Contents lists available at ScienceDirect

# Public Relations Review

# Examining characteristics of prebunking strategies to overcome PR disinformation attacks

Courtney D. Boman

*Department of Advertising and Public Relations, The University of Alabama, Reese Phifer 404A, Box 870172, Tuscaloosa, AL, 35487, United States*

A B S T R A C T

In 2020, Twitter launched a new strategy dubbed 'prebunking' in hopes of pre-emptively countering false information about voting by mail and election results. Prebunking was touted as a potential solution; however, little empirical research has tested the strategy to examine its effectiveness towards disinformation in the realm of public relations. Exploring PR-based disinformation attacks as a paracrisis, the purpose of this online quasi-experimental design study ($N = 965$) was to investigate the effects of an attack and how prebunking strategies, grounded in inoculation theory, can protect organizational outcomes (i.e., reputation and credibility) by interweaving positive psychology. In addition to examining effects on attitudinal components, this study also inspected dimensions of social amplification to determine if prebunking messages can thwart the spread of disinformation via Facebook. Findings provide insight into advancing the conceptual framework of proactive disinformation responses for crisis communication by demonstrating the success of using prebunking with autonomy support and explicit details to thwart adverse effects of disinformation.

## 1. Introduction

In the days leading up to the U.S. 2020 presidential election, Twitter launched a pilot campaign to help users identify disinformation shared on its platform (Ingram, 2020). The strategy, dubbed 'prebunking,' was initiated in hopes of pre-emptively countering false information about voting by mail and election results in the U.S. Prior to 2020, disinformation and its consequences were brought to the forefront of conversations in the aftermath of the 2016 national elections. With it, the disinformation phenomenon of astroturfing, which is the use of fake grassroots campaigns for political or organizational gain, has become more mainstream. However, there is limited support that provides an in-depth look at the effects of astroturfing within the context of crisis communication. Instead, many of the conversations regarding disinformation have encircled traditional journalism and have not focused on unethical tactics within the public relations field, such as astroturfing. This inhibits both scholars and public relations professionals from understanding its effect on individuals' attitudinal perceptions and behavioral intentions, a vital tool for PR professionals.

When it comes to the classification of astroturf attacks within crisis communication literature, the attacks are considered an advanced form of a paracrisis (Boman & Schneider, 2021). This type of attack falls under the umbrella of a paracrisis due to it being "a publicly visible crisis

threat that charges an organization with irresponsible or unethical behavior" (Coombs & Holladay, 2012, p. 402) Although Coombs and Holladay (2012) adapted the paracrisis as a part of the situational crisis communication theory (SCCT) framework, there is little guidance on the effectiveness of paracrisis response strategies (Honisch & Más Manchón, 2019). Due to the lack of theoretical testing of prebunking disinformation astroturf-based messages, it is unclear to both researchers and PR professionals facing such events on how the public perceives these messages, leaving an absence of knowledge on how to respond.

Thus, at least three research gaps can be identified from these observations. First, there is a need for more empirically tested response strategies to aid PR professionals in dealing with paracrises. It has been previously suggested within literature that no response, or strategic silence, may be warranted when faced with a paracrisis (Coombs, 2019). There is also evidence that the reform strategy is beneficial compared to reform, humor, and refute for self-imposed and external paracrises (Honisch & Más Manchón, 2019). Unlike other paracrises, astroturf attacks provide a unique lens to examine response types, as the messaging primarily uses false information with only hints of truth. This makes it unnecessary for the attacked organization to fully take responsibility for what the attacker says transpired, as suggested within the reform strategies or stealing thunder.

Secondly, paracrises are primarily seen as a phase within issue

management that can be recognized during issue monitoring before morphing into a full crisis (Freberg, 2012; Jin et al., 2014). Taking into account research that has been done to examine how *debunking* misinformation can be successfully implemented after information has been widely spread (Jin et al., 2020), the current research endeavor sets out to test characteristics of *prebunking*. Prebunking, which is a form of proactive communication, can be used before disinformation against an organization provided via astroturf is broadly disseminated. Organizations that detect a potential attack during environmental scanning may be able to use prebunking to reach audience members prior to the full-scale attack, similar to the strategy Twitter used during the U.S. presidential campaign. Recently, researchers have found that proactive strategies, specifically prebunking and PR supportive messaging, are more successful than not responding (i.e., strategic silence) on post-attack outcomes (Boman & Schneider, 2021). These initial findings regarding the potential success of prebunking within crisis communication have not examined specific content strategies, such as use of autonomy support, within the messages. Thus, it is important to examine specific message characteristics to increase the efficacy of prebunking.

Lastly, there is an ongoing societal issue of disinformation and how to decrease its persuasiveness. Social amplification is a concept that can be used to measure the potential for a crisis to perpetuate through social media by asking questions such as individuals' likelihood of reacting, sharing, or commenting on a post negatively or positively on social media (Strekalova, 2017). This type of social media engagement is significant to the spread of information post-crisis because engagement, such as sharing a post, is necessary for reaching wider audiences (Breland et al., 2017; Strekalova, 2017; Swani et al., 2013). This engagement has the potential to expand the reach of a Facebook post exponentially. While protecting an organization's reputation is essential by decreasing the disinformation's persuasiveness, it is even more critical to examine how, and if, organizations can help individuals overcome disinformation attacks on a broader collective level by preventing its amplification.

To examine these three gaps within the literature, the current study works to expand upon the well-developed theoretical models supporting post-crisis communication to empirically support the use of pre-crisis proactive communication. Proactive communication was implemented to assess *prebunking* a disinformation attack that took the form of astroturfing. By interweaving the disciplines of positive psychology, persuasion communication, and crisis communication, this study tests if the use of explicit details and autonomy-support within prebunking messages increases resistance to the persuasive effects of astroturf attacks.

## 2. Literature review

### 2.1. Viewing astroturf attacks as a crisis

Research regarding crisis communication has exploded during the last twenty years, leaving researchers and PR professionals with strong directions on what a crisis is, preparing for a crisis, and communicating to stakeholders reactively within the post-crisis stage. Less common within crisis communication literature is empirically-tested and theoretically-supported proactive response strategies for organizations to use before a crisis. The current study explores if the use of such strategies (i.e., prebunking) could be beneficial for organizations when undergoing an astroturf attack. While the issue of astroturfing is not new, over the last few years its use has become more prevalent in the digital environment. Astroturf attacks provide a unique issue, as it presents itself as a grassroots campaign supported by an individual's peers or members of the general public. However, as the name suggests, it is not organically grown and is considered "fake grassroots" produced by shell organizations, usually for political or organizational gain. These campaigns work to intentionally spread harmful or inaccurate disinformation, which can be particularly troubling for organizations on the receiving end of an attack (Cho et al., 2011; Keller et al., 2020).

Despite the growing popularity of astroturf campaigns, a limited number of studies within the crisis communication paradigm provide insights into the effects of astroturf on audience attitudes and perceptions. In an experiment, Pfau et al. (2007) investigated the effects of corporate front-group stealth campaigns, similar to astroturfing. After individuals were confronted with the disguised corporate messages, the opinions of those initially favoring restrictive policies on different issues significantly decreased. Another study demonstrated that people exposed to astroturf websites became more uncertain than those who saw real grassroots websites about the causes of global warming and humans' role in the phenomenon (Cho et al., 2011). More recently, Zerback et al. (2020) examined the psychological effects of online astroturfing in the context of Russian propaganda and found that overtime, individuals' political opinions and attitudes were influenced by the astroturf messaging, even when participants were inoculated.

The current study works to extend literature that has examined the effects of PR astroturf campaigns on attitudes, along with the initial examination of content strategies within prebunking messages. As research has illustrated the consequences of disinformation (e.g., Cho et al., 2011; Pfau et al., 2007; Zerback et al., 2020), this study is designed to inform individuals of potential disinformation to assist organizations in preventing astroturf-induced crises and thwart its spread on social media.

### 2.2. Creating resistance to persuasive attempts

Astroturf campaigns are a prime example of persuasive communication being deployed, as it is "human communication that is designed to influence others by modifying their beliefs, values, or attitudes" (Simons, 1976, p. 21). Research examining persuasion communication has been an area well studied since the mid-1930s, with researchers exploring areas spanning relational, economic, political, and social change (Dainton & Zelley, 2004). When considering how to limit the persuasiveness of a future message or attack, researchers have worked to examine how to increase or decrease the persuasive attempt through motivated resistance and outcome resistance using theoretical frameworks such as cognitive dissonance theory (Festinger, 1957), social judgement theory (Sherif & Hovland, 1961), principles of congruity (Tannenbaum & Norris, 1965), psychological reactance theory (Brehm, 1966), inoculation theory (McGuire, 1961a), and elaboration likelihood model theory (Petty & Cacioppo, 1986).

Motivated resistance is when an individual does not want to be influenced and is motivated to resist attempts (Ringold, 2002). To do so, the basis of this study examined how mechanisms put forth by inoculation theory can serve as a foundation for proactive communication strategies to create resistance towards disinformation. Similar to a medical vaccine inoculating patients, the theory states that individuals presented with a forewarning message about an attack will be less affected by the persuasive attempt (Eagly & Chaiken, 1993; McGuire, 1961a). Developed by social psychologist William McGuire, the theory proposes that rather than persuading an individual, two-sided messages should be used to create resistance to persuasion (McGuire, 1961a). This occurs due to individuals being exposed to the inoculation message strengthening attitudes and deflecting subsequent attack messages after receiving an inoculation message (Pfau et al., 1997).

