Situational Factors Influencing Receptivity to Bullshit

Mitch Brown, Lucas A. Keefer, Shelby J. McGrew

[a] Fairleigh Dickinson University, Teaneck, NJ, USA. [b] University of Southern Mississippi, Hattiesburg, MS, USA. [c] Southeast Louisiana Veteran's Health Care System, New Orleans, LA, USA.


Received: 2019-06-17 • Accepted: 2019-09-02 • Published (VoR): 2019-10-24

Handling Editor: Jarosław Piotrowski, Cardinal Stefan Wyszyński University, Warsaw, Poland

Corresponding Author: Mitch Brown, Fairleigh Dickinson University, 1000 River Rd, Teaneck, NJ 07666, USA. E-mail: mbrown41@fdu.edu

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]

Abstract

Individuals are motivated to maintain a sense of meaning, and enact cognitive processes to do so (e.g., perceiving structure in the environment). This motivation to find meaning may ultimately impact humans’ interpretation of “bullshit”, statements intended to convey profundity without any. Conversely, subtle cues threatening the meaningfulness of bullshit may elicit greater skepticism. Three studies tested situational factors predicted to heighten or diminish susceptibility to bullshit by changing motivations to seek meaning. We employed diverse methods including symbolic meaning threat (Study 1), social exclusion (Cyberball; Study 2), and manipulating cognitive fluency (Study 3). Taken together, the results indicate basic processes shaping the detection of meaning have implications for the appraisal of ambiguously insightful information.

Keywords

bullshit, meaning, critical thinking, threat, exclusion

In navigating social interactions, individuals must understand meanings conveyed by others. Such navigation is imperative in identifying trustworthy others while also being cognizant of the fact that certain individuals may be attempting to produce a positive impression of oneself as competent and prosocial (Leary & Kowalski, 1990; Paulhus & Reid, 1991; Sedikides & Gebauer, 2010). For example, individuals must distinguish true from false (or distorted) information, something done with considerable accuracy, albeit imperfectly (Bond & DePaulo, 2006). This process requires identifying less-than-truthful in-
formation, which could include displays of so-called bullshit. *Bullshit* refers to the use of misleading information, often typified by vague or ambiguous wording, produced in the service of impression management, despite its lack of contextual truth (Frankfurt, 2005; Pennycook, Cheyne, Barr, Koehler, & Fugelsang, 2015a). Bullshit is typically intended to convey profundity and competence to others. Although sometimes passing as a mere strategy to wow a friend or woo a romantic partner, the costs of bullshit can be high when taken to the level of public misinformation or blanket misconceptions about the nature of the political and physical world (Luks, 2017).

To determine why individuals are swayed by bullshit, we explored the unique possibility that an individual’s susceptibility to bullshit depends on basic processes that shape the detection of meaning. We draw on a wealth of literature exploring the view that people strive to live meaningful lives with purpose and therefore attempt to instill thematic coherence in information around them. Although this claim has long been a staple of humanistic and existential perspectives in psychology, renewed interest in meaning demonstrates the contemporary significance of this claim, particularly related to the presence of, and search for, meaning (Steger, Frazier, Oishi, & Kaler, 2006). Individuals with meaning in life report greater life satisfaction, more positive emotions, higher levels of optimism, and better self-esteem (Compton, 2000; Steger & Kashdan, 2007; Steger, Oishi, & Kashdan, 2009; Zika & Chamberlain, 1987). Conversely, search for meaning corresponds with negative outcomes, including depression and anxiety (Steger et al., 2006; Steger et al., 2009).

Although individual differences in peoples’ abilities to distinguish between truth and bullshit have been demonstrated (Pennycook et al., 2015a), less is known about the extent to which meaning-detection processes may elicit greater openness to bullshit or skepticism toward it. The current paper offers three preliminary experimental tests of the extent to which processes that decrease perceived meaning influence receptivity to bullshit in which we identify situations with individuals becoming more receptive to bullshit and those in which individuals become less receptive.

**Bullshit Receptivity**

Group living is essential to human survival, as it facilitates individuals’ access to resources and mates while ensuring continued protection from predation and exploitation. This reliance has since necessitated individuals to attempt to inflate their social value to others, even if such information is not necessarily truthful. Although lying involves recognition of the truth, a process that philosophers regard as bullshit is characterized by a lack of regard for the truth in the service of appearing knowledgeable (Frankfurt, 2005). The primary goal of “bullshitting” is to impress others with a mere display of insight or intelligence, which could create an impression of oneself being more competent than one actually is (Leary & Kowalski, 1990; Paulhus & Reid, 1991). This would afford bullshitters the opportunity to reap the benefits of group living.
Despite its lack of deep meaning, people nonetheless vary in their receptivity toward bullshit and may ultimately find meaning in it. *Bullshit receptivity* is an individual difference in how much meaning individuals find in pseudo-profound language that sounds impressive, but lacks clear semantic content (e.g., “Hidden meaning transforms unparalleled abstract beauty;” Pennycook et al., 2015a). Initial research finds that this dimension correlates with a susceptibility to heuristics and cognitive biases, paranormal beliefs, and endorsement of conspiracy theories that is prevalent across various political ideologies (Nilsson, Erlandsson, & Västfjäll, 2019; Pennycook, Fugelsang, & Koehler, 2015b). This receptivity additionally predicts reduced willingness to engage in prosocial behavior, suggesting bullshit may simply serve to foster social connections without deeper expectations in the future (Erlandsson, Nilsson, Tinghög, & Västfjäll, 2018).

