

RESEARCH ARTICLE | SPECIAL ISSUE: FROM VULNERABILITY TO VIGILANCE

Counteracting socially endorsed misinformation through an emotion-fallacy inoculation

Cecilie S. Traberg^{1*}, Thomas Morton², & Sander van der Linden¹

Received: February 29, 2024 | **Accepted:** June 2, 2024 | **Published:** June 24, 2024 | **Edited by:** Jonas R. Kunst

¹Department of Psychology, School of the Biological Sciences, University of Cambridge. ²Department of Psychology, Faculty of Social Sciences, University of Copenhagen. *Please address correspondence to Cecilie Steenbuch Traberg, cso35@cam.ac.uk, Department of Psychology, University of Cambridge, Downing Street, CB2 3EB Cambridge, United Kingdom. This article is published under the Creative Commons BY 4.0 license. Users are allowed to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

This study ($N = 755$) explores the efficacy of an emotion-fallacy inoculation in reducing susceptibility to emotionally misleading news and investigates the impact of persuasive social cues on its effectiveness. Results show that inoculation significantly reduces the perceived reliability of misinformation ($d = 0.23$), enhances participants' confidence in their reliability ($d = 0.26$), and improves veracity discernment ($d = 0.23$). Findings also reveal that social cues increase the perceived reliability ($d = 0.44$) and perceived accuracy of misinformation ($d = 0.38$), even among inoculated individuals. However, the impact of inoculation remains consistent, suggesting that, while social cues enhance the persuasiveness of misinformation, they do not diminish the effectiveness of the inoculation intervention. Finally, participants acknowledge the influence of social cues more on others than on themselves, indicating a third-person consensus effect. The findings highlight the persistent influence of social cues, even in the presence of inoculation, emphasising the need for nuanced interventions to address the complex interplay between emotions, misinformation, and social influence in the digital age.

Keywords: social influence, misinformation, inoculation, interventions, persuasion

1. INTRODUCTION

In 2020, the World Health Organization (WHO) declared an “infodemic” where the amount of (false) information, both offline and online, is contributing to confusion and making it difficult to identify trustworthy information (Pertwee et al., 2022). Similarly, in 2024, the World Economic Forum Global Risks Report named misinformation as the top global risk to society in the next few years (Global Risks Report, 2024) – and with good reason. Misinformation has been shown to impact important decisions like getting vaccinated (Loomba et al., 2021), and whom to vote for during elections (Gunther et al., 2019), but also support for key issues such as climate change (van der Linden et al., 2017), and compliance with public health guidance. As such, the infiltration of misinformation into information ecosystems poses a substantial threat to a range of pressing global issues spanning health, environment, democracy, and social justice (T. Lee, 2019; Lewandowsky et al., 2017; Maleki et al., 2021; Treen et al., 2020; van der Linden et al., 2017).

One proposed method to combat misinformation is rooted in inoculation theory (McGuire, 1964). Much like how individuals can be immunised against viral contagion, inoculation theory suggests that they can also be preemptively vaccinated against undesirable persuasive attacks. However, news increasingly flows through social media sites (Auxier & Anderson, 2021), which contain a host of other persuasive social cues that complicate the information environment and warrant attention (Traberg, Harjani, et al., 2024). In this paper, we examine whether it is possible to inoculate news consumers against misinformation, even in the presence of persuasive social cues. As heavy emotional appeals have been identified as one of the ‘fingerprints’ of misinformation (Carrasco-Farré, 2022), here we focus on emotionally misleading news and develop an emotion-fallacy inoculation intervention to tackle this. Below, we outline research demonstrating the social context of misinformation

susceptibility followed by a brief review of research on inoculation theory.

1.1 The Social Context of Misinformation Susceptibility

Research has demonstrated that ‘social proof’ - exposure to social cues that indicate the judgments of others - has a significant influence on our own perceptions of information (Asch, 1951). For example, social interaction can lead to belief acquisition (Moussaïd et al., 2013) and consensus within social groups is a predictor of individual attitudes (Kobayashi, 2018; Lewandowsky et al., 2019). Expressing, sharing, and validating information within groups and communities represents a virtually ineluctable part of the internet (Dvir-Gvirsman, 2019; Kim, 2018; S. S. Lee et al., 2021). A significant proportion of news consumers get at least some of their news from social media sites (Auxier & Anderson, 2021; Walker & Matsa, 2021) and people often discuss socio-cultural issues with their social circles (Sehulster, 2006). As such, news consumers are very likely to encounter the judgments of others in ways that might impact how they assess the veracity of misinformation. For instance, articles are liked and commented on as well as being shared on social media. Such online commentary has been found to influence perceptions of wider social consensus (Traberg, Harjani, et al., 2024), with consequences for individual beliefs (Lewandowsky et al., 2019).