When conceptualizing inoculation, McGuire assumed that individuals are a) inexperienced in defending their beliefs and b) unmotivated to do so. To overcome these two impediments, inoculation works to provide individuals with details and resources on defending beliefs while also motivating the individual to defend said beliefs (Wan & Pfau, 2004). Within this process, the mechanism of threat is crucial as it works to motivate an individual to prepare for future challenges by being delivered explicitly or implicitly. This leads to the second crucial mechanism, presenting counter-belief arguments the individual may encounter (Ivanov et al., 2018). This act provides individuals with both materials and guidance on using counterarguments in defense against

future attacks.

### 2.3. Enhancing message characteristics of inoculation messages

After decades of scholarship, studies have produced substantial evidence that inoculation messages work (Ivanov et al., 2016). However, scholars continue to empirically test inoculation theory examining mechanisms affecting the messages' persuasiveness, including threat, timing of the message, refutational preemption, and involvement (Banas & Rains, 2010). A key example of this is research exploring refutational preemption, or the argument and/or evidence provided by inoculation messages. Research findings have differed, especially when considering how explicit an inoculation warning must be. This concept is referred to as refutational same and refutational different preemptions. These two terms refer to whether the inoculation argument is the same as the actual attack (refutational same) or if the attack is novel (refutational different) (Wan & Pfau, 2004). Throughout inoculation literature, there have been varying conclusions on whether there are differing effects of using refutational same or refutational different-based messages (e.g., McGuire, 1961a; McGuire & Papageorgis, 1961; Pfau et al., 1990; Wan & Pfau, 2004). However, a meta-analysis examining 41 inoculation-based studies found no differences between the two types of treatments (Banas & Rains, 2010).

#### 2.3.1. Source derogation

To continue testing ways that enhance the effects of inoculation, the current study tested if providing explicit details about what astroturf is, along with the motivation behind the campaign, can further an individual's ability to realize persuasive attempts compared to a traditional refutational same message. The goal of this message was to provide source derogation, which has been found to be a successful strategy in reducing persuasion throughout the principle of congruity literature (Tannenbaum & Norris, 1965). While both principles of congruity (Tannenbaum & Norris, 1965) and inoculation theory (McGuire, 1961a) look at how timing of a persuasive message influences outcomes, the principle of congruity focuses on a favorable source making a negative assertion against a favorably viewed concept (Tannenbaum, 1967). Congruity messages focus on testing both the effectiveness of source derogation and refutational arguments. The driving force behind source derogation is to reduce perceived credibility and attitudes towards the attack message and source. Supporting this idea, Miller et al. (2013) found that combining source derogation and refutation into the same inoculation message increased reactance and individual's awareness of potential threat to their freedom. The researchers suggested future research focusing on resistance to persuasion examine "enhancing motivation and sensitivity to future threats via inoculation messages, over focusing on threat or specific content of a message" (Miller et al., 2013, p. 150).

#### 2.3.2. Autonomy supportive language

The concept of language intensity has been well studied throughout the area of resistance to persuasion (see Burgoon, 1976; Burgoon & Miller, 1971; Miller & Burgoon, 1979; Stacks & Sellers, 1986). Language intensity has been conceptualized as "linguistic cues that indicate direction and semantic distance from a neutral point" (Burgoon & King, 1974, p. 32). Within language intensity research, the use of controlling language has also been explored, finding that the use of high-controlling language (e.g., a command) can increase reactance or rejection of a message (Miller et al., 2007, 2017).

Differing from previous research in this area, the development of the current study was guided through positive psychology's self-determination theory (SDT) (Deci & Ryan, 1985). Unlike studies throughout persuasion literature that have manipulated language intensity emphasizing attitude change (Burgoon et al., 1975), SDT addresses when people will be internally versus externally motivated to engage in a specific behavior. The theory presents motivation as a vital

way in understanding how an individual's behavior is energized and directed (Gagné & Deci, 2005). To do so, SDT addresses two main ideas: the first being that people have an inherent desire to be proactive and internally motivated. Secondly, humans have three primary needs: competence, relatedness, and autonomy (Deci & Ryan, 1985; Ryan & Deci, 2000, 2017). The current study focuses on addressing the latter.

Similar to characteristics of "low-intensity" messages found within the language intensity literature, autonomy supportive communication is considered low-controlling. By doing so, it can be perceived as legitimate, less threatening, and less demotivating (Miller et al., 2007; Mouratidis et al., 2010). The characteristics of autonomy supportive messaging include providing meaningful choices, acknowledging potential anxieties, and providing strategies that actively allow a person to explore options. When all three of these items are met within a message, evidence shows that individuals can be motivated to further think about or overcome a particular issue (Sheldon & Filak, 2008).

On the other side of the continuum, psychological reactance literature has worked to examine the use of high-controlling language to boost resistance to persuasion by focusing on inducing anger and negative cognitions (e.g., Ivanov et al., 2012; Miller et al., 2013). However, SDT research shows that using controlling language may cause external motivation, leading individuals to feel like pawns in societal concerns. If autonomous motivation is hindered, negative outcomes such as lack of effort and low performance may occur (Gagné & Deci, 2005). However, when individuals are internally motivated, they may take personal responsibility for their actions which motivates reflection regarding one's interests and feelings (Koestner & Losler, 1996). For example, Mouratidis et al. (2010) found that feedback given to athletes using autonomy supporting versus controlling communication was associated with autonomous motivation and led to favorable outcomes. More recently, Wilbur et al. (2021) found promising results that autonomy supportive inoculation messages could boost participants' resistance to extremist propaganda. This suggests that using a level of language intensity that incorporates autonomy support can "boost" internal motivation an individual experiences and aid individuals in overcoming persuasive effects within disinformation.

#### 2.3.3. Using motivation to overcome persuasive attempts

Looking further into the mechanism of motivation presented by SDT, one will find that the theory presents a continuum of motivation levels. These include externally regulated (individual's behavior is initiated and maintained by external agencies), introjected regulation (regulation is controlling individual's behavior), identified regulation (behavior is congruent with individual goals and identify), and integrated regulation (behavior is an integral part of who an individual is) (Gagné & Deci, 2005). This continuum focuses on the internalization of one's "values and behavioral regulations and on the degree to which they have been fully integrated with one's self" (Gagné & Deci, 2005, p. 343). It has been noted that autonomy support is the most important mechanism for predicting identification and integration within SDT (e.g., Gagné & Deci, 2005). Studies within the medical field have found that autonomy support led patients to become autonomously motivated for health behavior change, which ultimately led to greater maintained health behavior change (Williams et al., 1998). This is relevant as the foundation of the current study is to increase an individual's ability to overcome disinformation by being motivated to engage in behavior, such as critically thinking, while viewing an astroturf attack.

### 2.4. Using prebunking to influence crisis outcomes

Inoculation theory has recently been applied to a broadening spectrum of topics, including pre-crisis communication (e.g., Einwiller & Johar, 2013; Ivanov et al., 2016, 2018). Mirroring reactive strategies use of the term *debunking* to overcome misinformation (i.e., Jin et al., 2020), inoculation-based crisis communication strategies have been referred to as *prebunking* (i.e., Boman & Schneider, 2021). An example of

inoculation within crisis communication includes a study by Ivanov et al. (2016) which found that inoculation-based messages may be useful in enhancing the public's belief that government protective agencies can prevent and minimize the effects of politically motivated acts of violence. In another study focusing on a financial crisis, Dillingham & Ivanov (2017) compared inoculation to supportive messaging and no response provided by a company. It was found that inoculation strategies successfully managed stakeholder perceptions of weakness compared to the other strategies.

Previously held attitudes and attitude change have served as fundamental components when examining how to increase resistance to future persuasive attempts. While the social judgement theory may not be overly present in current research, its conceptual ideas still play a prominent role in understanding attitude within persuasion (O'Keefe, 2009). In broad terms, the theory proposes that an individual's reaction to a persuasive attempt depends on previously held positions or attitudes on a topic formed by knowledge, experience, interest, and self-interest (Sherif & Hovland, 1961). This approach relies on gaining insight into individuals' attitudes then analyzing those responses through what is referred to as the contrast effect (falls far away to held attitudes) and the assimilation effect (falls close to held attitudes) (O'Keefe, 1990). To simplify, the theory is built on the idea that individuals have judgements that position persuasive efforts within the range of acceptable, objectional, or neither. The more acceptable an idea is, the more likely an individual sees it as reasonable or worthy of consideration. Depending on how the persuasive effort is classified compared to the original anchor or stance, attitude change is created (Petty et al., 1992).

Due to the interest of examining how autonomy supportive language influences an individual's motivation to critically think about both the prebunking message and the future attack, the following question was asked to gauge that impact on attitude change.

**RQ1.** How, if at all, does the overall change in attitudes pre- and post-astroturf attack towards the attacked organization differ between proactive communication using strategic silence (no response) vs. prebunking strategies?