Two cognitive mechanisms have been proposed to explain variability in receptivity to bullshit. First is a cognitive bias toward accepting information as true or factual, whereas the other is an inability to detect meaningful differences between bullshit and other information (Pennycook et al., 2015a). In addition, the ability to detect conflicts between bullshit and reality may help critical thinkers reject meaningless information. Accordingly, cognitive monitoring failures may lead to an inability to identify bullshit (Pennycook et al., 2015a). The ability to distinguish conflict involves purposeful, analytical thinking and more bullshit receptive audiences may just be poorer at this task (Pennycook et al., 2015b). Such mechanisms may ultimately play a role in navigating and discerning between information based on its veracity. For example, individuals who deliberate at length over information, as measured by cognitive reflections tests, are more skeptical of “fake news,” a potential real-world corollary to bullshit (Pennycook & Rand, 2018).

Although initial work has demonstrated that bullshit receptivity is a meaningful individual difference, little is known about contextual factors that may influence this dimension. Given that individual differences can vary based on situational factors (Dunlop, 2015), it would be sensible to identify situations wherein receptivity may vary. To explore this topic, we draw upon prior research concerning threat management, particularly as it pertains to relevant context of meaning, and processing fluency to determine whether people find meaning in intrinsically meaningless bullshit based on these influences. We specifically predicted that those situationally cued to seek greater meaning following a symbolic threat (Study 1) and exclusionary experiences (Study 2) will be more receptive to bullshit. Conversely, we predicted that those experiencing diminished fluency (an index of meaningfulness) will be less receptive (Study 3).

**Study 1**

One situational factor that could heighten individuals’ receptivity to bullshit is a threat to one’s global sense of meaning in life. The Meaning Maintenance Model proposes that threat to this sense of meaning (i.e., orderliness and coherence in the world) results in
compensatory behaviors that ease the aversive arousal associated with uncertainty (Heine, Proulx, & Vohs, 2006). Work in this area has demonstrated a small set of psychological tools that individuals can employ to restore a sense of comprehensibility to one’s existence. For example, individuals with a threatened sense of meaning sometimes abstract information from the environment to find meaningful patterns in complexity (see Proulx & Inzlicht, 2012).

Experimental research indicates that meaning threats through absurdism foster a search for coherence, particularly in stimuli without inherent meaning. For example, reading the work of Franz Kafka, an author whose work is capable of fostering meaning threat (Camus, 1955), heightens comprehension of an artificial grammar by motivating individuals to detect novel patterns in letter strings (Proulx & Heine, 2009). Individuals have been shown to seek meaning in response to the threat posed by surrealist art (e.g., Kafka, Monty Python), which ultimately fosters a desire for structure (Proulx, Heine, & Vohs, 2010). These findings suggest that meaning threat may enhance efforts to seek out information that affords some degree of coherence, even if it is bullshit that is not objectively true or meaningful. Our first study tested the possibility that meaning threat may increase bullshit sensitivity precisely because bullshit seems to offer some cognitive closure (however misguided) to the threatened individual. To do so, we employed the validated Kafka manipulation used in past research and measured perceptions of bullshit.

Method
Participants
We recruited 109 participants from a public university in Southeastern U.S. in exchange for course credit. Of these, three were excluded from the analysis for failing to respond to the outcome measure, but no other exclusion criteria were employed. This left a final sample for analysis of 106 (22 Men, 83 Women, 1 Other; \( M_{\text{Age}} = 20.87, SD = 4.73; 47\% \) White). We had no a priori predictions for power in all studies and sought to recruit as many participants as possible within the given semester in which they were conducted. A sensitivity analysis indicated that we were appropriately sampled to detect medium-small effects in this study (Cohen’s \( f = 0.13, \beta = 0.80 \). It should be additionally noted that participants in each study were precluded from participating in the other studies presented in this paper.

Materials and Procedure
Meaning threat — Upon entering the laboratory, participants initially read either a meaningful or absurd parable to differentially activate meaning threat (Proulx et al., 2010). The meaningful parable was a version of Aesop’s fable, The Tortoise and the Hare, a ubiquitous story emphasizing the rewards of humility and hard work (\( n = 53 \)). The absurd parable, which served to elicit a meaning threat, was an excerpt from Kafka’s An Imperial Message, a story about an emperor’s messenger being sent to deliver a message...
but is unable to find his way out of the palace \( (n = 53) \). Although this story initially provides expectations of success, the messenger’s efforts prove futile, imparting the message of failure’s inevitability. Previous research has found Kafka to effectively invoke meaning threat in participants (Proulx et al., 2010), so no additional manipulation check was included.