Recent research shows that being exposed to social proof in the form of fabricated social consensus information (e.g., fabricated percentages of previous participant judgements) impacts misinformation susceptibility (Traberg, Harjani, et al., 2024). Moreover, being exposed to this type of ‘local’ consensus information also impacts perceptions of wider public consensus (which we refer to as *perceived consensus*) – that is, the percentage of the public that news consumers think would find the information reliable.

Although research documents that individual

judgements are impacted by social proof, individuals themselves may underestimate their own susceptibility to such cues. Work on social influence shows that this is often undetected by influenced individuals (e.g., Nolan et al., 2008). Similarly, work on media effects demonstrates that people tend to perceive media influences as more pronounced on others than on themselves (i.e., the ‘third person media effect’; Sherrick, 2016), which Douglas and Sutton (2004) suggest stems from an underestimation of the persuasibility of the self. We propose the existence of a similar phenomenon regarding social cues, which we term the “*third person consensus effect*”. If this effect exists, where individuals believe that social (consensus) cues wield a greater influence on others than on themselves, it warrants attention from researchers in the misinformation field. Such a perception could lead individuals to develop a false sense of immunity to social influence, potentially resulting in an underestimation of their vulnerability to misinformation.

1.2 Inoculation Theory: A Potential Solution

In response to the spread of misinformation, researchers have investigated whether it is possible to prevent its psychological effects before people are exposed to it (Cook et al., 2017; Lewandowsky & van der Linden, 2021; Traberg et al., 2022, 2023). A problem with misinformation is that even when explicit warnings or retractions are made, misinformation can continue to impact reasoning (Lewandowsky et al., 2012). Inoculation theory (McGuire, 1964) proposes that much like vaccinations against viral contagion, individuals can be psychologically vaccinated against persuasive attacks (Roozenbeek, Traberg, et al., 2022; Traberg et al., 2022; van der Linden, 2022). In contrast to post-hoc measures such as fact-checks, corrections or retractions, inoculation theory represents a *pre-emptive* solution to countering misinformation. A growing body of evidence supports the use of inoculation interventions to counter both specific misinformation content (Mason et al., 2023; Spampatti et al., 2023; van

der Linden et al., 2017) and the underlying manipulative techniques contained in misinformation (Banas et al., 2023; Cook et al., 2017; Lewandowsky & Yesilada, 2021; Roozenbeek, Traberg, et al., 2022; Roozenbeek, van der Linden, et al., 2022; Roozenbeek & van der Linden, 2020).

According to the theory (McGuire, 1961), inoculation interventions should consist of two components. A *threat* component forewarns people that a persuasive attack is imminent to cognitively prepare and motivate the mind to be ready to engage with the second component. A *refutational pre-emption* provides people with the cognitive tools to refute future misinformation, either through exposing specific false arguments (van der Linden et al., 2017), or misleading techniques that underlie misinformation more generally (Cook et al., 2017; Roozenbeek, Traberg, et al., 2022; Roozenbeek, van der Linden, et al., 2022).

Technique-based inoculations have witnessed particular attention due to their potential for scalability (Roozenbeek, van der Linden, et al., 2022). However, while inoculation might help individuals be more discerning in their information environment, information is typically received in the context of other people (Clarkson et al., 2013; Hertz et al., 2017; Tormala et al., 2009). As a result, social processes might also influence the degree to which information is accepted as true. For this reason, researchers have started calling for investigations of how contextual cues in the misinformation environment may impact the efficacy of inoculation interventions (Traberg, 2024).

1.3 Emotional Deception & Inoculation

Emotions play a key role in the spread of misinformation (Carrasco-Farré, 2022). Firstly, people are more susceptible to misinformation in an emotional state (Martel et al., 2020). Secondly, emotions play a role in the spread of misinformation with emotionally deceptive news making up part of the “fingerprints” of misinformation (Carrasco-Farré, 2022). Furthermore,

online misinformation has been shown to be significantly more emotional than non-deceptive content (Paschen, 2019; Peng et al., 2023). Although emotions are a valuable part of human communication (Juez & Mackenzie, 2019), when used deceptively to evoke outrage, anger, or other strong emotions (Roozenbeek, van der Linden, et al., 2022), they can hinder our ability to critically assess information and manipulate attention away from the evidence by evoking irrelevant cues (i.e. the appeal-to-emotion-fallacy: Blassnig et al., 2019; Hamlin, 1970). For this reason, technique-based inoculation interventions target emotional deception as a key misinformation strategy (Roozenbeek, van der Linden, et al., 2022).