When looking at the overall purpose of crisis responses, the aim is to mitigate the negative impacts of a crisis, including negatively changing individual's attitudes towards an organization and its reputation. Having a positive reputation can reduce stakeholder uncertainty about organizational performance, motivate consumers to buy products, attract high-quality employees, encourage outside investors, and retain essential transaction partners such as suppliers and distributors (Gardberg & Fombrun, 2006). It generates perceptions among employees, customers, investors, competitors, and the general public about what a company is, what it does, and what it stands for (Fombrun & van Riel, 1997). To protect reputation and other key crisis outcomes, research must understand how organizations can navigate the online environment and how stakeholders interpret the delivery of a crisis response, such as message characteristics of the disseminated content.

Based on previous research within positive psychology and inoculation, the current study tested if combining these strategies outperform solo use within prebunking or strategic silence when examining organizational reputation.

**RQ2.** How, if at all, do individuals' perceptions of the attacked organization's reputation differ between proactive communication using strategic silence (no response) vs. prebunking strategies?

When someone is unable or unmotivated to process a message, they rely on a source's credibility and expertise. Thus, an important post-crisis outcome to consider is perceived source credibility. Source credibility refers to the believability that a message receiver assigns to a communicator and involves considering the source's expertise and trustworthiness (de Meulenaer et al., 2018; Lin et al., 2017; Yoo & Gretzel, 2011). Typically, persuasion and motivation are increased when a message receiver perceives the source to be credible (Eagly et al.,

1978; Lin et al., 2017; Wilson & Sherrell, 1993). It is common for people to rely on the sources they trust, "internalizing information from reliable sources and rejecting information from sources they consider unreliable" (Malka et al., 2009, p. 635). Two research groups (Earle & Cvetkovich, 1995; Poortinga & Pidgeon, 2003) found that people rarely evaluate sources systematically. Instead, they evaluate the levels of similarity between their values and the source's values. This is especially true with complex issues. A central component of an effective inoculation message is refutational preemption, which is the provision of counterarguments with which individuals can defend against future persuasive attacks (Banas & Rains, 2010; Jackson et al., 2015; Parker et al., 2016). Thus, it is important to gauge if the organization presenting the refutational content is perceived as credible once the counterarguments are provided.

**RQ3.** How, if at all, does an individual's perceived credibility of the organization being attacked differ between proactive communication using strategic silence (no response) vs. prebunking strategies?

Since inoculation-based messages work to prebunk future arguments that may threaten the receiver, it was of interest to examine to what extent the messages aid in discrediting or lowering the attacking organization's perceived credibility. Studies using principles of congruity have found that source derogation, such as in messages containing explicit detail, can assist in prompting resistance to persuasive attempts (Compton & Pfau, 2008). Thus, the following question was asked to examine specific characteristics (i.e., autonomy support and explicit details) within the prebunking messages in regards to source credibility:

**RQ4.** How, if at all, does an individual's perceived credibility of the attacking organization differ between proactive communication using strategic silence (no response) vs. prebunking strategies?

To fulfill the third gap within the literature, a less common outcome studied within crisis communication, social amplification, was examined (Chong & Choy, 2018; Pidgeon et al., 2003). When a social media user sees that another person or organization has liked or shared a post, it represents the value of the information for that user (Gittelman et al., 2015; Strekalova, 2017; Sun et al., 2014). This can be referred to as the process of social amplification, which has primarily been used for risk perception research using the social amplification of risk framework (SARF). This framework is rooted in social experience and explains how information can start with one source and become dramatized as shared (or amplified) by others (Chong & Choy, 2018; Kasperson & Kasperson, 1996). Therefore, while a single entity can publish crisis-inducing information about another organization, it can be amplified by an individual or a group of people like an organization or institution (Kasperson & Kasperson, 1996). Similar to word-of-mouth literature that has been applied within crisis communication, amplification of social media posts can have adverse implications on crisis outcomes such as reputation and credibility. As such, the following research question was posed:

**RQ5.** How, if at all, does the use of prebunking messaging (vs. strategic silence) thwart individuals from amplifying the disinformation through 'reacting,' 'sharing,' and 'posting' on an astroturf video posted to Facebook?

### 2.5. Post-crisis debunking

While the primary goal of this study was to explore the effects of prebunking on an individual's resistance to astroturf campaigns, it also examined the differing effects of reactive debunking. The overall goal of strategic crisis communication is to reduce the damage a crisis inflicts on an organization and its stakeholders. Post-crisis communication is often referred to as the reactive stage and is one of the most explored areas of crisis communication research. Building off previous studies examining debunking misinformation (i.e., Jin et al., 2020; van der Meer & Jin, 2020), the current study examined the use of refutation as a debunking

strategy. Refutation was selected as it is a strategy put forth by SCCT as an appropriate response for paracrisis, suggesting an organization argues that the actions or policies being conducted are appropriate (Coombs, 2019). Additionally, it mirrors the content called for by inoculation messages, allowing minimal modifications between the conditions, outside of timing.

Regardless of whether proactive or reactive crisis communication is being used, the ultimate goals are largely the same, protect an organization's image post-attack (Compton, 2017). The dividing difference is in the timing of the strategy, which occurs either before or after the onset of the crisis. Thus, this research endeavor set out to test the difference between prebunking and debunking communication efforts on post-crisis outcomes.

**RQ6**.   How, if at all, do the effects of a refutation response differ within prebunking and debunking on influencing crisis outcomes (i.e., attitudes, perceived source credibility, organizational reputation, and social amplification)?

## 3. Methodology

To examine how prebunking with explicit details and autonomy supportive language can mitigate the damage of an astroturf attack, a between-subjects online quasi-experiment was conducted. Nine experimental conditions were examined for this study, including (pre/de) bunking only, (pre/de)bunking with autonomy support, (pre/de)bunking with explicit details, (pre/de)bunking combined strategy with explicit details and autonomy support, and a message control using strategic silence (no refutational message, astroturf attack only).

### 3.1. Sample overview

Amazon's Mechanical Turk (MTurk) was used to collect data ($N = 1561$). Following the advice from Berinsky et al. (2012), attention checks were used throughout the study to help ensure data quality. After individuals who incorrectly answered the attention checks, did not have audio or video, or did not complete the questionnaire were removed ($n = 596$), a total of 965 participant responses were analyzed. The average age of the sample was 38.60 ($SD = 17.50$), with 483 males (50.1 %) and 474 females (49.1 %), with 0.8% preferring not to say or intersex ($n = 8$). Additional demographic information can be found in Table A1 in Appendix A.

### 3.2. Stimuli development

The stimuli for this experiment used content and branding that represent what is published by both the Humane Society of the United States (HSUS) and Humane Watch to increase the levels of ecological validity. HSUS was selected as the organization is continually attacked by astroturf campaigns from Humane Watch. Only minimal edits occurred between conditions to ensure that content stayed consistent throughout all treatments when accounting for the manipulations, reading levels, and length. Total word counts and the Flesch-Kincaid readability test were used to gauge the comprehensibility of messages and evaluate message equivalence (Kincaid et al., 1975).

Previously conducted inoculation studies served as a template for the stimuli, which provide participants with a forewarning message enacting a sense of threat (Banas & Miller, 2013; Ivanov et al., 2012). However, a unique element of this experiment was that it tested the differences between an inoculation message that included only a broad forewarning element, along with messages that included autonomy support and explicit details (see exemplars of stimuli in Table B1 in Appendix B). Messages containing autonomy-supportive language followed previous guidelines set forth by SDT literature that state autonomy supportive messages provide individuals with control, freedom, and support (Deci & Ryan, 1985; Sheldon & Filak, 2008). The stimuli

that contained explicit details stated what astroturfing is and how Humane Watch (the attacking organization) is funded. As such, the (pre/de)bunking messages from HSUS were broken into four different strategies using (pre/de)bunking, including inoculation only, inoculation with autonomy support, inoculation with explicit details of the attack, and inoculation with both autonomy support and the explicit details. The only modification that occurred between the (pre/de) bunking was the timing of delivery.

### 3.3. Procedure

After reading the informed consent script approved by the Institutional Review Board at a university in the U.S., participants proceeded through the experiment, which took approximately 20 min. First, participants were asked to fill out a feeling thermometer that gauged attitudes towards seven organizations, including HSUS and Humane Watch.

Once this initial measurement was collected, participants were randomized into one of nine conditions. The first group ($n = 413$) was randomized to see one of four prebunking messages, which included prebunking base ($n = 106$), prebunking with autonomy support ($n = 107$), prebunking with explicit details ($n = 102$), and prebunking combining the former two strategies ($n = 98$). These participants were then shown a 30-second attack astroturf video produced by Humane Watch. The second group ($n = 431$) mirrored the first group, with the only difference being that the attack astroturf video was seen first, followed by the reactive refutational debunking messages from HSUS including debunking base ($n = 109$), with autonomy support ($n = 104$), with explicit details ($n = 115$), and combined strategies ($n = 103$). The last group received a message control ($n = 121$), seeing only the astroturf attack video. While not a true control, this group served as the referent group to compare the differences between an organization choosing, or not choosing, to respond to an astroturf attack. After viewing the stimuli, all participants were asked to recall the organization allegedly spreading false information (Humane Watch) and who the (pre/de)bunking message was from (HSUS). Next, participants received dependent measures displayed in the form of an online questionnaire. Lastly, demographics were collected, and a randomized code was generated for participants to receive payment.