**Mood** — Participants then responded to 19 state negative affect items (e.g., “Afraid,” “Irritable”) from the Positive Affect Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) to assess the potential effects of meaning threat on mood. Participants rated the extent to which they felt each of these states using a 5-point scale (0 = *Very slightly or not at all*, 4 = *Extremely*) and the items formed a reliable composite \( (\alpha = 0.95) \). Ratings were summed to provide an estimate of overall negative affect.\(^1\) Although Proulx and colleagues (2010) found no difference in state affect between the story conditions we employed, we sought to follow their procedure and to determine whether any effects of story condition may be due to heightened negative affect following threat. We then included a distractor task asking participants to rate the importance of objects for a hypothetical camping trip as a means to bolster meaning maintenance motivation (Greenberg, Pyszczynski, Solomon, Simon, & Breus, 1994).

**Bullshit receptivity** — Participants responded to the profundity of 40 statements presented in a random order. Twenty of these statements were mundane filler statements expressing broad truisms (e.g., “Lazy people don’t usually succeed in life”). The other 20 were categorically bullshit from the Bullshit Receptivity Scale (BRS; Pennycook et al., 2015a), derived from the “New Age Bullshit Generator” and the Twitter account of Deepak Chopra. Statements were vague sentences containing buzzwords and jargon designed to create an impression that the speaker was profound without necessarily being clear, true, or meaningful (e.g., “The future will be an astral unveiling of inseparability”). Participants rated each statement’s profundity on 5-point Likert-type scales (1 = *Not at all Profound*, 3 = *Fairly Profound*, 5 = *Very Profound*) with higher scores reflecting perceptions of greater profundity. Both bullshit \( (\alpha = 0.91) \) and filler statements \( (\alpha = 0.88) \) demonstrated acceptable reliabilities. Although typically addressed as a concept in the United States and Canada (e.g., Pennycook et al., 2015a), more recent findings have indicated a cross-cultural understanding of what Frankfurt (2005) initially described as bullshit (e.g., Blanchette & da Silva, 2012; Erlandsson et al., 2018; Nilsson et al., 2019).

In this and all subsequent studies, mean ratings of the bullshit and filler statements were significantly positively correlated \( (all r < 0.50) \). However, given that such correlations were only moderate at most, it was sensible to treat them as separate variables in our analysis. For all samples, we tested whether the correlation between mean ratings of

---

\(^1\) One participant skipped an item in the PANAS and their data was excluded test-wise for all analyses involving affect. They were included in other analyses to maximize power.
the two statements also differed significantly from 1 (it did in all cases; $p s < 0.001$). In other words, although statement ratings correlated, they were not reducible to a single dimension due to unique variance for each statement type.

**Personal need for structure** — Finally, participants completed the 12-item Personal Need for Structure scale (Neuberg & Newsom, 1993). The measure asks participants to rate their agreement ($1 = \text{Strongly Disagree}; 6 = \text{Strongly Agree}$) with statements assessing their tolerance for ambiguity and spontaneity (e.g., “It upsets me to go into a situation without knowing what I can expect from it). In the current sample, one reverse-scored item was not reverse-loading (item-total correlation: $r = 0.30$) and was thus excluded from analysis. The remaining items formed a reliable composite ($\alpha = 0.76$).

**Results and Discussion**

**Mood**

Summed mood scores were submitted to a Poisson regression model by condition (0 = Control; 1 = Threat) (to account for the count nature of the data and the large positive skew [skewness $= 2.27$]). This test returned a significant difference between conditions, $b = 0.43, SE = 0.07, 95\% \text{ CI} [.295, .570], z = 5.88, p < .001$. Specifically, while the model implied a negative affect score of 5.827 in the Control condition, there was an estimated increase of 1.537 ($e^{0.43}$) in the Threat condition.$^2$

**Bullshit Receptivity**

We submitted receptivity scores to a 2 (Condition: Threat vs. Control) × 2 (Statement: Bullshit vs. Mundane) mixed-model ANOVA over the latter factor. Effects were qualified by a Condition × Statement interaction, $F(1, 104) = 3.95, p = .050, \eta_p^2 = 0.037$. Neither main effect was significant, $F s < 0.22, ps > 0.645$. Simple effects revealed that participants in the control condition found the mundane statements to be marginally more profound ($M = 3.01, SD = 0.70$) than the bullshit statements ($M = 2.84, SD = 0.68$), $F(1, 104) = 3.01, p = .086, \eta_p^2 = 0.028$. In the threat condition, the mundane ($M = 2.88, SD = 0.72$) and bullshit ($M = 2.98, SD = 0.67$) ratings did not differ, $F(1, 106) = 1.15, p = .285, \eta_p^2 = 0.001$. Viewed another way, no differences emerged in comparing the receptivity of either type of statement as a function of condition, $F s < 1.15, ps > 0.280$ (see Figure 1).

---

$^2$ Testing the same model in a standard linear regression rendered this test in the same direction, but non-significant, $b = 3.15, SE = 2.25, t(103) = 1.40, p = .16$. However, as noted, the data violate assumptions and are best analyzed in a way that properly handles their extreme skew.
Personal Need for Structure

Personal Need for Structure scores did not differ by condition, $F(1, 104) = 0.09, p = .767, d = .056$ ($M_{\text{Threat}} = 4.02, SD = 0.72; M_{\text{Control}} = 3.98, SD = 0.70$). Individual differences in PNS did not moderate the effect of condition on bullshit ratings, mundane ratings, or the difference score (all $\beta$s < 0.03; all $p$s > 0.75).