Despite the demonstrated effects of social cues on individual information judgements, it is unknown whether inoculation interventions can withstand these social influences. Is it possible to inoculate individuals against misinformation even when their social environment appears to endorse misinformation as reliable? In the current work we set out to address this question. Firstly, we investigate whether an emotion-fallacy inoculation intervention can protect individuals against the influence of misleading emotional news, and secondly, whether this intervention is successful in the presence of strong social influences. Finally, we examine the potential for third person consensus effects.

2. METHODS

The purpose of this study was to test the efficacy of an inoculation intervention in the presence of persuasive social cues in a 2 (inoculation vs control) by 2 (social cues vs no social cues) randomised control trial design. The study was approved by the Cambridge Psychology Research Ethics Committee (PRE.2022.117).

2.1 Participants

The sample was UK-based. A power analysis indicated that 580 participants would be

required to detect the hypothesised effects based on previously found inoculation (Roozenbeek, van der Linden, et al., 2022) and social cue effect sizes (Traberg, Harjani, et al., 2024). After excluding 5 participants who did not pass the attention checks the final sample was $N = 755$ with 48.4% identifying as liberal (i.e., below 4 on a 1-7 Left-Right political orientation scale), 54.6% female, 66% had at least a bachelor's degree and $M_{\text{age}} = 43$, $SD_{\text{age}} = 13.3$. 94% of participants had social media accounts with 78% reporting they used it at least occasionally and 59% of participants got the majority of their news from online news sites or on social media.

Participants were randomly assigned to one of the four conditions with $n = 192$ in the *Inoculation with social cues* condition, $n = 186$ in the *Inoculation without social cues* condition, $n = 188$ in the *Control with social cues* condition and $n = 189$ in the *Control without social cues* condition.

2.2 Procedure

A study on “News Evaluation” was advertised via *Prolific* and the study was run on *Qualtrics*. Upon electronic consent, participants were randomly assigned to one of four conditions: 1) Inoculation with social cues, 2) Inoculation without social cues, 3) Control with social cues, or 4) Control without social cues.

In the inoculation conditions, participants were exposed to a two-part text-based inoculation message. They were asked to read the message. The page could not be skipped until at least 30 seconds had passed. In the control conditions, participants were given a word-search task, with a 30 second timer that ensured they stayed on the page for the same duration. In all conditions, participants were subsequently shown a series of misleading and neutral headlines and asked to make a series of judgements about them (see “Measures”). Participants who were assigned to the conditions with social cues (both the inoculation group and control/word search group) were informed that a high percentage of previous participants had

Table 1*Misleading and Neutral Headlines Included*

Headline	Misleading	Neutral
Horrific nuclear meltdown causes chaos and despair in local town, nuclear energy no solution for climate change residents claim.	x	
Heartbreaking story: baby elephant gets horribly hurt after falling off a ledge, mother elephant cries for HOURS!	x	
Students accused of cheating note how outrageous and inhumane housing conditions make it difficult to study.	x	
Physical fitness keeps your brain in good shape.		x
Netflix to include mobile games for subscribers.		x
Apple, Google and Amazon named as most valuable brands in the world.		x

judged each headline to be reliable (always above 75

percent). Participants were also informed that their judgements would be used to calculate the average percentage for future participants. In the *no social cues* conditions, participants were not shown this statement. Then, participants answered a series of demographic questions, third person effect questions and provided responses to psychological scales. Finally, participants were debriefed.

2.3 Materials

2.3.1 Inoculation Message

The inoculation intervention consisted of a two-part message (see Figure 1). The threat component forewarned participants about the threat of misinformation. The cognitive refutation (pre-bunk) element of the message informed participants about the use of emotionally deceptive language. The message also provided information about the appeal-to-emotion fallacy, which involves manipulating emotions to divert attention or influence perceptions away from the evidence. To avoid fear-mongering about the use of emotions in news in general, the message highlighted that emotions are a natural and valuable part of human communication, and that they can enhance the impact of messaging when used honestly and transparently.

2.3.2 Control Task

In line with previous inoculation research using control conditions (van der Linden et al., 2017) participants in the control condition were given a neutral word-search puzzle.

2.3.3 Headlines

Participants made judgements regarding six headlines which they were informed had been published online between 2021-2022: 3 emotionally misleading headlines and 3 neutral headlines (see Table 1). The emotionally misleading headlines were created for the purpose of the research and used the appeal-to-emotion fallacy: That is, they used emotional language to divert attention from an issue. e.g., “*Horrific nuclear meltdown causes chaos and despair in local town, nuclear energy no solution for climate change residents claim.*” In this instance, the mention of ‘nuclear meltdown’ does not pertain to the potential of nuclear power as a solution to climate change. The neutral headlines were real headlines that had been published in 2021 and did not make use of emotions, instead stating ‘matter of fact’ happenings or news.