### 3.4. Pre-test

Prior to launching the study, a pre-test ($N = 157$) was conducted via MTurk to ensure that the manipulations to inoculation and autonomy support mechanisms had the desired effect on messages published by HSUS. Following the procedure set forth by Banas and Richards (2017) motivational threat was measured as a key mechanism of inoculation. Participants were randomized into two groups, one group who received an inoculation message from HSUS and one group who received a supportive message from HSUS, which acted as a control message. Following previous recommendations, the supportive message solely focused on bolstering information about HSUS, while not mentioning Humane Watch or a forthcoming astroturf attack (McGuire, 1961a). After reading the assigned message from HSUS, individuals received four items on a seven-point Likert-type scale to measure their perceived motivational threat (Banas & Richards, 2017; Richards & Banas, 2018). An independent samples $t$-test found significant differences between participants who received the message containing inoculation and those who received the control supportive message, $t(1, 72) = -3.65, p < .001$, two-tailed. Participants exposed to the inoculation message experienced more motivational threat after reading the message than participants exposed to the control message. Therefore, the manipulation of the inoculation prebunking strategy from HSUS was deemed successful.

To analyze if the manipulations for autonomy support were successful, perceived autonomy support was also measured in the pre-test. Guided by the self-determination theory, the manipulation check asked, "While reading the message, I felt like HSUS provided me with choices

and options on what to believe and support" (Sheldon & Filak, 2008). An independent samples *t*-test found significant differences between the prebunking message containing autonomy support and the control message, *t*(1, 62) = -2.13, *p* < .04, two-tailed. Participants exposed to the autonomy-supportive message perceived it as significantly more supportive than those who received the inoculation only message, which did not contain autonomy supportive messaging. Therefore, the manipulation was deemed successful.

### 3.5. Dependent measures

All items used for each dependent variable were adopted from previous research. Due to the sample size (*N* = 965), a confirmatory factor analysis (CFA) was conducted with the lavaan package (Rosseel, 2012) for R to ensure the validity of each multi-item measurement. Model fit for the measurement models were adequate based on the criteria from Hu and Bentler (1999)[1] . In addition, Cronbach's alpha coefficients ranged from .84 to .96, surpassing the acceptable threshold of .70 (Nunnally, 1978).

### 3.6. Attitudinal feeling thermometer

Participants were asked to evaluate a series of organizations, including HSUS and Humane Watch, on a 100-point feeling thermometer (0 = dislike a great deal, 100 = like a great deal). If individuals were unfamiliar with the organization, they were instructed to select "50 = neither dislike/like." A feeling thermometer is commonly used in political communication experimental design literature to avoid priming participants within the experiment (Warner et al., 2020; Warner & McKinney, 2013; Zavala-Rojas, 2014). (Pre-crisis HSUS: *M* = 76.78, *SD* = 20.66; Pre-crisis Humane Watch: *M* = 61.64, *SD* = 20.74, Post-crisis HSUS: *M* = 68.29, *SD* = 26.03; Post-crisis Humane Watch: *M* = 46.21, *SD* = 29.43).

### 3.7. Post-crisis organizational reputation

Using an adaptation of Coombs and Holladay's (1996, 2002) organizational reputation scale, participants were asked to evaluate their perception of the organization's reputation post-crisis. The five-item scale included statements such as, "The organization is concerned with the well-being of its publics" and "Under most circumstances, I would be likely to believe what the organization says." Participants responded on a seven-point Likert scale ranging from (1) strongly disagree to (7) strongly agree (*M* = 4.58, *SD* = 6.02, *α* = .84).

### 3.8. Perceived source credibility

Participants were asked about their perceived credibility towards HSUS and Humane Watch. The credibility scale was adapted from McCroskey and Teven (1999) and contained a series of 12 statements with bipolar answers including, "I perceive this organization to be trustworthy/untrustworthy, intelligent/unintelligent, honest/dishonest" (Averaged scale for HSUS: *M* =

---

[1] Final CFA model goodness-of-fit for organizational reputation (5-item scale): $\chi^2(3) = 13.955$, p < .001, robust root mean square error of approximation (rRMSEA) = .072 (.037-.112), robust comparative fit index (rCFI) = .994, robust non-normed fit index/Tucker Lewis index (rNNFI/TLI) = .981, standardized root mean residual (SRMR) = .015; perceived credibility (12-item scale): $\chi^2(32) = 55.587$, p < .01, robust root mean square error of approximation (rRMSEA) = .040 (.022-.058), robust comparative fit index (rCFI) = .995, robust non-normed fit index/Tucker Lewis index (rNNFI/TLI) = .991, standardized root mean residual (SRMR) = .027; social amplification (3-item scale): $\chi^2(0) = 0$, p < .001, robust root mean square error of approximation (rRMSEA) = .000 (.000-.000), robust comparative fit index (rCFI) = 1.00, robust non-normed fit index/Tucker Lewis index (rNNFI/TLI) = 1.00, standardized root mean residual (SRMR) = .000.

**Table 1**
ANOVA Comparative Summary of Effects of the Proactive Disinformation Communication Strategies on Attitudinal Change towards Attacked Organization.

| Condition | n | Avg. Decrease in Attitude M (SD) |
|---|---|---|
| Strategic silence | 121 | 23.20 (*35.84*)abcd |
| Prebunking only | 106 | 10.44 (*23.43*)a |
| + Autonomy support | 107 | 4.85 (*19.83*)b |
| + Explicit detail | 102 | 7.94 (*24.74*)c |
| + Combined strategy | 98 | 1.27 (*21.50*)d |

Shared subscripts represent statistically significant differences using Fisher's LSD: b,c,d = *p* < .001, a = *p* < .01.
Represents average decrease in attitude from pre-test to post-test, 100-point scale.

**Table 2**
ANCOVA Comparative Summary of Effects of the Proactive Disinformation Communication Strategies on Credibility and Reputation.

| Condition | n | Attacked Org. Reputation M (se) | Attacked Org. Cred. M (se) | Attacking Org. Cred. M (se) | Social Amplification M (se) |
|---|---|---|---|---|---|
| Strategic silence | 121 | 3.88 (*.10*)abcd | 3.88 (*.11*)fghi | 4.38 (*.12*)klmn | 3.84 (*.16*)pqrs |
| Prebunking only | 106 | 4.50 (*.11*)ae | 4.62 (*.12*)fj | 3.68 (*.12*)k | 2.71 (*.17*)p |
| + Autonomy support | 107 | 4.79 (*.12*)b | 4.91 (*.12*)g | 3.74 (*.12*)lo | 2.52 (*.17*)q |
| + Explicit details | 102 | 4.78 (*.11*)c | 4.86 (*.12*)h | 3.27 (*.13*)m | 2.54 (*.18*)r |
| + Combined strategy | 98 | 5.00 (*.12*)de | 5.09 (*.12*)ij | 3.28 (*.13*)no | 2.48 (*.18*)s |

Shared subscripts represent statistically significant differences using Fisher's LSD: a,b,c,d,f,g,h,i,k,l,m,n,o,p,q,r,s = *p* <= .001, j,o = *p* <= .01.

4.69, *SD* = 15.84, *α* = .95; Averaged scale for Humane Watch: *M* = 3.74, *SD* = 16.40, *α* = .95).

### 3.9. Social amplification

This study asked participants how likely they are to engage with the attack video via Facebook (i.e., share, positively react, comment) (Barger et al., 2016). This measurement helped in determining if proactive messaging could help deter the spread of false information. This three-item scale was evaluated on a 7-point Likert scale ranging from (1) very unlikely to (7) very likely (*M* = 2.87, *SD* = 5.92, *α* = .93).

## 4. Results

### 4.1. Attitudinal change towards the attacked organization

Research question 1 asked which proactive communication strategy would cause the least amount of attitudinal change between pre-and post-astroturf video viewings. To answer this question, individuals received a 100-point attitudinal scale before entering the experiment and after viewing the stimuli. An ANOVA found that there were significant differences between the message strategies used and attitudinal change experienced towards HSUS, *F*(4, 529) = 11.70, *p* < .001, Power = 1.00, $\eta^2$ = .08 (Table 1). It was found that individuals who did not receive a prebunking message forewarning about the attack experienced an average decrease of 23.20 points on their views towards HSUS after watching the astroturf attack video. It was found that individuals who received a prebunking message which included autonomy support or explicit details on average, had less than a 7.94-point decrease on a 100-point scale in attitudinal change post-attack. Individuals who received a prebunking message, which included both autonomy support and explicit details, only experienced a 1.27 change in attitude.