Results provided preliminary evidence that evoking meaning threat can elicit motivation to instill thematic coherence in available information, even if such information is meaningless bullshit. Specifically, participants who read a story that motivated meaning management viewed bullshit and mundane statements as equivocally profound, whereas those in the control condition viewed bullshit as marginally less profound. This aligns with previous research indicating that threat instills a search for meaning (e.g., Proulx & Heine, 2008, 2009). That is, meaning threat makes individuals more susceptible to the intended use of bullshit to make meaningless statements seem more profound. Furthermore, we observed no moderation of this effect by variation in the need for clear or unambiguous information.

Nonetheless, the difference between receptivity to bullshit and mundane statements in the control condition was only marginal. This marginal difference could reflect that the manipulation was not particularly impactful in threatening meaning in this population. Although some previous findings demonstrate that Kafka is itself threatening to individuals’ sense of meaning (Proulx et al., 2010), we next considered a stronger threat manipulation that is more engaging and therefore capable of eliciting diverse threats. This multifaceted threat allowed us to better understand situational factors that might influence bullshit receptivity by considering both threats of meaning in influencing receptivity to bullshit but also additional related threats that could similarly heighten receptivity.
Study 2

Although meaning threat appeared to influence receptivity to bullshit to some degree, Study 1’s manipulation provided only modest support. We sought to conceptually replicate these findings in Study 2 using a more impactful manipulation. We specifically utilized Cyberball, an experimental paradigm that elicits an exclusionary experience and affords the opportunity to consider additional threats elicited by ostracism that is both more engaging and capable of eliciting perceptions of threat more immediately (Hartgerink, Van Beest, Wicherts, & Williams, 2015).

Feelings of ostracism elicit perceptions of threats to reflexive (belonging, self-esteem) and reflective needs (control and meaning; Williams, 2009). Reflective needs refer to post-evaluative motivational states, wherein individuals attempt to instill coherence in their lives, with exclusionary experiences thwarting satisfaction of one’s ability to have a meaningful existence and exercise control in life (Oaten, Williams, Jones, & Zadro, 2008). Individuals additionally endorse more religious thinking following an exclusionary experience, potentially in the service of re-instilling thematic coherence in their existence (Aydin, Fischer, & Frey, 2010). Thus, although we are still interested in identifying the effects of meaning threats on receptivity to bullshit, Cyberball affords us the latitude to consider multiple threats in the same paradigm, given how pervasively threatening ostracism is.

Given the robust effects of a Cyberball manipulation in threatening reflective needs (e.g., Goodwin, Williams, & Carter-Sowell, 2010) and our prediction that these needs would motivate people to seek meaning in bullshit, we predicted that threats to both meaning and control would heighten bullshit receptivity. To test this, we specifically explored whether social exclusion would heighten receptivity to bullshit with meaning and control threats serving as mediators of that effect.

Results from Study 1 indicated that participants in the meaning threat condition also reported worse mood compared to those in the control condition. Some have argued that affect serves as a cue for meaningfulness (e.g., King, Hicks, Krull, & Del Gaiso, 2006), which suggests that affect itself is another potential predictor of receptivity. We therefore also expect negative mood to reduce bullshit receptivity. Finally, given that exclusion similarly influences perceived threats toward reflexive needs (Williams, 2009), and that bullshit appears to serve an interpersonal function of impression management (Montoya & Bauer, 2013), we thought it prudent to consider threats to belonging and self-esteem simultaneously as candidate mediators. However, we remain agnostic regarding these threats’ effects and offer no a priori hypotheses.
Method

Participants
We recruited 375 participants from a public university in Southeastern U.S. (275 Women, 99 Men, 1 non-disclosing; $M_{\text{Age}} = 19.56, SD = 3.56$; 59.2% White) in exchange for course credit through online data collection. A sensitivity analysis nonetheless indicated that we had sufficient power to detect smaller effects ($f = 0.07$, $\beta = 0.80$). We excluded no participants from the final analyses.

Materials and Procedure
Consenting participants were randomly assigned to one of two conditions in Cyberball, a simulated cooperative task designed to elicit feelings of inclusion or exclusion (Williams & Jarvis, 2006). Participants played an online ball-tossing game with three computer avatars, with the avatars eventually continuing to include participants in the game or excluding participants. Following the manipulation, participants responded to 12 items assessing the extent to which participants felt that their four basic needs were threatened (3 items each): belonging, self-esteem, meaning, and control needs ($1 = \text{Not at All}; 9 = \text{Very Much So}$; Zadro, Williams, & Richardson, 2004). Participants also completed 4 items assessing mood (e.g., $1 = \text{Sad}; 9 = \text{Happy}$) and a single item assessing feelings of rejection ($1 = \text{Rejected}; 9 = \text{Accepted}$). Participants also indicated the extent to which they received the ball during the task using a sliding scale from 0 to 100%.