2.4 Measures

Participants were asked to make the following judgements about news headlines:

2.4.1 Perceived Reliability of Headlines

After reading each headline, participants were asked: If 1 is “*Not at all reliable*” and 7 is “*Very reliable*”, how reliable is the above headline? (1-

7 scale). The reliability ratings were averaged separately for misleading headlines (Cronbach's $\alpha = 0.78$) and neutral headlines (Cronbach's $\alpha = 0.78$).

2.4.2 Confidence

For each headline, participants were also asked: "If 1 is *Not at all confident*" and 7 is *Very confident*", how confident are you in your reliability judgement above?" (1-7 scale). The confidence ratings were averaged separately for misleading headlines (Cronbach's $\alpha = 0.87$) and neutral headlines (Cronbach's $\alpha = 0.82$).

2.4.3 Perceived Accuracy

For each headline, participants were also asked: "If 1 is *Not at all likely*" and 7 is *Very likely*", how likely is it that the content of the above headline is an accurate description of reality?" (1-7 scale). The perceived accuracy ratings were averaged separately for misleading headlines (Cronbach's $\alpha = 0.78$) and neutral

headlines (Cronbach's $\alpha = 0.75$).

2.4.4 Perceived Consensus

For each headline, participants were asked: "What percentage of the general public do you think would believe the above headline?" (0-100%). The perceived consensus ratings were averaged separately for misleading headlines (Cronbach's $\alpha = 0.80$) and neutral headlines (Cronbach's $\alpha = 0.75$).

The following variables were also computed:

2.4.5 Reliability Discernment

Average perceived reliability of neutral headlines minus average perceived reliability of misleading headlines.

2.4.6 Accuracy Discernment

Average perceived accuracy of neutral headlines minus average perceived accuracy of misleading headlines.

Figure 1

Two-Component Inoculation Message: Threat and Pre-Emption.



In addition to the above main measures, we also collected data for the following variables:

2.4.7 Third Person Consensus Effects

Participants were asked to report on two measures:

Social cue influence on the self: On a scale from 1-7, where 1 is "Not at all" and 7 is "A lot", please rate how much you think other people's judgments influence your own opinions and behaviour.

Social cue influence on others: On a scale from 1-7, where 1 is "Not at all" and 7 is "A lot", please rate how much you think other people's judgments influence the general public's opinions and behaviours.

Third person consensus effect: Computed as *social cue influence on others* – *social cue influence on the self*. A positive score indicates that participants judge others to be more susceptible to social cues than the self.

2.4.8 Attention Checks

We included two attention checks. In the first, participants were asked to select 'Disagree' for the question: "Imagine you are planning a trip to the beach. Please select 'Disagree' for this question." In the second, they were told: "In the previous questions, you were asked to rate the reliability of various headlines. Please select 'Reliability 6' in this question."

2.4.9 Demographic Variables

Socio-demographic variables included gender (male, female, other), age (birth year), political orientation (measured on a 7-point Likert scale where 1 is very left-wing and 7 is very right-wing), level of education (less than high school degree, high school graduate, bachelor's degree, master's degree, doctoral degree or professional degree), current use of social media ("I don't have any accounts"; "I have one or more accounts but I hardly ever use them"; "I have one or more accounts, and I use them occasionally"; "I have one or more accounts and I

use them often"; "I have one or more accounts and I use them on a daily basis") and main news consumption source ("I don't really follow the news", "Social Media", "TV and radio"; "Print Media (newspapers, magazines)"; "Word of mouth", "Online news sites (excluding social media)").

3. RESULTS

The main hypotheses and analyses were pre-registered: <https://osf.io/8mdhp/>. Exploratory analyses are explicitly highlighted as such. Data for this study is available on: <https://osf.io/5vzxcg/>.

3.1 Inoculation Effects

Based on previous research we hypothesised that inoculated participants would show greater immunity to misinformation compared to those in the control condition on several key outcome variables. We specifically hypothesise:

H1. Participants who are inoculated will judge misinformation headlines to be (A) less reliable, (B) less accurate and (C) be more confident in their reliability judgements than participants in the control condition.