## 4.2. Perceived organizational reputation of the attacked organization

To examine research question 2, an ANCOVA controlling for pre-attitudes was used to determine if there were significant differences between prebunking strategies used regarding the attacked organization's reputation compared to those who only received the attack video (control). It was found that there were statistically significant differences between the conditions, $F(4, 527) = 18.34$, $p < .001$, Power = 1.00, $\eta^2 = .12$ (Table 2). Pairwise comparisons indicated that individuals who did not receive a prebunking message prior to the attack (control message) had significantly lower perceptions of HSUS' reputation after viewing the attack video ($M = 3.88$, $se = .10$) than all four prebunking treatments: prebunking only ($M = 4.50$, $se = .11$, $p < .001$), with autonomy support ($M = 4.79$, $se = .12$, $p < .001$), with explicit details ($M = 4.78$, $se = .11$, $p < .001$), and combined strategy ($M = 5.00$, $se = .12$, $p < .001$). Furthermore, it was found that individuals who received the prebunking message using the combined strategy of both autonomy support and explicit details perceived significantly higher levels of organizational reputation post-attack than individuals who received the prebunking only strategy ($p = .001$).

## 4.3. Perceived credibility of attacked organization

To examine research question 3, the perceived source credibility of the organization being attacked (HSUS) by the astroturf campaign was examined. An ANCOVA found that there were significant differences between the message treatments and control, $F(4, 527) = 17.17$, $p < .001$, Power = 1.00, $\eta^2 = .12$ (Table 2). Pairwise comparisons found that individuals who did not receive a prebunking message (control) had significantly lower perceptions of HSUS' credibility than all four prebunking treatments ($M = 3.88$, $se = .11$, $p < .001$). It was also found that individuals who received the prebunking combined strategy reported significantly higher perceived source credibility ($M = 5.09$, $se = .12$) than those who received prebunking only ($p = .01$).

## 4.4. Perceived credibility of the attacking organization

Research question 4 examined if perceived source credibility of the attacking organization differed between prebunking strategies used compared to those who only received the attack video (control). An ANCOVA found that there were significant differences between the message treatments and control, $F(4, 527) = 13.96$, $p < .001$, Power = 1.00, $\eta^2 = .10$ (Table 2). Pairwise comparisons found that individuals who did not receive a prebunking message had significantly lower perceptions of Humane Watch's credibility ($M = 4.38$, $se = .12$) than individuals who received prebunking only ($M = 3.68$, $se = .12$, $p < .001$), with autonomy support ($M = 3.74$, $se = .12$, $p = .002$), with explicit details ($M = 3.27$, $se = .13$, $p < .001$), and combined strategy ($M = 3.28$, $se = .13$, $p < .001$). Meaning that individuals that did not receive a prebunking message from HSUS were more likely to view Humane Watch as being more credible than those who received one of the four prebunking messages. Further, individuals who received the prebunking combined strategy reported significantly lower amounts of perceived credibility towards the attacking organization than those who received prebunking with only autonomy support ($M_{diff} = -.46$, $p = .01$).

## 4.5. Amplification of disinformation

Research question 5 examined the ability of prebunking messaging to thwart individuals from amplifying the disinformation through 'positively reacting,' 'sharing,' and 'posting' on an astroturf video posted to Facebook compared to those who did not receive a prebunking message. An ANCOVA controlling for pre-attitudes towards the organizations found that there were significant differences between the conditions, $F(4, 527) = 15.00$, $p < .001$, Power = 1.00, $\eta^2 = .09$ (Table 2). Pairwise comparisons showed that individuals who received strategic silence

reported having greater intentions to amplify the astroturf video ($M = 3.84$, $se = .16$) than those who received prebunking only ($M = 2.71$, $se = .17$, $p < .001$), with autonomy support ($M = 2.52$, $se = .17$, $p < .001$), with explicit details ($M = 2.54$, $se = .18$, $p < .001$), and with combined strategies ($M = 2.48$, $se = .18$, $p < .001$). There were no significant differences found between the four prebunking messages.

## 4.6. Prebunking vs. debunking

Lastly, research question 6 asked how the effects of a refutation response differ within prebunking and debunking on crisis outcomes (i. e., credibility, organizational reputation, attitudes, and social amplification). A series of independent t-tests were performed comparing the prebunking condition to the debunking condition. There were no significant differences found on attitude change ($p = .80$), perceived credibility towards Humane Watch ($p = .19$), or perceived credibility towards HSUS ($p = .20$). However, there were statistically significant difference between the conditions when examining HSUS's perceived organizational reputation, $t(1, 842) = 1.93$, $p = .05$, $d = 13$. This result indicates that individuals who received a prebunking message ($M = 4.77$, $SD = 1.21$) perceived HSUS to have a higher reputation post-attack than those who received the debunking message ($M = 4.61$, $SD$ 1.17). In addition, there were statistically significant differences for intention to amplify the astroturf video, $t(1, 841.56) = -2.56$, $p = .01$, $d = .17$. The result suggests individuals who received a prebunking message ($M = 2.56$, $SD = 1.85$) have less intention to amplify the message than those who received a debunking message ($M = 2.89$, $SD = 1.98$).

## 4.7. Discussion

The primary goal of this research endeavor was to explore the plausibility and effectiveness of organizations using proactive communication by inoculating individuals against astroturf attacks through prebunking. The use of prebunking statistically outperformed debunking and strategic silence when looking at minimizing damage to organizational reputation and thwarting social amplification. These findings provide initial evidence that prebunking could be used as a stand-alone strategy to combat astroturf attacks. Furthermore, the additive effect of manipulating language intensity by using autonomy support, along with source derogation by using explicit details, indicates a promising outlook for organizations and for decreasing the spread of disinformation.

Findings from this study provide a better understanding of what proactive strategies could be used to thwart the effects of disinformation astroturf attempts to protect crisis outcomes such as source credibility, organizational reputation, and overall attitudes towards an organization. Prebunking messages using a solo or combined strategy with autonomy support and explicit details consistently outperformed not responding to the attack (i.e., strategic silence). These tentative findings support interweaving components of the self-determination theory, such as autonomy support, along with inoculation, into public relations and crisis communication strategies. This supports findings within self-determination theory in that individuals prefer autonomy supportive language, which implies the individual has the right to decide what to believe (Mouratidis et al., 2010; Vansteenkiste et al., 2004). Parallel to Mouratidis et al.'s (2010) findings of athletes preferring to receive autonomy supportive corrective feedback, the use of autonomy support when dealing with overcoming disinformation consistently provided higher outcomes. In addition, using source derogation by providing explicit details about the attacking organization and the intentions of the attack could be of benefit.

The preliminary empirical evidence from this study suggests that organizations should proactively communicate by using prebunking strategies when an astroturf attack can be anticipated, providing support against previous literature proposing that no response (i.e., strategic silence) may be warranted when faced with a paracrisis (Coombs, 2019). The current study found that while there were no consistent statistically

significant differences present for the differing prebunking strategies of autonomy support and explicit details, the mean scores for the combined strategy outperformed solo usage for all variables.

### 4.8. Theoretical implications

The first theoretical contribution of this study was the examination of how both message strategies to combat astroturf disinformation and the attack itself can influence perceived source credibility. As a conceptual framework, source credibility refers to the believability that an individual receiving a message assigns to the source's expertise and trustworthiness. This study found that prebunking disinformation, using inoculation, decreased the perceived credibility towards the attacking organization (Humane Watch). The organization was perceived as more credible when an individual did not receive any type of message regarding the disinformation through strategic silence (no response from HSUS). Findings suggest it is possible to encourage stakeholders to discredit the attacking organization by lowering their credibility through messaging. While it was found that individuals who received the prebunking treatment perceived Humane Watch to be the least credible, using autonomy support could amplify these positive effects, as both the solo use and combined strategy had the lowest credibility scores. Similarly, the same strategies that worked to lower Humane Watch's perceived source credibility increased that of the organization being attacked (HSUS). Overall, it was found that using prebunking strategies outperformed the effects of strategic silence. In particular, individuals who received the prebunking combined strategy, infused with autonomy support and explicit details, perceived HSUS as having the highest credibility.

One of the most examined variables within crisis communication literature is organizational reputation. Having a positive reputation can reduce stakeholder uncertainty about organizational performance, motivate consumers to buy products, attract high-quality employees, encourage outside investors, and retain essential transaction partners (Gardberg & Fombrun, 2006). This study provides a clearer picture of how paracrises can threaten organizational reputation and what responses should be considered. It was found that prebunking outperformed the use of strategic silence and debunking in its ability to protect organizational reputation. Looking further into the data for potential ways organizations can mitigate damage to reputation, results showed that using the combined factors of autonomy support and explicit details could be a strong strategy, as it produced the highest average mean for organizational reputation. As reputation is a dominant dependent variable in scholarship and a valued outcome in practice, this outcome provides insight into how reputation is impacted in these emerging crises.

Expanding out from looking at organizational-based outcomes, the spread of disinformation on social media is a systemic societal concern. This study approached this concern through the lens of an organization minimizing the spread of an astroturf disinformation video on Facebook. This research contends that the potential for social amplification is a valuable variable to include when examining message strategy impact on post-crisis reputation. By using the dimensions of the social amplification scale present in this study, scholars and PR professionals could potentially anticipate the reactions of stakeholders on social media and the threat of further perpetuating a crisis. Measuring the likelihood of interacting with crisis information online is a specific behavioral intention that can travel exponentially faster than negative word of mouth. Interacting with a social media post (e.g., liking, commenting, and sharing) increases the relevance of that topic for other users, further disseminating the message (Strekalova, 2017). When a crisis response is amplified, it reaches a greater audience, and the potential for negative perceptions to perpetuate is extended. This implies that the platform should be recognized as a tool that can deescalate the interactions surrounding disinformation.