Results and Discussion

Manipulation Check
We initially conducted two t-tests to identify the efficacy of the Cyberball manipulation. Excluded participants reported greater feelings of rejection ($M = 6.60, SD = 2.06$) compared to included participants ($M = 3.18, SD = 1.98$), $t(372.44) = 16.36, p < .001, d = 1.69$. Furthermore, included participants reported receiving the ball more frequently ($M = 31.26\%, SD = 16.80$) than excluded participants ($M = 16.76\%, SD = 16.35$), $t(371) = 8.44, p < .001, d = 0.61$.

Primary Analyses
We further submitted our candidate mediators to a series of t-tests (see Table 1 for relevant means and correlations). As predicted, excluded participants reported greater
As an initial analysis for bullshit receptivity, we conducted a 2 (Condition: Exclusion vs. Inclusion) × 2 (Statement: Bullshit vs. Mundane) mixed-model ANOVA with repeated factors over the latter factor. A main effect of statement emerged, such that participants perceived greater profundity in bullshit (\(M = 2.96, SD = 0.65\)) than in mundane statements (\(M = 2.87, SD = 0.83\)), \(F(1, 373) = 5.99, p < .015, \eta_p^2 = 0.016\). Neither a Condition main effect nor interaction emerged, \(Fs < 0.14, ps > 0.710\).

Because of the possibility of an indirect effect on bullshit receptivity through the proposed candidate mediators (Rucker, Preacher, Tormala, & Petty, 2011), and the temporal nature of ostracism effects (Williams, 2009), we correlated each perceived threat and mood with receptivity. We additionally calculated a difference score from receptivity to bullshit and mundane statements in which higher scores would reflect greater receptivity to bullshit. We utilized a difference score to reduce the number of omnibus analyses, thus mitigating the likelihood of a Type I Error, and to make our work align with previous correlational work utilizing the measure that uses difference scores to adjust for baseline rates (e.g., Pennycook et al., 2015a). See Table 1 for correlations and descriptive statistics.

Because our candidate mediators were impacted by the Cyberball manipulation, we finally modeled a multiple mediation model treating Condition as the predictor, each need threat and mood as mediating variables, and the difference score for receptivity as the outcome. This was conducted using Model 4 of PROCESS (Hayes, 2013) using 5,000 bootstraps as a single analysis to minimize the number of analyses. Although correlations between some threats were as high as \(r = 0.82\), each threat is nonetheless theoretically distinct from the other (Williams, 2009), and predicts perceptual acuity toward social information in their own unique capacity. For example, threats to self-esteem mediated the link between social exclusion and preferences for Duchenne smiles (Bernstein,
Thus, to ensure the most granular understanding of our analyses, we considered each threat and mood separately within the same model.

We found that threats to belonging needs mediated the association between Condition and Bullshit Receptivity as an indirect effect, such that greater perceived threats of belonging heightened receptivity, 95% CI [0.052, 0.423]. Threats to self-esteem also mediated the link, such that esteem threat reduced receptivity, 95% CI [-0.255, -0.004]. Finally, mood mediated the link, such that Cyberball’s decrease in positive mood elicited a corresponding decrease in receptivity, 95% CI [0.035, 0.149]. Threats to meaning or control needs did not mediate this association. See Figure 2 for mediation pathways.

![Figure 2. Mediation pathways in Study 2.](image)

**Note.** Total Effect = 0.07, SE = 0.07, 95% CI [-0.06, 0.21]. Effects reported are unstandardized.

†p < .10. *p < .05. **p < .01.

Results from this study provided mixed support for our hypotheses. Although an exclusionary experience itself did not elicit bullshit receptivity, we were nonetheless able to identify various indirect processes through which receptivity emerges following social exclusion. That is, negative mood and threats to both belonging and self-esteem mediated the association between exclusion and receptivity. Unexpectedly, threats to meaning and control did not mediate this link. This could suggest that Cyberball’s elicitation of reflective need threats may be specific to meaning and control within affiliative domains and may not be a generalizable threat.
Interestingly, threats to belonging and self-esteem needs mediated the association between exclusion and bullshit receptivity, but in opposing directions. For belonging needs, heightened perceptions of threat subsequently increased receptivity. This receptivity could reflect sensitivity to the proposed interpersonal function of bullshit in impression management (Frankfurt, 2005; Leary & Kowalski, 1990). Humans have a fundamental need to affiliate with others and perceive others’ bullshit displays as especially ingratiating (Baumeister & Leary, 1995; Montoya & Bauer, 2013). Threats to belongingness could motivate individuals to recognize the social value of bullshit in identifying desirable others, which could have subsequently satisfying affiliative needs (Lambert et al., 2013). These findings align with previous research indicating that the thwarting of reflexive needs following social exclusion specifically heightens interest in highly affiliative others (Brown, Sacco, & Medlin, 2019).