Three independent samples t-tests were run with Bonferroni corrections for multiple testing, confirming the three hypotheses (A, B & C) showing that participants who were inoculated judged misinformation headlines to be significantly less reliable ($t(752.91) = 3.10$, $p = 0.002$, $M_{Control} = 3.73$, 95% CI [3.62, 3.85], $M_{Inoculation} = 3.48$, 95% CI [3.37, 3.59], $d = 0.23$), were more confident in their reliability judgement ($t(751.95) = -3.54$, $p < 0.001$, $M_{Control} = 4.50$, 95% CI [4.38, 4.62], $M_{Inoculation} = 4.82$, 95% CI [4.69, 4.95], $d = 0.26$) and judged the headlines to be less accurate ($t(750.41) = 2.90$, $p = 0.004$, $M_{Control} = 3.68$, 95% CI [3.57, 3.80], $M_{Inoculation} = 3.44$, 95% CI [3.33, 3.55], $d = 0.21$). We therefore confirm *H1A*, *H1B* and *H1C*.

As there has been scholarly discussion regarding the efficacy of inoculation interventions in improving discernment – that is, the ability to discern between true and false – we also put

this to the test in the following hypothesis:

H2. Participants who are inoculated will show (A) higher reliability discernment and (B) accuracy discernment than participants in the control condition.

Confirming *H2*, two independent samples *t*-tests run with Bonferroni corrections for multiple testing showed that participants who were inoculated had higher reliability discernment ($t(746.41) = -3.23, p = 0.001, M_{Control} = 1.89, 95\% \text{ CI } [1.75, 2.03], M_{Inoculation} = 2.23, 95\% \text{ CI } [2.08, 2.39], d = 0.24$) and higher accuracy discernment ($t(752.62) = -3.19, p = 0.001, M_{Control} = 1.89, 95\% \text{ CI } [1.76, 2.01], M_{Inoculation} = 2.18, 95\% \text{ CI } [2.05, 2.31], d = 0.23$). That is, inoculated participants were better at discerning between the veracity of misleading and neutral information than those in the control condition.

An exploratory analysis showed that participants who were inoculated believed a significantly larger percentage of the general public would judge the headlines to be reliable (which we refer to as ‘perceived consensus’) ($t(742.27) = -4.65, p < 0.001, M = 65.07, 95\% \text{ CI } [63.56, 66.58]$) compared to in the control condition ($M = 59.68, 95\% \text{ CI } [57.98, 61.38], M_{diff} = -5.49, d = 0.34$). Figure 2 illustrates the results of the six comparisons above.

3.2 Social Cue Effects and Interaction with Inoculation

A key question posed in this paper is whether even those who are inoculated are impacted by social proof. *H3* tests whether this is the case. *H4* expands on this assumption by testing whether social cues have a greater impact on those who are not inoculated.

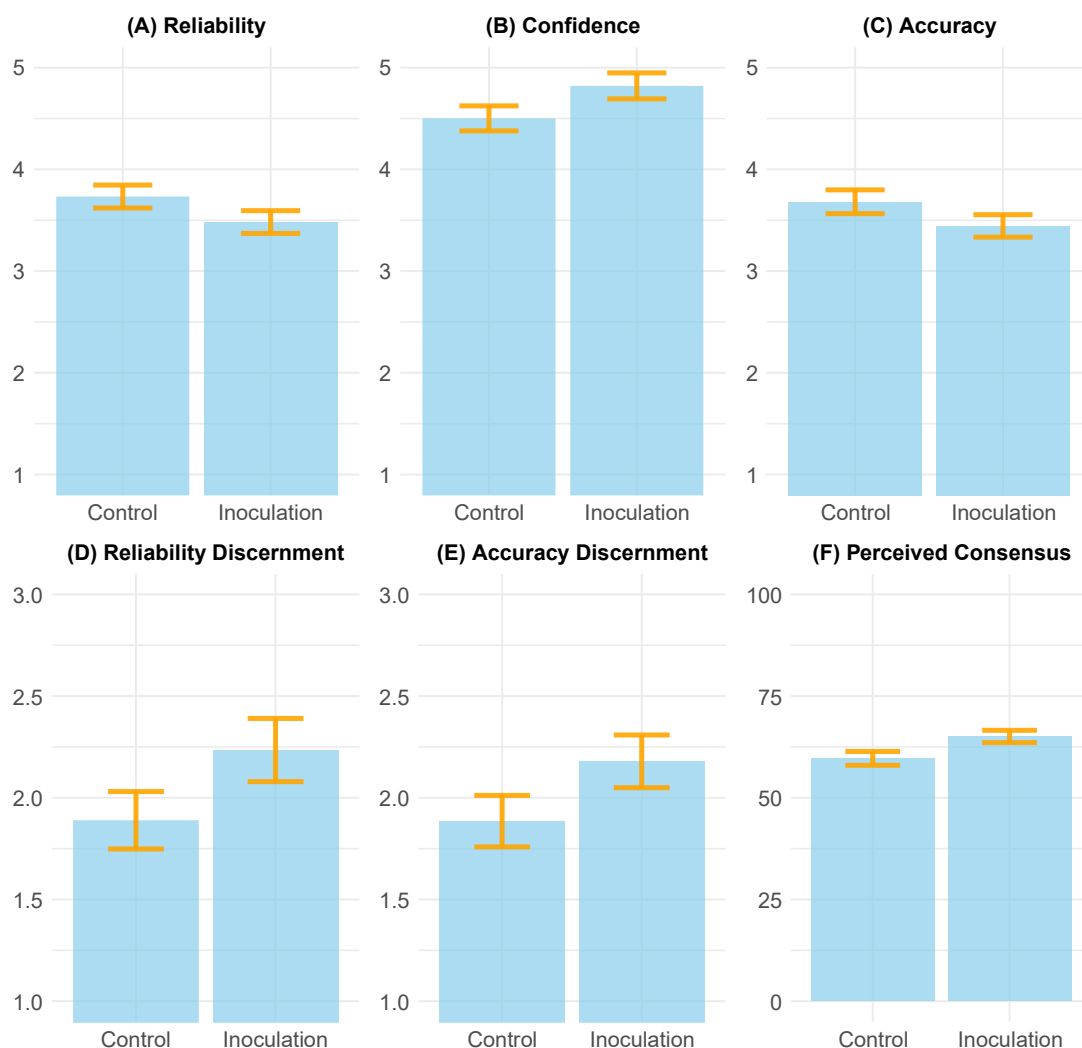
H3. Participants in the inoculation with social cues condition will judge misinformation headlines to be (A) more reliable and (B) more accurate than participants in the inoculation without social cues condition.