Findings from this study show that the way forewarning content is framed can influence individuals' intentions to engage with disinformation-based content on Facebook through positive engagement. A particularly exciting finding from this study is initial confirmation that prebunking can possibly assist in preventing the spread of disinformation on social media platforms. Those who did not receive a prebunking message reported significantly higher intentions to partake in amplifying or retransmitting the disinformation on Facebook. This provides evidence that alerting individuals to the idea of disinformation being spread in advance can help deter further transmission of the message.

### 4.9. Practical implications

Beyond the theoretical points made above, this study is applicable for communication professionals responsible for protecting the reputation and reducing harm to an organization's brand. Initial findings demonstrate that the steps taken before a disinformation attack occurs in communicating can help thwart the extensions of the crisis from it spreading on social media. The empirically-tested strategies of using autonomy support and explicit details provide PR professionals with actionable steps to create content that can work to negate the effects of disinformation. Providing details about who the attacker is and what the motive is for spreading disinformation, along with offering meaningful choices and acknowledging potential anxieties, are all message strategies that organizations can execute.

Regarding responding to a paracrisis post-attack, findings suggest that organizations should respond to disinformation paracrisis attacks rather than ignoring the attack. Using prebunking significantly improved crisis outcomes compared to using strategic silence (no response). Within this experiment, participants who did not receive a message from HSUS experienced an increase in negative crisis outcomes, including lower perceived attitudes towards HSUS, lower perceived credibility, lower organizational reputation, and more intention to amplify the astroturf Facebook video. This is an important finding as paracrisis literature suggests that responding to paracrisis situations is optional. However, it appears that when the scenario involves disinformation regarding use of funds or an organization's intent, a response is warranted.

### 4.10. Limitations and future research direction

The current study has several limitations that can be addressed in future studies. A primary limitation of this study is that while experimental design has many benefits, it is still conducted in an experimental setting. As such, the lifecycle of this crisis played out in a 20-minute capsule, which does not represent the actual timescale in which crises occur. For example, the astroturf attack followed immediately after the forewarning message from HSUS. This inhibits this study's ecological validity since, in reality, individuals would most likely not receive a threat immediately following the warning. It is hypothesized that the lack of time-lapse between presentation of the stimuli could have led to the lack of significant differences between the (pre/de)bunking conditions. Within this same vein, this study only examined the individual effects of (pre/de)bunking administered at a single point. Future research should further examine the additive or "boosting" effect found within the principle of congruity and inoculation literature when using (pre/de)bunking (e.g., McGuire, 1961b; Pfau et al., 2004; Tannenbaum & Norris, 1965). Moreover, disinformation and astroturf campaigns influence various industries, and it is crucial to further test if the strategies that successfully thwarted adverse effects for HSUS translate into other fields and other cultures.

In addition, future research is needed to bolster the assumptions and claims within this study for generalizability, including using a true control to increase internal and external validity. This line of research could be strengthened further by including several measurements such as motivation to process the experimental messages (e.g., Maheswaran & Chaiken, 1991) and an in-depth attitudinal scale (e.g., Burgoon et al.,

1978; Pfau et al., 2007). Lastly, two lines of promising future tasks based on previously conducted persuasion research include examining the complex role of involvement of public groupings and further developing a continuum for message intensity and control presented within (pre/de)bunking messages.

### 4.11. Conclusion

This research endeavor examined how persuasive attempts from astroturf attacks, a specific form of disinformation, can be counteracted to protect organizational outcomes (i.e., attitudes, organizational reputation, credibility), along with thwarting the amplification of the astroturf messages on social media. The present study's main contribution is showcasing the effectiveness of proactive messaging, specifically the utility of using prebunking to prepare for an astroturf-based attack. Even more promising was the initial look at language intensity in combining the use of explicit details and autonomy support within prebunking messages. This initial assessment of prebunking illustrates a promising strategy to effectively expose the nature of astroturf campaigns, working to preserve an organization's reputation and minimize the spread of disinformation.

### Declaration of Competing Interest

I do not have any interests that might be interpreted as influencing the research, and APA ethical standards were adhered to in the research process. This manuscript has not been previously published, nor is it under consideration for publication elsewhere.

### Appendix A. Participant demographics

**Table A1**
Participant Demographics.

| | n (%) |
|---|---|
| Gender | |
| Female | 483 (50.1 %) |
| Male | 474 (49.1 %) |
| Intersex | 1 (.1 %) |
| Prefer not to say | 7 (.7 %) |
| Race | |
| White | 730 (75.6 %) |
| Black/African American | 138 (14.3 %) |
| Asian | 62 (6.4 %) |
| Other | 24 (2.5 %) |
| American Indian/Alaska Native | 8 (.8 %) |
| Native Hawaiian/Pacific Islander | 3 (.3 %) |
| Ethnicity | |
| Spanish/Hispanic/Latino/Other | 63 (6.5 %) |
| Mexican/Mexican American/Chicano | 36 (3.7 %) |
| Puerto Rican | 11 (1.1 %) |
| Cuban | 10 (1.0 %) |
| Prefer not to say | 29 (3.0 %) |
| Age (range) | 38.60 (19−78) |
| Highest Level of Education | |
| Less than high school degree | 4 (.4 %) |
| High school diploma or GED | 80 (8.3 %) |
| Some college, no degree | 157 (16.3 %) |
| Associate degree or equivalent | 105 (10.9 %) |
| Bachelor's degree | 443 (45.9 %) |
| Graduate or professional degree | 176 (18.2 %) |
| Household Income | |
| ~ $30,000 | 165 (17.1 %) |
| $30,000 - $59,999 | 368 (38.1 %) |
| $60,000 - $99,999 | 302 (31.3 %) |
| $100,000 ~ | 130 (13.5 %) |

*Note.* Cells display counts of each category with percentages in parentheses in the overall column, except for cells referred to *age* with the range in parentheses.

### Appendix B. Stimuli exemplars

**Table B1**
Exemplars of pre/debunking stimuli.

| Condition | Exemplars |
|---|---|
| Full (pre/de)bunking base message | "At the Humane Society of the United States (HSUS), we sometimes get attacked by groups like Humane Watch that do not like the way we do things. The attack video being shared online by Humane Watch tries to persuade viewers that HSUS misuses the funds it collects. That we use the money for ourselves, instead of spending it on the animals. For example, you may hear that we do not use funds to support local shelters. Humane Watch wants you to believe that HSUS misleads the public on how we spend funds. That HSUS does not have the well-being of animals as our top priority. What this claim doesn't tell you is that in 2018 alone, HSUS spent $139.1 million on animal protection programs. This money provides emergency care and homes to more animals than any other organization in the United States. Since 1954, HSUS has taken on puppy mills, factory farms, trophy hunts, animal testing, and other cruel industries. With our partners, we rescue and care for thousands of animals every year. As you reflect on your beliefs about HSUS and our dedication to animal protection, remember our current wins and our fight for animal rights. Regardless of the false information Humane Watch provides, no organization in the world is more dedicated to protecting animals than we are." |
| +Autonomy support | "…Although we hope to convince you that the attack is not accurate or legitimate, we realize that you have the right to decide this for yourself. It's up to you, not us!…" |
| +Explicit details | "…The Humane Watch video being spread online tells viewers that HSUS misuses its funds. It states that we use the money for ourselves, instead of spending it on animal welfare. For example, you may hear that we do not use funds to support local shelters. Humane Watch wants you to believe that HSUS misleads the public on how we spend funds…" |
| +Combined strategy | "…Sometimes at the Humane Society of the United States (HSUS), we get attacked by groups like Humane Watch. These groups spread misleading information in an attempt to defund us. We want you to know who Humane Watch is, so you can decide for yourself on what to believe…" |

### References

Banas, J. A., & Miller, G. (2013). Inducing resistance to conspiracy theory propaganda: Testing inoculation and metainoculation strategies. *Human Communication Research, 39*(2), 184–207. https://doi.org/10.1111/hcre.12000.

Banas, J. A., & Rains, S. A. (2010). A meta-analysis of research on inoculation theory. *Communication Monographs, 77*(3), 281–311. https://doi.org/10.1080/03637751003758193.

Banas, J. A., & Richards, A. S. (2017). Apprehension or motivation to defend attitudes? Exploring the underlying threat mechanism in inoculation-induced resistance to persuasion. *Communication Monographs, 84*(2), 164–178. https://doi.org/10.1080/03637751.2017.1307999.

Barger, V., Peltier, J. W., & Schultz, D. E. (2016). Social media and consumer engagement: A review and research agenda. *Journal of Research in Interactive Marketing, 10*(4), 268–287. https://doi.org/10.1108/JRIM-06-2016-0065.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: amazon.com's mechanical turk. *Political Analysis, 20*(3), 351–368. https://doi.org/10.1093/pan/mpr057.

Boman, C. D., & Schneider, E. J. (2021). Finding an antidote: Testing the use of proactive crisis strategies to protect organizations from astroturf attacks. *Public Relations Review, 47*(1), Article 102004. https://doi.org/10.1016/j.pubrev.2020.102004.

Brehm, J. (1966). *A theory of psychological reactance*. Academic Press.