Although exclusion similarly elicited greater threats to self-esteem, its mediating role was in the opposite direction of belonging, such that perceiving one’s self-esteem as threatened reduced receptivity to bullshit. Bullshit information may be ingratiating, yet remains nonetheless misleading, with perceivers recognizing bullshitters as being “full of it” (Montoya & Bauer, 2013). Given that perceiving one’s self-esteem as threatened may operate similarly to negative affect, these results could be unsurprising. Further, a sociometer account of social selection posits that individuals become especially sensitive toward deceptive social information following social exclusion, which would facilitate the identification, and avoidance, of suboptimal affiliative opportunities (Leary & Baumeister, 2000). For example, social exclusion heightens individuals’ perceptual acuity toward Duchenne smiles, a genuine emotional display that veridically signals affiliative intent, from non-Duchenne smiles, a deceptive display (Bernstein, Young, Brown, Sacco, & Claypool, 2008). Alternatively, individuals with lower self-esteem may have been attempting to improve their self-worth by derogating the statements. The dual direction of reflexive need threats predicting receptivity may represent a balancing of competing motives and explain the absence of any main effect of condition in the current study. Whereas belonging concerns may facilitate ingratiation toward bullshit in the service of ingratiation, self-esteem may temper the ingratiation to prevent exploitation because of its misleading nature.

The previous two studies found evidence that certain temporal threats increase susceptibility to bullshit. However, it may also be possible that other factors decrease susceptibility to bullshit by reducing perceived meaning. If individuals not experiencing a temporal threat are more capable of thinking analytically (e.g., Epley, Akalis, Waytz, & Cacioppo, 2008; Swami, Voracek, Stieger, Tran, & Furnham, 2014), it would seem sensible to predict that engaging critical thinking faculties would ameliorate susceptibility to bullshit. Study 3 leveraged work on cognitive fluency and meaning to test whether reductions in fluency inoculate individuals against finding meaning in the meaningless.
Study 3

The purpose of Study 3 was to identify measures to buffer individuals from becoming receptive to bullshit. Thinking critically about the nature of bullshit may reduce individuals’ receptivity to it. Previous research has shown that fostering analytic thinking ultimately heightens individuals’ endorsement of scientific findings over other beliefs (e.g., endorsing evolution over creationism; Gervais, 2015) and reduces beliefs in conspiracy theories, an established correlate of bullshit receptivity (Hart & Graether, 2018; Swami et al., 2014).

One process through which changes in critical thinking may occur is cognitive fluency. Cognitive fluency refers to the relative ease or difficulty associated with processing information (Alter & Oppenheimer, 2009). Whereas some thinking is quick, intuitive, and effortless (System 1 thinking), other times thinking is slow, deliberate, and effortful (System 2). System 1 exhibits a level of fluency and is typically engaged when information is presented in easily understood capacities. Conversely, information that is difficult, confusing, or complex is often relegated to the disfluent System 2 approach, which could foster critical thinking, as individuals would have to engage with such information more critically than with System 1 (Alter, Oppenheimer, Epley, & Eyre, 2007). Disfluency affords individuals the opportunity to think abstractly about concepts (Alter & Oppenheimer, 2008).

Cognitive fluency (or lack thereof) exerts powerful effects on thinking generally and the perception of meaning specifically (e.g., Schwarz, 2004, 2005). For instance, those who could more easily recall features of their “true self” were more likely to report meaning in life (Schlegel, Hicks, King, & Arndt, 2011). Other research finds a strong correlation between meaning in life scores and perceived ease of completing the measure (although this link was moderated by religiosity; Davis & Hicks, 2016). germane to this research is work on cognitive disfluency’s effect on perceptions of meaning in life (Trent, Lavelock, & King, 2013). When reading statements on meaning-in-life inventory that were written in either disfluent font (e.g., five different fonts per item; very small print) or more fluent, easy-to-read font (e.g., large 48-point font), participants’ meaning-in-life ratings were significantly higher for individuals given easy-to-read fonts compared to the disfluent fonts.

Individuals’ presented with disfluent information think critically and analytically, as evidenced by better performance on syllogistic reasoning tasks (Alter et al., 2007). This reduced reliance may suggest greater use of logical faculties over the more intuitive thinking that might cause people to find significance in bullshit. Furthermore, asking participants about their belief in a higher power in disfluent fonts fosters greater heightened critical thinking, and subsequently religious disbelief, relative to those similarly tasked in easy-to-read fonts (Gervais & Norenzayan, 2012). Such skepticism toward religion also predicts less receptivity to bullshit (Pennycook, Ross, Koehler, & Fugelsang, 2016). Taken together, these results indicate that engaging critical thinking capacities could cause individuals to show greater skepticism toward bullshit statements.
In Study 3, we built on this research showing that disfluency reduces perceived meaning. Specifically, we presented participants with the same bullshit and normal statements employed in the first study, but experimentally manipulated presentation to lower fluency, which would foster differences in critical thinking. We predicted that decreased fluency, because it engages critical thinking and decreases perceived meaning, would elicit reduced receptivity to bullshit.

**Method**

**Participants**

We recruited 319 participants from a public university in Southeastern U.S. (239 Women, 79 Men, 1 reporting “Other”; $M_{Age} = 20.42, SD = 4.53; 60\%$ White) in exchange for course credit through online data collection. We excluded no participants from the final analyses. A sensitivity analysis indicated that the current sample was sufficient to detect small effects (Cohen’s $f = 0.08, \beta = 0.80$).