H4: There will be an interaction between social cue condition and inoculation such that

participants in the control condition (who are not inoculated) will be more impacted by social cues than those who are inoculated, showing (A) higher perceived reliability and (B) perceived accuracy of misleading headlines.

To examine *H3* and *H4*, we ran a 2x2 factorial ANOVA to assess whether the difference between the control and inoculation conditions was significantly smaller when social cues were present: That is, did the impact of being inoculated (compared to the control) depend on whether social cues were included or not? The ANOVA model showed a significant main effect of inoculation ($F(1,751) = 10.37, \eta_p^2 = 0.01, p = 0.001$), a main effect of social cues ($F(1,751) = 35.75, \eta_p^2 = 0.05, p < 0.001$) but no interaction between the two ($F(1,751) = 0.19, \eta_p^2 < 0.01, p = 0.659$). That is, inoculated participants judged misleading headlines as less reliable ($M = 3.48, 95\% \text{ CI } [3.37, 3.59]$) than those in the control ($M = 3.73, 95\% \text{ CI } [3.62, 3.84], p = 0.001, d = 0.23$). Furthermore, the presence ($M = 3.84, 95\% \text{ CI } [3.73, 3.95]$) (versus absence) ($M = 3.37, 95\% \text{ CI } [3.26, 3.48]$) of social cues increased the perceived reliability of misleading information ($p < 0.001, d = 0.43$). Interestingly, although the interaction between social cues and inoculation was not found significant (rejecting *H4A*), the effect size for social cues was larger ($\eta_p^2 = 0.05$) than for inoculation ($\eta_p^2 = 0.01$). Results of these analyses are visualised in Figure 3 below.

This result suggests that although it is possible to inoculate individuals in the presence of social cues, social cues make information more persuasive even for those who have been inoculated. We therefore confirm *H3A*. Post-hoc tests with planned contrasts between the *inoculation with social cues* condition and the *inoculation without social cues* conditions revealed that inoculated participants were significantly impacted by social cues ($t(751) = -3.92, p < 0.001, d = 0.40$). Inoculated participants who were exposed to social cues judged misleading headlines to be significantly more reliable ($M = 3.70, 95\% \text{ CI } [3.54, 3.85]$), than inoculated participants

Figure 2*Judgements Compared Across Control vs Inoculation Conditions.*

Note. Error bars show 95% confidence intervals.

who did not see social cues ($M = 3.26$, 95% CI [3.10, 3.42]).

