2019; Li, Chan, and Kim 2019). As social interactions become increasingly digitally mediated, people rely more heavily on online cues to guide their judgments and decisions, highlighting the importance of understanding perceptions stemming from one's online activities.

Practical Implications

Our research offers important practical insights for online influencers and their brand partners. Depression and anxiety are rising globally (Goodwin et al. 2020), yet stigma prevents many individuals from addressing their mental health (Corrigan 2000). It is thus critical to understand how the use of platform features designed to protect one's mental health could be perceived and to illuminate the potential downstream consequences of such actions. Although disabling comments may limit cyberbullying in the short term and improve influencers' immediate mental health, the results of Study 1 suggest that this strategy could breed animosity and generate additional disapproval. Critically, the goal of this work is not to discourage public figures from utilizing social media features that protect their mental health but rather to shed light on possible public reactions to their online activities, thereby enabling them to make more informed decisions for their personal and professional well-being. Additionally, by identifying an intervention grounded in mental health awareness to mitigate the backlash against disabling comments, we hope to empower influencers seeking to better manage online hostility without compromising their professional livelihood.

Moreover, our work emphasizes how seemingly innocuous online activities could have important professional ramifications for influencers' brand partnerships. Global spending on influencer marketing campaigns reached \$34.1 billion in 2023 and is projected to surpass \$47.8 billion by the end of 2027 (Statista 2023; West 2023), an expenditure level that has even spurred the rapid development of an industry centered around connecting influencers with brands (Geyser 2024b). Our studies collectively demonstrate that the decision to disable social media comments can reduce influencer persuasiveness, which emphasizes the importance of ensuring communication between brands and influencers to optimize their strategic partnerships.

Limitations and Future Research

This research presents some limitations as well as avenues for further investigation. First, because neither Study 5 nor Study 6 fully captured our conceptual model, we could not explicitly test whether, as theorized, self-protection reasonableness specifically moderates the relationship between consumer voice receptiveness and sincerity perceptions, our proposed mediators. Further, the nature of the caption utilized in Study 6 may have amplified the influencer's voice receptiveness perceptions relative to those used in other studies. Future work could examine other ways to communicate influencers' need for self-protection without highlighting their receptiveness to consumer voice as prominently.

Second, although our studies demonstrate the robustness of our proposed mechanisms and addressed alternative processes like reactance, the negative impact of disabling comments is likely multiply determined. Indeed, it would be interesting to examine whether this phenomenon broadly captures a negative halo effect that is triggered by the lack of consumer voice receptiveness. For example, by signaling lower voice receptiveness, could influencers be evaluated negatively along other dimensions beyond sincerity, including competence, diligence, and attractiveness? This question is substantively important, as it suggests that an influencer's sincerity is not the only attribute at stake when disabling comments. Rather, violating expectations of consumer voice receptiveness could create a negative lens that shapes interpersonal judgments of the influencer along other critical dimensions.

Third, it may be worth exploring the emotional and motivational antecedents that propel consumers to seek out other users' comments in the first place. When people experience emotions like envy and schadenfreude (Takahashi et al. 2009), they may be especially motivated to witness the public shaming and subsequent downfall of influencers and public figures, further exacerbating consumer reactions to disabling comments. Future research could thus investigate the various antecedents that cause people to peruse the comments section following a specific incident, as well as the emotional reactions that emerge upon discovering that comments have been turned off.

Notably, although the backlash against disabling comments is attenuated when consumers consider it reasonable for influencers

to prioritize self-protection, the unfortunate reality is that not all influencers will be afforded the same amount of grace. People with marginalized identities (e.g., BIPOC, LGBTQ+) are often targets of online harassment (Vogels 2021) and thus may be more inclined to disable comments. Yet marginalized groups regularly experience bias and discrimination in the marketplace (Bone, Christensen, and Williams 2014; Crockett 2017; Lamberton et al. 2024; Scott et al. 2024). As a result, consumers may hold marginalized influencers to a higher standard when deciding whether it would be reasonable for them to prioritize self-protection, possibly penalizing them even more heavily for engaging in the same activity. Future research should therefore examine how influencers' marginalized identities could impact consumer responses to disabling comments.