**Materials and Procedure**

Participants indicated the profundity of the same 20 BSR statements from Study 1 mixed with 20 filler statements. Importantly, such statements were either written in a standard Calibri font (sample) that afforded easy readability (i.e., Control, $n = 157$) or in unpredictably chosen fonts in light grey to reduce ease of readability to evoke critical thinking (sample, $n = 161$). That is, participants would ostensibly have a more difficult time reading these manipulated fonts and would have to engage with the items more strenuously in the critical thinking condition than in the control condition. Such disfluency manipulations are common to directly foster critical thinking and reduce intuitive judgments (Alter et al., 2007; Gervais & Norenzayan, 2012; Trent et al., 2013). Therefore, to be consistent with previous research, we did not utilize a direct manipulation check. Both bullshit ($\alpha = 0.88$) and filler statements ($\alpha = 0.87$) demonstrated acceptable reliabilities. This was followed by demographics and debriefing.

**Results and Discussion**

We conducted a 2 (Font: Fluent vs. Disfluent) × 2 (Statement: Bullshit vs. Mundane) mixed-model ANOVA with repeated factors over the latter factor. A main effect of statement emerged, such that participants found the bullshit statements to be significantly more profound ($M = 2.95, SD = 0.56$) than the mundane statements ($M = 2.77, SD = 0.66$), $F(1, 316) = 23.50, p < .001, \eta^2_p = 0.069$. Another main effect of condition emerged, such that participants found the statements to be less profound in the critical thinking condition ($M = 2.78, SD = 0.58$) than in the control condition ($M = 2.92, SD = 0.63$), $F(1, 316) = 7.07, p = .008, \eta^2_p = 0.022$. However, effects were not qualified by a 2-way interaction, $F(1, 316) = \ldots$
0.14, \( p = .710, \eta^2_p = 0.000 \), suggesting that disfluency’s effects were equivalent for both statement types.

Despite the predicted interaction not emerging, disfluent fonts ultimately reduced receptivity to all statements as profound due to participants thinking more analytically. This prompted us to consider whether these reductions in receptivity elicited categorical skepticism. We conducted one-sample \( t \)-tests to determine the categorical profundity of bullshit in both conditions. We compared responses to bullshit statements against a scalar midpoint of 3, which indicates perceptions of a statement as being fairly profound. Whereas the fluent condition elicited a perception of bullshit as being fairly profound (i.e., no different from the midpoint; \( M = 3.02, SD = 0.57 \), \( t(156) = 0.50, p = .61, d = 0.08 \)), participants in the disfluent font condition found the bullshit to be categorically less profound (\( M = 2.88, SD = 0.54 \), \( t(160) = 2.68, p = .008, d = 0.42 \). In other words, critical thinking categorically fostered skepticism toward bullshit, as evidenced by its above-chance score below the midpoint.

Results from Study 3 suggest engaging critical thinking faculties may be an important determinant in one’s receptivity to meaningless bullshit. We found that participants exposed to disfluent statements (i.e., in unpredictable fonts) tended to rate the bullshit statements as less profound. These findings are consonant with the skepticism found in previous literature toward certain intuitive beliefs (Gervais & Norenzayan, 2012; Trent et al., 2013), and suggest that critical thinking helps individuals distinguish between truth and bullshit, which indicates it may be an important cognitive tool used to make sense of the world.

**General Discussion**

Across three studies, we identified situational factors shown to influence perceived meaning that cause individuals to heighten and diminish their receptivity toward bullshit. Additionally, we made efforts to identify specific processes through which these changes occur. Specifically, we found evidence that various threats heighten receptivity (Studies 1 and 2) and processing disfluency reduces such receptivity (Study 3). These studies built upon early work on bullshit receptivity (e.g., Pennycook et al., 2015a) by showing preliminary evidence that stable individual differences vary across situations and that such receptivity can increase and decrease.

These studies contribute to a growing corpus of work identifying how temporally activated motivations to instill meaning facilitate identification of patterns from semantically meaningless, or artificial, content (e.g., Proulx & Heine, 2008). Such a response may be adaptive in the service of ensuring individuals have a coherent view of their environment and sense of self, which could subsequently direct goal attainment (Sedikides & Skowronski, 2003). Recognizing bullshit as profound could be a proxy for finding the meaning necessary for intentional actions. However, this meaning threat management...
appears to have boundary conditions insofar as meaning threat only appeared to heighten receptivity toward general semantic processes, but not meaning threats related to social contexts. Within social contexts, belonging threat heightened receptivity to bullshit, which may ultimately speak to the interpersonal function of such displays via impression management (Leary & Kowalski, 1990). Given both the overall likability of bullshitters (Montoya & Bauer, 2013) and an interest in affiliative conspecifics following exclusionary experiences (e.g., Brown et al., 2019), these findings seem sensible. Future research would benefit from considering bullshit’s evolutionary function more extensively by assessing affiliative interest in bullshitters following an exclusionary experience.

These findings can provide a starting point for future research to identify how situational factors alter bullshit receptivity. In an era of “fake news” and social media saturation, it is important to understand both when people are drawn to meaningless things that sound good and when they establish sufficient skepticism. Indeed, previous research indicates that endorsement of fake news is rooted in a lack of critical thinking (Pennycook & Rand, 2018). These results are consonant with previous findings in that fostering critical thinking, potentially through reduced meaning searches (see King & Hicks, 2009), heightens skepticism toward manipulations of the truth. This further implicates fake news as an outgrowth of bullshit, thus suggesting our data have the capability to mitigate the effects of potentially problematic media presentations. More specifically, these data have considerable application to the development of interventions to inform individuals about the biases that lead to receptivity to fake news as well as potential critical thinking tasks that can heighten one’s ability to recognize bullshit more readily.