Breland, J. Y., Quintiliani, L. M., Schneider, K. L., May, C. N., & Pagoto, S. (2017). Social media as a tool to increase the impact of public health research. *American Journal of Public Health, 107*(12), 1890–1891. https://doi.org/10.2105/AJPH.2017.304098.

Burgoon, M. (1976). Propensity of persuasive attack and intensity of pretreatment messages as predictors of resistance to persuasion. *The Journal of Psychology, 92*(1). https://doi.org/10.1080/00223980.1976.9921345.

Burgoon, M., & King, L. B. (1974). The mediation of resistance to persuasion strategies by language variables and active-passive participation. *Human Communication Research, 1*(1), 30–41. https://doi.org/10.1111/J.1468-2958.1974.TB00251.X.

Burgoon, M., & Miller, G. R. (1971). Prior attitude and language intensity as predictors of message style and attitude change following counterattitudinal advocacy. *Journal of Personality and Social Psychology, 20*(2). https://doi.org/10.1037/h0031678.

Burgoon, M., Jones, S. B., & Stewart, D. (1975). Toward a message-centered theory of persuasion: Three empirical investigations of language intensity. *Human Communication Research, 1*(3), 240–256. https://doi.org/10.1111/J.1468-2958.1975.TB00271.X.

Burgoon, M., Miller, M. D., Cohen, M., & Montgomery, C. L. (1978). An empirical test of a model of resistance to persuasion. *Human Communication Research, 5*(1). https://doi.org/10.1111/j.1468-2958.1978.tb00620.x.

Cho, C. H., Martens, M. L., Kim, H., & Rodrigue, M. (2011). Astroturfing global warming: It isn't always greener on the other side of the fence. *Journal of Business Ethics, 104*(4), 571.

Chong, M., & Choy, M. (2018). The social amplification of haze-related risks on the internet. *Health Communication, 33*(1), 14–21. https://doi.org/10.1080/10410236.2016.1242031.

Compton, J. A. (2017). Image prepare: Image repair, inoculation theory, and anticipated attacks on credibility. *The International Journal of the Image, 8*(1), 1–9. https://doi.org/10.18848/2154-8560/CGP/v08i01/1-9.

Compton, J. A., & Pfau, M. (2008). Inoculating against pro-plagiarism justifications: Rational and affective strategies. *Journal of Applied Communication Research, 36*(1), 98–119. https://doi.org/10.1080/00909880701799329.

Coombs, T. (2019). *Ongoing crisis communication: Planning, managing, and responding* (5th ed.). SAGE Publications, Inc.

Coombs, T., & Holladay, S. (1996). Communication and attributions in a crisis: An experimental study in crisis communication. *Journal of Public Relations Research, 8*(4), 279–295. https://doi.org/10.1207/s1532754xjprr0804_04.

Coombs, T., & Holladay, S. (2002). Helping crisis managers protect reputational assets. *Management Communication Quarterly, 16*(2), 165–186. https://doi.org/10.1177/089331802237233.

Coombs, T., & Holladay, S. (2012). The paracrisis: The challenges created by publicly managing crisis prevention. *Public Relations Review, 38*(3), 408–415. https://doi.org/10.1016/j.pubrev.2012.04.004.

Dainton, M., & Zelley, E. D. (2004). Explaining theories of persuasion. *Applying communication theory for professional life* (pp. 103–131).

de Meulenaer, S., de Pelsmacker, P., & Dens, N. (2018). Power distance, uncertainty avoidance, and the effects of source credibility on health risk message compliance. *Health Communication, 33*(3), 291–298. https://doi.org/10.1080/10410236.2016.1266573.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.

Dillingham, L. L., & Ivanov, B. (2017). Inoculation messages as a preemptive financial crisis communication strategy with inexperienced investors. *Journal of Applied Communication Research, 45*(3), 274–293. https://doi.org/10.1080/00909882.2017.1320571.

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Eagly, A., Wood, W., & Chaiken, S. (1978). Causal inferences about communicators and their effect on opinion change. *Journal of Personality and Social Psychology, 36*(4), 424–435. https://doi.org/10.1037/0022-3514.36.4.424.

Earle, T. C., & Cvetkovich, G. (1995). *Social trust: Toward a cosmopolitan society*. Greenwood Publishing Group.

Einwiller, S. A., & Johar, G.v. (2013). Countering accusations with inoculation: The moderating role of consumer-company identification. *Public Relations Review, 39*(3), 198–206. https://doi.org/10.1016/j.pubrev.2013.03.002.

Festinger, L. (1957). *A theory of cognitive dissonance* (Vol. 2). Stanford University Press.

Fombrun, C. J., & van Riel, C. B. M. (1997). The reputational landscape. *Corporate Reputation Review, 1*(2), 5–13. https://doi.org/10.1057/palgrave.crr.1540024.

Freberg, K. (2012). Intention to comply with crisis messages communicated via social media. *Public Relations Review, 38*(3), 416–421. https://doi.org/10.1016/j.pubrev.2012.01.008.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior, 26*(4). https://doi.org/10.1002/job.322.

Gardberg, N. A., & Fombrun, C. J. (2006). Corporate citizenship: Creating intangible assets across institutional environments. *The Academy of Management Review, 31*(2), 329–346. https://doi.org/10.5465/AMR.2006.20208684.

Gittelman, S., Lange, V., Crawford, C. A. G., Okoro, C. A., Lieb, E., Dhingra, S. S., & Trimarchi, E. (2015). A new source of data for public health surveillance: Facebook likes. *Journal of Medical Internet Research, 17*(4). https://doi.org/10.2196/jmir.3970.

Honisch, S. V., & Más Manchón, L. (2019). The effects of paracrisis origin and response strategy on Facebook audience's perceived organisational reputation and behavioural intentions. *Corporate Reputation Review, 23*(3), 133–144. https://doi.org/10.1057/s41299-019-00070-4.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118.

Ingram, D. (2020). *Pre-bunks*. NBC News. https://www.nbcnews.com/tech/tech-news/twitter-launches-pre-bunks-get-ahead-voting-misinformation-n1244777.

Ivanov, B., Burns, W. J., Sellnow, T. L., Petrun Sayers, E. L., Veil, S. R., & Mayorga, M. W. (2016). Using an inoculation message approach to promote public confidence in protective agencies. *Journal of Applied Communication Research, 44*(4), 381–398. https://doi.org/10.1080/00909882.2016.1225165.

Ivanov, B., Miller, C. H., Compton, J. A., Averbeck, J. M., Harrison, K. J., Sims, J. D., … Parker, J. L. (2012). The attitude base as a moderator of the effectiveness of inoculation strategy. *The Journal of Communication, 62*(4), 701–718. https://doi.org/10.1111/j.1460-2466.2012.01658.x.

Ivanov, B., Sellnow, T., Getchell, M., & Burns, W. (2018). The potential for inoculation messages and postinoculation talk to minimize the social impact of politically motivated acts of violence. *Journal of Contingencies and Crisis Management, 26*, 414–424. https://doi.org/10.1111/1468-5973.12213.

Jackson, B., Whiddett, R., Anthony, D. R., Dimmock, J. A., & Compton, J. A. (2015). Preempting performance challenges: The effects of inoculation messaging on attacks to task self-efficacy. *PloS One, 10*(4). https://doi.org/10.1371/journal.pone.0124886.

Jin, Y., Liu, B. F., & Austin, L. L. (2014). Examining the role of social media in effective crisis management: The effects of crisis origin, information form, and source on publics' crisis responses. *Communication Research, 41*(1), 74–94.

Jin, Y., van der Meer, T. G. L. A., Lee, Y. I., & Lu, X. (2020). The effects of corrective communication and employee backup on the effectiveness of fighting crisis misinformation. *Public Relations Review, 46*(3), Article 101910. https://doi.org/10.1016/j.pubrev.2020.101910.

Kasperson, R. E., & Kasperson, J. X. (1996). The social amplification and attenuation of risk. *The Annals of the American Academy of Political and Social Science, 545*, 95.

Keller, F. B., Schoch, D., Stier, S., & Yang, J. H. (2020). Political astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication, 37*(2), 256–280. https://doi.org/10.1080/10584609.2019.1661888.

Kincaid, J., Fishburne, R., Rogers, R., & Chissom, B. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Naval technical training*. U. S. Naval Air Station.

Koestner, R., & Losier, G. F. (1996). Distinguishing reactive versus reflective AutonomY. *Jourrml of Persormlity, 64*(2).

Lin, T., Hwang, L., & Lai, Y. (2017). Effects of argument quality, source credibility and self-reported diabetes knowledge on message attitudes: An experiment using diabetes related messages. *Health Information & Libraries Journal, 34*(3), 225–235. https://doi.org/10.0.4.87/hir.12181.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology, 61*(1). https://doi.org/10.1037/0022-3514.61.1.13.

Malka, A., Krosnick, J. A., & Langer, G. (2009). The association of knowledge with concern about global warming: Trusted information sources shape public thinking. *Risk Analysis, 29*(5), 633–647. https://doi.org/10.1111/j.1539-6924.2009.01220.x.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs, 66*(1), 90–103. https://doi.org/10.1080/03637759909376464.

McGuire, W. J. (1961a). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. *Journal of Abnormal & Social Psychology, 63*(2), 326. https://doi.org/10.1037/h0048344.