Additionally, future work can identify situations where disabling comments might bolster influencer effectiveness. For instance, if influencers reframe their decision to turn off comments as a measure designed to protect their followers from hateful messages or misinformation rather than as a sign of their dismissiveness, could the repercussions of disabling comments be lessened and perhaps enhance sincerity? Likewise, in certain contexts, might disabling comments signify an influencer's unequivocal support for the focal cause (e.g., social justice and political advocacy) by showing their unwavering dedication to the issue and unwillingness to entertain opposing viewpoints? In sum, it would be interesting to explore situations where this phenomenon reverses.

Moreover, although our findings indicate that it can be more reputationally costly to disable comments than to keep negative comments visible, the observed pattern of results is likely dependent on both the context and content of the displayed comments. For instance, future work could examine how consumers might respond to an influencer receiving extremely negative criticism. It is plausible that the negativity could become so corrosive that disabling comments would be the better option reputationally, yet consumers support brands when they are criticized unfairly (Allard, Dunn, and White 2020). However, if the negative comments reveal new information (e.g., they offer novel insights into the influencer's transgressions) or sway public opinion by providing social proof (Cialdini 2007), keeping comments visible could turn out to be just as damaging. As such, future research can further investigate the role of comment content in shaping consumer reactions.

Finally, it would be interesting to examine the potential consequences of the language used by platforms to communicate comment disabling. For example, Twitter specifies that comments have been limited to a select group of users, whereas Facebook specifies that comments have been turned off for everyone. Future research could identify more tactful strategies to communicate a user's decision to disable comments, possibly generating less aversive reactions. Additionally, it may be worth exploring other wellness features on these platforms, such as the ability to hide the number of likes received (Witte 2019). Concealing likes could limit unhealthy social comparisons without conveying a lack of receptiveness to consumer voice and may even signal that the influencer is less motivated by

extrinsic factors like popularity, thereby bolstering sincerity perceptions. Such intriguing questions await further investigation.

In conclusion, we find that for public figures like online influencers, the intersection of mental health and impression management is rife with difficult decisions. Although there is growing awareness of the potential harm caused by social media comments, influencers are often placed in an unenviable position where they must determine the level of negativity they can tolerate for their profession's sake. We hope that our research represents an important first step toward understanding the interpersonal and professional consequences that influencers incur for establishing boundaries with the public. Moreover, by pinpointing actionable strategies to reduce the tension between mental health protection and consumer engagement, we hope that our work aids those who are looking to make social media a safer and more productive marketing channel.

Acknowledgments

The authors would like to thank Kelly Goldsmith, Kelly Haws, Dawn Iacobucci, Andrea Morales, and Adriana Samper for their helpful feedback on previous versions of the manuscript. Further, the authors are grateful to the *JM* review team for their thoughtful guidance and support throughout the review process. Finally, the authors are grateful to the TIDE Lab at the Culverhouse College of Business and the Insights Lab at Vanderbilt.

Authors Contributions

The authors contributed equally to this work and are listed in alphabetical order.

Coeditor

Cait Lamberton

Associate Editor

Donna L. Hoffman


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Freeman Wu  <https://orcid.org/0000-0003-4379-3736>