**Limitations and Future Directions**

Despite identifying a level of process in the current program of research, one limitation is our reliance on simple designs. Although this affords a simple interpretation and maximizes statistical power, it leaves open the possibility for several confounds. One clear example of this is Study 1: Our reliance on a simple two-cell design (following past research) affords little information on why effects occurred and/or whether the neutral story serves as a truly neutral control group. Although previous research demonstrates that satisfaction of salient meaning needs can itself be a control in some paradigms (Dvir, Kelly, & Williams, 2019), future research would benefit from providing an additional control group in which meaning is neither threatened nor satisfied.

Additionally, although we found instances in which meaning and belonging threats induced greater bullshit receptivity, it may be that an unpleasant experience is the primary basis of these effects. Indeed, negative affect was an especially robust mediator in Study 2. However, these various differences in affective states could have ultimately been a byproduct of precedingly activated motivational states, with one’s affect simply acting to motivate individuals to address the perceived threats. Although Study 2 provided evidence for concomitant mediation with threats to belonging and self-esteem, future re-
search would benefit from further teasing apart affect and threatened states. One possible manner in which this could occur may include asking people to imagine a scenario in which they suddenly became deficient in a relevant need (e.g., Gaertner et al., 2012), wherein participants would be instructed to consider measures to replenish this need in subsequent tasks.

Additionally, this work neglects the role of individual differences established in previous work to focus on experimental designs. However, because individuals vary in their dispositional receptivity to bullshit, it is likely that any effects of situational variables will be moderated by those baseline differences. The dependent measure we employed is the same one used to assess individual differences, which means that any future test would have to accept the costs of giving participants the same items before and after a manipulation or find another method to assess bullshit receptivity. For example, a future study could consider dispositional levels of personal need for structure, as individual differences in PNS are robust predictors for how individuals infer ambiguous or unfamiliar information (Moskowitz, 1993; Schaller, Boyd, Yohannes, & O’Brien, 1995). It would also be advantageous to consider individual differences related to affiliative motives (e.g., trait self-esteem), given that dispositionally high self-esteem buffers individuals against endorsing conspiracy theories (Cichocka, Marchlewksa, & de Zavala, 2016). Given an established association between individual differences in need for cognition with skepticism toward bullshit (Pennycook et al., 2015a), future studies could consider this trait’s interactive effects with analytic processing induced from the disfluency paradigm in Study 3.

Participants high in need for cognition could have already been resistant to perceiving bullshit as profound, which could have precluded our ability to identify the full extent of the manipulation’s efficacy. It could be the case that those with a lower need for cognition could have been susceptible to information that relies on intuitive thinking styles (Pennycook & Rand, 2018; Pfattheicher & Schindler, 2016). Such knowledge of individuals’ utilization of intuitive heuristics could subsequently inform efforts to develop interventions targeting the basis for bullshit endorsement to ensure individuals become immune to its effects (e.g., Pennycook, Cannon, & Rand, 2018).

This speaks to a general value in further considering individual differences that may moderate situational factors influencing bullshit receptivity. For instance, research could additionally consider differences in the academic fields of respondents as a potential individual difference moderator. Within psychology and other social sciences, bullshit and misinformation have become widely discussed topics, yet it is unclear what impact this awareness might have on behavior. Future studies could consider how those highly familiar with misinformation differ from individuals with less familiarity.

One area in which future research can benefit is the consideration of participants’ measurable sensitivity to bullshit beyond simple receptivity. That is, although the current findings demonstrate that individuals perceive bullshit as more profound, it is less clear whether this receptivity is rooted in impression management strategies to appear intelli-
gent or if these individuals consider these statements actually profound and set a lower criterion to perceiving bullshit as profound (Green & Swets, 1966). Indeed, individuals appear aware of when others are “full of it” (Montoya & Bauer, 2013), but various situational factors may ultimately mute individuals’ skepticism toward claims in the service of ingratiating. A study could initially prime individuals with meaning or belonging threat and then present each statement in a categorization task in which individuals indicate whether a statement is bullshit or not (Bernstein et al., 2008). If the basis for threatening experiences’ heightened receptivity toward bullshit is a reduction in sensitivity to bullshit, individuals should be more prone to evaluating bullshit statements as categorically profound.

Conclusion

The current program of research expands upon a field of study investigating responses to misinformation and the impetus behind receptivity to it. Specifically, we identified situational factors that influence one’s proclivity to buy into bullshit or remain skeptical. Such findings may ultimately inform best practices to ensure the accurate dissemination of information to the public.

Funding: The authors have no funding to report.

Competing Interests: The authors have declared that no competing interests exist.

Acknowledgments: The authors have no support to report.

Data Availability: Datasets for the studies are freely available (see the Supplementary Materials section).

Supplementary Materials

All data and syntax for this paper are available via the OSF repository (for access see Index of Supplementary Materials below).

Index of Supplementary Materials


References