McGuire, W. J. (1961b). The effectiveness of supportive and refutational defenses in immunizing and restoring beliefs against persuasion. *Sociometry, 24*(2). https://doi.org/10.2307/2786067.

McGuire, W. J., & Papageorgis, D. (1961). The relative efficacy of various types of prior belief-defense in producing immunity against persuasion. *Journal of Abnormal and Social Psychology, 62*(2), 327–337. https://doi.org/10.1037/h0042026.

Miller, M. D., & Burgoon, M. (1979). The relationship between violations of expectations and the induction of resistance to persuasion. *Human Communication Research, 5*(4). https://doi.org/10.1111/j.1468-2958.1979.tb00642.x.

Miller, C. H., Ivanov, B., Sims, J., Compton, J. A., Harrison, K. J., Parker, K. A., Parker, J. L., & Averbeck, J. M. (2013). Boosting the potency of resistance: Combining the motivational forces of inoculation and psychological reactance. *Human Communication Research, 39*(1), 127–155. https://doi.org/10.1111/J.1468-2958.2012.01438.X.

Miller, C. H., Lane, L. T., Deatrick, L. M., Young, A. M., & Potts, K. A. (2007). Psychological reactance and promotional health messages: The effects of controlling language, lexical concreteness, and the restoration of freedom. *Human Communication Research, 33*(2), 219–240. https://doi.org/10.1111/J.1468-2958.2007.00297.X.

Miller, C. H., Lane, L. T., Deatrick, L. M., Young, A. M., & Potts, K. A. (2017). Psychological reactance and promotional health messages: The effects of controlling language, lexical concreteness, and the restoration of freedom. *Human Communication Research, 33*(2), 219–240. https://doi.org/10.1111/j.1468-2958.2007.00297.x.

Mouratidis, A., Lens, W., & Vansteenkiste, M. (2010). How you provide corrective feedback makes a difference: The motivating role of communicating in an autonomy-supporting way. *Journal of Sport & Exercise Psychology, 32*(5), 619. https://doi.org/10.1123/jsep.32.5.619.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

O'Keefe, D. J. (1990). *Persuasion: Theory and research*. Sage Publications, Inc.

O'Keefe, D. J. (2009). Theories of persuasion. In R. L. Nabi, & M. B. Oliver (Eds.), *The SAGE handbook of media processes and effects* (pp. 269–282). Sage Publications, Inc.

Parker, K. A., Rains, S. A., & Ivanov, B. (2016). Examining the "blanket of protection" conferred by inoculation: The effects of inoculation messages on the cross-protection of related attitudes. *Communication Monographs, 83*(1), 49–68. https://doi.org/10.1080/03637751.2015.1030681.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Communication and persuasion* (pp. 1–22). New York: Springer. https://doi.org/10.1007/978-1-4612-4964-1_1.

Petty, R. E., Cacioppo, J. T., & Haugtvedt, C. P. (1992). Ego-involvement and persuasion: An appreciative look at the Sherifs' contribution to the study of self-relevance and attitude change. *Social judgment and intergroup relations* (pp. 147–174).

Pfau, M., Compton, J. A., Parker, K. A., Wittenberg, E. M., An, C., Ferguson, M., Horton, H., & Malyshev, Y. (2004). The traditional explanation for resistance versus attitude accessibility. *Human Communication Research, 30*(3). https://doi.org/10.1111/j.1468-2958.2004.tb00735.x.

Pfau, M., Haigh, M. M., Sims, J., & Wigley, S. (2007). The influence of corporate front-group stealth campaigns. *Communication Research, 34*(1), 73.

Pfau, M., Sorenson, J., Kenski, H. C., & Nitz, M. (1990). Efficacy of inoculation strategies in promoting resistance to political attack messages: Application to direct mail. *Communication Monographs, 57*(1), 25–43. https://doi.org/10.1080/03637759009376183.

Pfau, M., Tusing, J., Koerner, A. F., Lee, W., Godbold, L. C., Penaloza, L. J., Shu-Huei, V., & Hong, Y. Y.-H. (1997). Enriching the inoculation construct the role of critical components in the process of resistance. *Human Communication Research, 24*(2). https://doi.org/10.1111/j.1468-2958.1997.tb00413.x.

Pidgeon, N. F., Kasperson, R. E., & Slovic, P. (2003). *The social amplification of risk.* Cambridge University Press.

Poortinga, W., & Pidgeon, N. F. (2003). Exploring the dimensionality of trust in risk regulation. *Risk Analysis, 23*(5), 961–972. https://doi.org/10.1111/1539-6924.00373.

Richards, A. S., & Banas, J. A. (2018). The opposing mediational effects of apprehensive threat and motivational threat when inoculating against reactance to health promotion. *Southern Communication Journal, 83*(4), 245–255. https://doi.org/10.1080/1041794X.2018.1498909.

Ringold, D. J. (2002). Boomerang effects in response to public health interventions: Some unintended consequences in the alcoholic beverage market. *Journal of Consumer Policy, 25*(1). https://doi.org/10.1023/A:1014588126336.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2). https://doi.org/10.18637/jss.v048.i02.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* The Guilford Press.

Sheldon, K. M., & Filak, V. (2008). Manipulating autonomy, competence, and relatedness support in a game-learning context: New evidence that all three needs matter. *The British Journal of Social Psychology, 47*(2), 267–283. https://doi.org/10.1348/014466607X238797.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change. - PsycNET.* Yale University Press. https://psycnet.apa.org/record/1963-06591-000.

Simons, H. W. (1976). *Persuasion: Understanding, practice, and analysis.* Addison-Wesley Publishing Co.

Stacks, D. W., & Sellers, D. E. (1986). Toward a holistic approach to communication: The effect of "pure" hemispheric reception on message acceptance. *Communication Quarterly, 34*(3). https://doi.org/10.1080/01463378609369640.

Strekalova, Y. A. (2017). Health risk information engagement and amplification on social media: News about an emerging pandemic on Facebook. *Health Education & Behavior, 44*(2), 332–339. https://doi.org/10.1177/1090198116660310.

Sun, N., Rau, P. P.-L., & Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior, 38*, 110–117. https://doi.org/10.1016/j.chb.2014.05.022.

Swani, K., Milne, G., & Brown, P. (2013). Spreading the word through likes on Facebook. *Journal of Research in Interactive Marketing, 7*(4). https://doi.org/10.1108/JRIM-05-2013-0026.

Tannenbaum, P. (1967). The congruity principle revisited: Studies in the reduction, induction, and generalization of persuasion. *Advances in Experimental Social Psychology, 3*(C), 271–320. https://doi.org/10.1016/S0065-2601(08)60346-9.

Tannenbaum, P., & Norris, E. L. (1965). Effects of combining congruity principle strategies for the reduction of persuasion. *Soins; La Revue de Référence Infirmière, 28*, 145–157. https://doi.org/10.2307/2785647.

van der Meer, T. G. L. A., & Jin, Y. (2020). Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication, 35*(5), 560–575. https://doi.org/10.1080/10410236.2019.1573295.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology, 87*(2), 246–260. https://doi.org/10.1037/0022-3514.87.2.246.

Wan, H. H., & Pfau, M. (2004). The relative effectiveness of inoculation, bolstering, and combined approaches in crisis communication. *Journal of Public Relations Research, 16*(3), 301–328. https://doi.org/10.1080/1532-754X.2004.11925131.

Warner, B. R., & McKinney, M. S. (2013). To unite and divide: The polarizing effect of presidential debates. *Communication Studies, 64*(5), 508–527. https://doi.org/10.1080/10510974.2013.832341.

Warner, B. R., McKinney, M. S., Bramlett, J. C., Jennings, F. J., & Funk, M. E. (2020). Reconsidering partisanship as a constraint on the persuasive effects of debates. *Communication Monographs, 87*(2), 137–157. https://doi.org/10.1080/03637751.2019.1641731.

Wilbur, D., Sheldon, K. M., & Cameron, G. T. (2021). Autonomy supportive and reactance supportive inoculations both boost resistance to propaganda, as mediated by state autonomy but not state reactance. *Social Influence.* https://doi.org/10.1080/15534510.2021.1908910.

Williams, G. C., Deci, E. L., & Ryan, R. M. (1998). Building health-care partnerships by supporting autonomy: Promoting maintained behavior change and positive health outcomes. In A. L. Suchman, P. Hinton-Walker, & R. Botelho (Eds.), *Partnerships in healthcare: Transforming relational process* (pp. 67–87). University of Rochester Press.

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science, 21*(2), 101. https://doi.org/10.1007/BF02894421.

Yoo, K.-H., & Gretzel, U. (2011). Influence of personality on travel-related consumer-generated media creation. *Computers in Human Behavior, 27*(2), 609–621. https://doi.org/10.0.3.248/j.chb.2010.05.002.

Zavala-Rojas, D. (2014). Thermometer scale (feeling thermometer). *Encyclopedia of quality of life and well-being research* (pp. 6633–6634). Netherlands: Springer. https://doi.org/10.1007/978-94-007-0753-5_1028.

Zerback, T., Töpfl, F., & Knöpfle, M. (2020). The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them. *New Media & Society*, 1–19. https://doi.org/10.1177/1461444820908530.