References

- Abad-Santos, Alex (2023), "Drew Barrymore Tried to Live, Laugh, Scab Her Way Across the Picket Line. It Didn't Work," *Vox* (September 19), <https://www.vox.com/culture/2023/9/19/23880858/drew-barrymore-show-writers-strike-backlash-apology>.
- Abendroth, Lisa J. and Kristin Diehl (2006), "Now or Never: Effects of Limited Purchase Opportunities on Patterns of Regret Over Time," *Journal of Consumer Research*, 33 (3), 342–51.
- Abidin, Crystal (2015), "Communicative Intimacies: Influencers and Perceived Interconnectedness," *Ada: A Journal of Gender, New Media, and Technology*, 8 (November), <https://scholarsbank.uoregon.edu/xmlui/handle/1794/26365>.
- Allard, Thomas, Lea H. Dunn, and Katherine White (2020), "Negative Reviews, Positive Impact: Consumer Empathetic Responding to Unfair Word of Mouth," *Journal of Marketing*, 84 (4), 86–108.
- Appel, Gil, Luran Grewal, Rhonda Hadi, and Andrew T. Stephen (2020), "The Future of Social Media in Marketing," *Journal of the Academy of Marketing Science*, 48, 79–95.
- Ault, Susanne (2014), "Survey: YouTube Stars More Popular Than Mainstream Celebs Among U.S. Teens," *Variety* (August 5), <https://variety.com/2014/digital/news/survey-youtube-stars-more-popular-than-mainstream-celebs-among-u-s-teens-1201275245/>.
- Baker, Levi R. and James K. McNulty (2011), "Self-Compassion and Relationship Maintenance: The Moderating Roles of Conscientiousness and Gender," *Journal of Personality and Social Psychology*, 100 (5), 853–73.
- Barasch, Alixandra, Jonathan Z. Berman, and Deborah A. Small (2016), "When Payment Undermines the Pitch: On the Persuasiveness of Pure Motives in Fund-Raising," *Psychological Science*, 27 (10), 1388–97.
- Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel (2020), "Uniting the Tribes: Using Text for Marketing Insight," *Journal of Marketing*, 84 (1), 1–25.
- Bone, Sterling A., Glenn L. Christensen, and Jerome D. Williams (2014), "Rejected, Shackled, and Alone: The Impact of Systemic Restricted Choice on Minority Consumers' Construction of Self," *Journal of Consumer Research*, 41 (2), 451–74.
- Brehm, Jack W. (1966), *A Theory of Psychological Reactance*. Academic Press.
- Campbell, Margaret C. and Amna Kirmani (2000), "Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent," *Journal of Consumer Research*, 27 (1), 69–83.
- Cascio Rizzo, Giovanni Luca, Jonah Berger, Matteo De Angelis, and Rumen Pozharliev (2023), "How Sensory Language Shapes Influencer's Impact," *Journal of Consumer Research*, 50 (4), 810–25.
- Christison, Colleen (2022), "How Much Do Influencers Make in 2024?" *Hootsuite* (October 3), <https://blog.hootsuite.com/how-much-do-influencers-make/>.
- Chung, Jaeyeon, Yu Ding, and Ajay Kalra (2023), "I Really Know You: How Influencers Can Increase Audience Engagement by Referencing Their Close Social Ties," *Journal of Consumer Research*, 50 (4), 683–703.
- Cialdini, Robert B. (2007), *Influence: The Psychology of Persuasion*. HarperCollins.
- Corrigan, Patrick W. (2000), "Mental Health Stigma as Social Attribution: Implications for Research Methods and Attitude Change," *Clinical Psychology: Science and Practice*, 7 (1), 48–67.
- Crockett, David (2017), "Paths to Respectability: Consumption and Stigma Management in the Contemporary Black Middle Class," *Journal of Consumer Research*, 44 (3), 554–81.