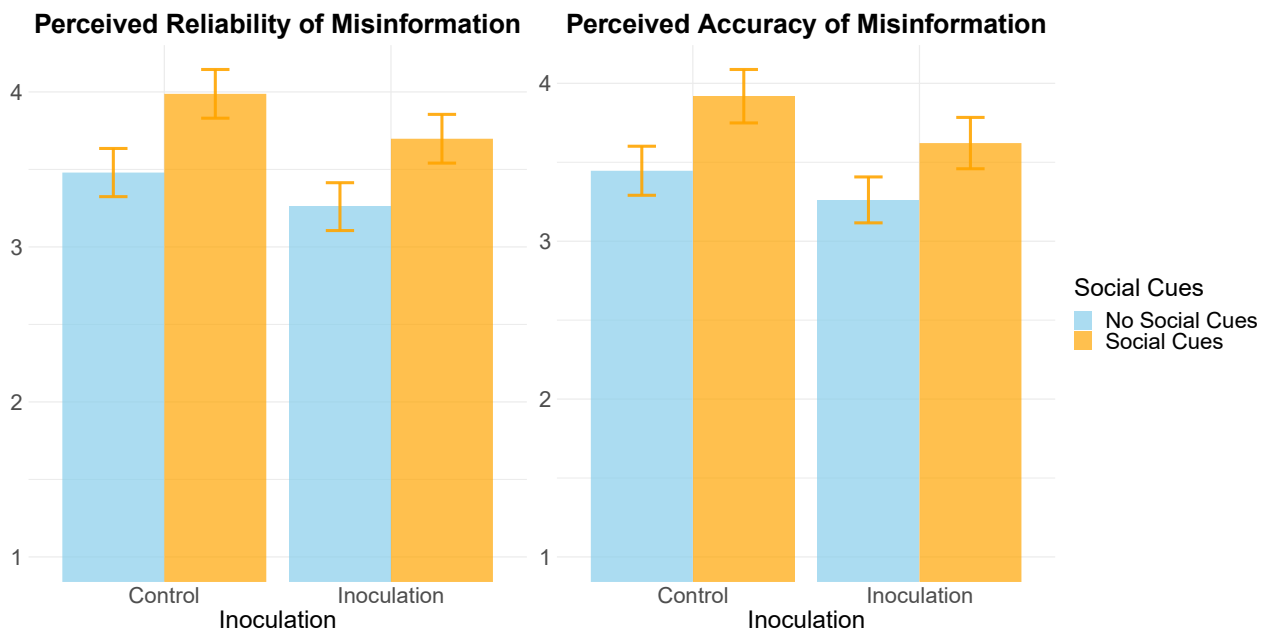
To assess the impact of social cues among inoculated individuals on perceived accuracy of misleading headlines, a 2x2 ANOVA with inoculation and social cues as the IVs and perceived accuracy as the DV was run. We find a significant main effect of inoculation ($F(1,751) = 8.97$, $p = 0.003$, $\eta_p^2 = 0.01$) with perceived accuracy being higher in the control condition ($M = 3.68$, 95% CI [3.57, 3.79]) compared to inoculation condition ($M = 3.44$, 95% CI [3.33, 3.55], $d = 0.22$), a main effect of social cues ($F(1,751) = 26.79$, $p < 0.001$, $\eta_p^2 = 0.034$), with perceived accuracy being higher in the social cues condition ($M = 3.77$, 95% CI [3.66, 3.88]) than the no social cues

condition ($M = 3.35$, 95% CI [3.24, 3.47], $d = 0.38$), but no interaction between the two ($F(1,751) = 0.49$, $p = 0.485$, $\eta_p^2 < 0.01$). Post-hoc tests with planned contrasts between the inoculation with social cues and the inoculation without social cues conditions showed that among inoculated participants, those who saw social cues judged headlines as more accurate ($M = 3.62$, 95% CI [3.47, 3.78]) than those who did not ($M = 3.26$, 95% CI [3.10, 3.42], $t(751) = -3.17$, $p = 0.002$, $d = 0.33$). As such we confirm $H3B$. As there was no significant interaction between social cues and inoculation however, we reject $H4B$.

In line with previous research showing that being exposed to social cues indicating the judgments of others can impact perceptions of

Figure 3

Perceived Reliability of Misinformation and Perceived Accuracy of Misinformation Across 4 Conditions.



Note. Error bars show 95% confidence intervals.

wider public consensus (perceived consensus) (Traberg, Harjani, et al., 2024), we test the following hypothesis:

H5: Participants in the social cues condition will show higher perceived general public consensus than those in the no social cues condition.

An independent samples *t*-test comparing social cue conditions on perceived general public consensus revealed that participants who had been exposed to social cues perceived significantly higher social consensus than those who had not been exposed to social cues ($t(743.18) = -8.02, p < 0.0001, M_{diff} = 9.05, 95\% \text{ CI } [6.84, 11.27], d = 0.58$). That is, being exposed to social cues suggesting a local group of previous participants had judged the misleading headlines to be reliable (compared to a control) led participants to believe that a higher percentage of the public would also find the headlines reliable. We therefore confirm H5.

3.3 Third Person Consensus Effects

As research has suggested that individuals underestimate the persuasibility of the self (Douglas & Sutton, 2004), we hypothesise that:

H6: Participants will judge other people as being more likely to be impacted by social cues than themselves.