- De Veirman, Marijke, Veroline Cauberghe, and Liselot Hudders (2017), "Marketing Through Instagram Influencers: The Impact of Number of Followers and Product Divergence on Brand Attitude," *International Journal of Advertising*, 36 (5), 798–828.
- Dibble, Jayson L., Tilo Hartmann, and Sarah F. Rosaen (2016), "Parasocial Interaction and Parasocial Relationship: Conceptual Clarification and a Critical Assessment of Measures," *Human Communication Research*, 42 (1), 21–44.
- Erickson, Rebecca J. (1995), "The Importance of Authenticity for Self and Society," *Symbolic Interaction*, 18 (2), 121–44.
- Fein, Steven, James L. Hilton, and Dale T. Miller (1990), "Suspicion of Ulterior Motivation and the Correspondence Bias," *Journal of Personality and Social Psychology*, 58 (5), 753–64.
- Ferchaud, Arianne, Jenna Grzeslo, Stephanie Orme, and Jared LaGroue (2018), "Parasocial Attributes and YouTube Personalities: Exploring Content Trends Across the Most Subscribed YouTube Channels," *Computers in Human Behavior*, 80 (March), 88–96.
- Filieri, Raffaele, Fulya Acikgoz, Chunyu Li, and Salma Alguezaui (2023), "Influencers' 'Organic' Persuasion Through Electronic Word of Mouth: A Case of Sincerity Over Brains and Beauty," *Psychology & Marketing*, 40 (2), 347–64.
- Firth, David (1993), "Bias Reduction of Maximum Likelihood Estimates," *Biometrika*, 80 (1), 27–38.
- Gallucci, Nicole (2019), "Taylor Swift Says Turning Off Instagram Comments Does Wonders for Self-Esteem," *Mashable* (March 6), <https://mashable.com/article/taylor-swift-elle-30-things-i-learned-turn-off-instagram-comments/>.
- Gerard, Tiegghan (@halfbakedharvest) (2023), "25 Minute Ginger Sesame Banh Mi Rice Bowls," Instagram video (March 6), <https://www.instagram.com/reel/CpePICLDvRt/>.
- Geyser, Werner (2024a), "20 of Instagram's Highest Paid Stars in 2024," *Influencer Marketing Hub* (January 30), <https://influencermarketinghub.com/instagram-highest-paid/>.
- Geyser, Werner (2024b), "The State of Influencer Marketing 2022: Benchmark Report," *Influencer Marketing Hub* (January 30), <https://influencermarketinghub.com/influencer-marketing-benchmark-report-2022/>.
- Gilbert, Daniel T. and Patrick S. Malone (1995), "The Correspondence Bias," *Psychological Bulletin*, 117 (1), 21–38.
- Gillath, Omri, Amanda K. Sesko, Phillip R. Shaver, and David S. Chun (2010), "Attachment, Authenticity, and Honesty: Dispositional and Experimentally Induced Security Can Reduce Self- and Other-Deception," *Journal of Personality and Social Psychology*, 98 (5), 841–55.
- Goodwin, Renee D., Andrea H. Weinberger, June H. Kim, Melody Wu, and Sandro Galea (2020), "Trends in Anxiety Among Adults in the United States, 2008–2018: Rapid Increases Among Young Adults," *Journal of Psychiatric Research*, 130 (November), 441–46.
- Grewal, Lauren and Andrew T. Stephen (2019), "In Mobile We Trust: The Effects of Mobile Versus Nonmobile Reviews on Consumer Purchase Intentions," *Journal of Marketing Research*, 56 (5), 791–808.
- Grose, Jessica (2023), "John Mulaney and the Insatiable Modern Audience," *New York Times* (May 3), <https://www.nytimes.com/2023/05/03/opinion/john-mulaney-maggie-smith.html/>.
- Grubbs, Joshua B. and Julie J. Exline (2016), "Trait Entitlement: A Cognitive-Personality Source of Vulnerability to Psychological Distress," *Psychological Bulletin*, 142 (11), 1204–26.
- Hassan, Syahida, Mohd Idzwan Jacob, Thuan Nguyen, and Suzana Zambri (2018), "Social Media Influencer and Cyberbullying: A Lesson Learned from Preliminary Findings," *Knowledge Management International Conference (KMICe) 2018 Proceedings*, 200–205, <https://soc.uum.edu.my/knice/proceedings/2018/pdf/CR54.pdf>.
- Hayes, Andrew F. (2017), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 2nd ed. Guilford Press.
- Horton, Donald and R. Richard Wohl (1956), "Mass Communication and Para-Social Interaction: Observations on Intimacy at a Distance," *Psychiatry*, 19 (3), 215–29.
- Hu, Li-tze and Peter M. Bentler (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling: A Multidisciplinary Journal*, 6 (1), 1–55.
- Images (2023), "Beauty Mogul Huda Kattan Apologises After Criticising Blogger Mina Hasan's Technique," (April 7), <https://images.dawn.com/news/1191680/>.
- Ito, Tiffany A., Jeff T. Larsen, N. Kyle Smith, and John T. Cacioppo (1998), "Negative Information Weighs More Heavily on the Brain: The Negativity Bias in Evaluative Categorizations," *Journal of Personality and Social Psychology*, 75 (4), 887–900.
- Jansen, Sue Curry and Brian Martin (2015), "The Streisand Effect and Censorship Backfire," *International Journal of Communication*, 9, 656–71.
- Jin, S. Venus and Ehri Ryu (2020), "'I'll Buy What She's #Wearing': The Roles of Envy Toward and Parasocial Interaction with Influencers in Instagram Celebrity-Based Brand Endorsement and Social Commerce," *Journal of Retailing and Consumer Services*, 55 (July), 102121.
- Jones, Edward E. and Richard E. Nisbett (1987), "The Actor and the Observer: Divergent Perceptions of the Causes of Behavior," in *Attribution: Perceiving the Causes of Behavior*. Lawrence Erlbaum Associates, 79–94.
- Just Jared (2023), "Fans Defend TikTok Beauty Guru Mikayla Nogueira Amid Claims She Fakes Her Accent," (May 11), <https://www.justjared.com/2023/05/11/fans-defend-tiktok-beauty-guru-mikayla-nogueira-amid-claims-she-fakes-her-accent/>.
- Kearney, Michael W. (2019), "rtweet: Collecting and Analyzing Twitter Data," *Journal of Open Source Software*, 4 (42), 1829.
- Kelley, Harold H. (1973), "The Processes of Causal Attribution," *American Psychologist*, 28 (2), 107–28.
- Kim, Dae-o (2020), "I Have Reported on 30 Korean Celebrity Suicides. The Blame Game Never Changes," *The Guardian* (January 3), <https://www.theguardian.com/music/2020/jan/04/i-have-reported-on-30-korean-celebrity-suicides-the-blame-game-never-changes/>.
- Kiritchenko, Svetlana, Xiaodan Zhu, and Saif M. Mohammad (2014), "Sentiment Analysis of Short Informal Texts," *Journal of Artificial Intelligence Research*, 50, 723–62.
- Kowert, Rachel and Emory Daniel Jr. (2021), "The One-and-a-Half Sided Parasocial Relationship: The Curious Case of Live Streaming," *Computers in Human Behavior Reports*, 4 (August–December), 100150.

- Labrecque, Lauren I. (2014), "Fostering Consumer-Brand Relationships in Social Media Environments: The Role of Parasocial Interaction," *Journal of Interactive Marketing*, 28 (2), 134-48.
- Lamberton, Cait, Rebecca Walker Naylor, and Kelly L. Haws (2013), "Same Destination, Different Paths: When and How Does Observing Others' Choices and Reasoning Alter Confidence in Our Own Choices?" *Journal of Consumer Psychology*, 23 (1), 74-89.
- Lamberton, Cait and Andrew T. Stephen (2016), "A Thematic Exploration of Digital, Social Media, and Mobile Marketing: Research Evolution from 2000 to 2015 and an Agenda for Future Inquiry," *Journal of Marketing*, 80 (6), 146-72.
- Lamberton, Cait, Tom Wein, Andrew Morningstar, and Sakshi Ghai (2024), "Marketing's Role in Promoting Dignity and Human Rights: A Conceptualization for Assessment and Future Research," *Journal of the Academy of Marketing Science* (published online March 1), <https://doi.org/10.1007/s11747-024-01008-x>.
- Leung, Fine F., Flora F. Gu, Yiwei Li, Jonathan Z. Zhang, and Robert W. Palmatier (2022), "Influencer Marketing Effectiveness," *Journal of Marketing*, 86 (6), 93-115.
- Leung, Fine F., Flora F. Gu, and Robert W. Palmatier (2022), "Online Influencer Marketing," *Journal of the Academy of Marketing Science*, 50 (2), 226-51.
- Li, Xueni (Shirley), Kimmy Wa Chan, and Sara Kim (2019), "Service with Emoticons: How Customers Interpret Employee Use of Emoticons in Online Service Encounters," *Journal of Consumer Research*, 45 (5), 973-87.
- Liden, Robert C. and Terence R. Mitchell (1988), "Ingratiation Behaviors in Organizational Settings," *Academy of Management Review*, 13 (4), 572-87.
- Lovett, Mitchell J., Renana Peres, and Ron Shachar (2013), "On Brands and Word of Mouth," *Journal of Marketing Research*, 50 (4), 427-44.
- Maynes, Timothy D. and Philip M. Podsakoff (2014), "Speaking More Broadly: An Examination of the Nature, Antecedents, and Consequences of an Expanded Set of Employee Voice Behaviors," *Journal of Applied Psychology*, 99 (1), 87-112.
- McCurry, Justin (2022), "South Korea Under Pressure to Crack Down on Cyberbullying After High-Profile Deaths," *The Guardian* (February 9), <https://www.theguardian.com/world/2022/feb/09/south-korea-under-pressure-to-crack-down-on-cyberbullying-after-high-profile-deaths/>.
- Morrison, Elizabeth W. (2014), "Employee Voice and Silence," *Annual Review of Organizational Psychology and Organizational Behavior*, 1 (1), 173-97.
- Murray, Sandra L., Jaye L. Derrick, Sadie Leder, and John G. Holmes (2008), "Balancing Connectedness and Self-Protection Goals in Close Relationships: A Levels-of-Processing Perspective on Risk Regulation," *Journal of Personality and Social Psychology*, 94 (3), 429-59.
- Murray, Sandra L., John G. Holmes, and Nancy L. Collins (2006), "Optimizing Assurance: The Risk Regulation System in Relationships," *Psychological Bulletin*, 132 (5), 641-66.
- Nguyen, Hoang (2023), "American Food Influencer Criticized for Misrepresenting Banh Mi," *VnExpress International* (March 15), <https://e.vnexpress.net/news/culture/american-food-influencer-criticized-for-misrepresenting-banh-mi-4581794.html/>.
- Nguyen, Nhung T., Anson Seer, and Nathan S. Hartman (2008), "Putting a Good Face on Impression Management: Team Citizenship and Team Satisfaction," *Journal of Behavioral and Applied Management*, 9 (2), 148-68.
- Ohanian, Roobina (1990), "Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness," *Journal of Advertising*, 19 (3), 39-52.
- Reinikainen, Hanna, Juha Munnukka, Devdeep Maity, and Vilma Luoma-aho (2020), "'You Really Are a Great Big Sister' - Parasocial Relationships, Credibility, and the Moderating Role of Audience Comments in Influencer Marketing," *Journal of Marketing Management*, 36 (3-4), 279-98.
- Reslen, Eileen (2023), "Jamie Foxx Apologizes for 'Antisemitic' Post: 'I Only Have Love in My Heart,'" *Page Six* (August 5), <https://pagesix.com/2023/08/05/jamie-foxx-apologizes-for-antisemitic-post-after-backlash/>.
- Schouten, Alexander P., Loes Janssen, and Maegan Verspaget (2020), "Celebrity vs. Influencer Endorsements in Advertising: The Role of Identification, Credibility, and Product-Endorser Fit," *International Journal of Advertising*, 39 (2), 258-81.
- Scott, Maura L., Sterling A. Bone, Glenn L. Christensen, Anneliese Lederer, Martin Mende, Brandon G. Christensen, et al (2024), "Revealing and Mitigating Racial Bias and Discrimination in Financial Services," *Journal of Marketing Research*, 61 (4), 598-618.
- Singh, Jagdip (1990), "Voice, Exit, and Negative Word-of-Mouth Behaviors: An Investigation Across Three Service Categories," *Journal of the Academy of Marketing Science*, 18, 1-15.
- Sokolova, Karina and Hajer Kefi (2020), "Instagram and YouTube Bloggers Promote It, Why Should I Buy? How Credibility and Parasocial Interaction Influence Purchase Intentions," *Journal of Retailing and Consumer Services*, 53 (March), 101742.
- Sprott, David, Sandor Czele, and Eric Spangenberg (2009), "The Importance of a General Measure of Brand Engagement on Market Behavior: Development and Validation of a Scale," *Journal of Marketing Research*, 46 (1), 92-104.
- Statista (2023), "Influencer Marketing Spending Worldwide and in the United States in 2022 and 2023," (September), <https://www.statista.com/statistics/1414663/influencer-marketing-spending-global-us/>.
- Stewart, Dodai (2022), "The Dazzling Life and Shocking Death of Cheslie Kryst," *New York Times* (September 27), <https://www.nytimes.com/2022/09/27/nyregion/cheslie-kryst-mental-health.html/>.
- Takahashi, Hidehiko, Motoichiro Kato, Masato Matsuura, Dean Mobbs, Tetsuya Suhara, and Yoshiro Okubo (2009), "When Your Gain Is My Pain and Your Pain Is My Gain: Neural Correlates of Envy and Schadenfreude," *Science*, 323 (5916), 937-39.
- Tolson, Andrew (2010), "A New Authenticity? Communicative Practices on YouTube," *Critical Discourse Studies*, 7 (4), 277-89.
- Trilling, Lionel (1973), *Sincerity and Authenticity*. Harvard University Press.
- Tuk, Mirjam A., Peeter W.J. Verlegh, Ale Smidts, and Daniel H.J. Wigboldus (2009), "Sales and Sincerity: The Role of Relational Framing in Word-of-Mouth Marketing," *Journal of Consumer Psychology*, 19 (1), 38-47.

- Valsesia, Francesca, Davide Proserpio, and Joseph C. Nunes (2020), "The Positive Effect of Not Following Others on Social Media," *Journal of Marketing Research*, 57 (6), 1152–68.
- Van Dyne, Linn and Jeffrey A. LePine (1998), "Helping and Voice Extra-Role Behaviors: Evidence of Construct and Predictive Validity," *Academy of Management Journal*, 41 (1), 108–19.
- Vogels, Emily A. (2021), "The State of Online Harassment," *Pew Research Center* (January 13), <https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/>.
- Ward, Fiona (2020), "Why Do People Really Post Mean Comments Online? The Experts Explain," *Hello Magazine* (February 19), <https://www.hellomagazine.com/healthandbeauty/health-and-fitness/2019012966949/why-people-post-mean-comments-on-social-media-psychologists-explain/>.
- West, Chloe (2023), "22 Influencer Marketing Statistics to Guide Your Brand's Strategy in 2024," *Sprout Social* (October 26), <https://sproutsocial.com/insights/influencer-marketing-statistics/>.
- Wies, Simone, Alexander Bleier, and Alexander Edeling (2023), "Finding Goldilocks Influencers: How Follower Count Drives Social Media Engagement," *Journal of Marketing*, 87 (3), 383–405.
- Witte, Rae (2019), "Should Instagram Get Rid of 'Likes'?" *Wall Street Journal* (July 25), <https://www.wsj.com/articles/should-instagram-get-rid-of-likes-11564071289/>.
- Xiong, Fei and Yun Liu (2014), "Opinion Formation on Social Media: An Empirical Approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24 (1), 013130.