Confirming H6, a paired samples *t*-test showed that participants believed other people would be more influenced by social cues ($t(754) = -22.61, p < 0.001, M_{diff} = 1.30, 95\% \text{ CI } [1.19, 1.42], d = 0.82$) than they would be themselves. Figure 4 illustrates the respective means for the perceived influence of social cues on self vs others.

As data for this variable was collected after the experimental manipulation, as a robustness check we assessed whether the social cue manipulation had a significant impact on the two outcome variables. An independent samples *t*-test showed that seeing social cues alongside news headlines did not have a significant

impact on the perceived influence of social cues on the self ($t(752.95) = -0.96$, $M_{diff} = 0.10$, 95% CI [-0.31, 0.11], $p = 0.339$, $d = 0.07$) nor on others ($t(751.18) = -0.48$, $M_{diff} = -0.04$, 95% CI [-0.19, 0.12], $p = 0.628$, $d = 0.04$).

3.3.1 Third Person Consensus Effect as a Moderator

As we find a significant difference between the perceived influence of social cues on others versus the self, we explored the impact of the computed third person consensus effect variable (*social cue influence on others – social cue influence on the self*) on the main outcome measure (*perceived reliability of misinformation*), as well as its potential moderating role.

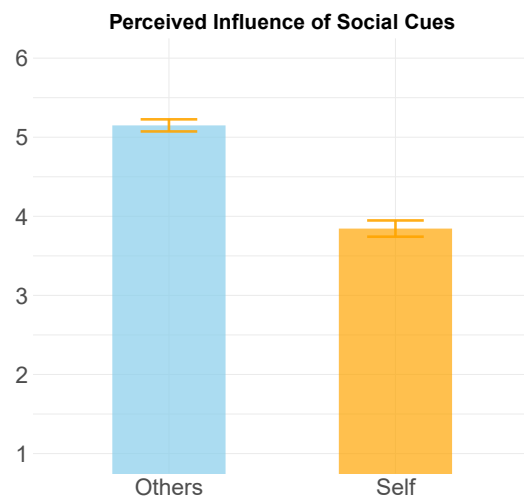
An exploratory regression analysis was run to assess the moderating role of the third person consensus effect on perceived reliability of misinformation. Results indicated that the main effect of the *third person consensus effect* on perceived reliability of misinformation was not significant ($\beta = -0.01$, $p = 0.729$). In line with previous analyses, the direct impacts of social cues ($\beta = 0.48$, $p < 0.001$) and inoculation ($\beta = -0.34$, $p = 0.001$) were significant. However, the interaction terms for the third person consensus effect with social cues and inoculation were not significant. That is, the interaction between third person consensus and social cues ($\beta = -0.01$, $p = 0.895$) and the interaction between third person consensus and inoculation ($\beta = 0.06$, $p = .214$) did not significantly moderate the relationships between these variables and perceived reliability of misinformation.

3.3.2 Inoculation and Neutral Headlines

In exploratory analyses, we evaluate the impact of inoculation and social cues on the evaluation of neutral headlines. First, we analyse the impact of inoculation and social cues on perceived reliability of neutral headlines. A 2x2 factorial ANOVA showed no significant main effect of inoculation ($F(1, 751) = 1.28$, $p = 0.258$), but a significant main effect of social cues ($F(1, 751) = 7.48$, $p = 0.006$) and no significant interaction

Figure 4

Perceived Effects of Social cues on Self vs Others.



Note. Error bars show 95% confidence intervals.

between the two ($F(1, 751) = 0.39$, $p = 0.532$). Tukey's post hoc tests revealed that the only significant contrast was between the inoculated participants who saw social cues ($M = 5.80$, 95% CI [5.64, 5.96]) and the control group who did not see social cues ($M = 5.49$, 95% CI [5.32, 5.65], $p = 0.031$, $d = 0.28$). Second, we analyse the impact of inoculation and social cues on perceived accuracy of neutral headlines. Here, a 2x2 factorial ANOVA showed no significant effect of inoculation, no significant effect of social cues and no significant interaction effects.

As such, neither inoculation nor the social cue manipulation significantly impacted perceptions of neutral headlines. Nevertheless, it is worth noting that prior research suggests social cues may only influence perceptions when they oppose one's own views (Traberg, Harjani, et al., 2024). In this case, it is therefore not surprising that social cues indicating previous participants judged headlines as *reliable* did not impact judgements of neutral news.

4. DISCUSSION

The findings of this study shed light on the interplay between inoculation, social cues, and

susceptibility to emotionally misleading news. The results indicate that an emotion-fallacy inoculation intervention effectively reduces susceptibility to misleading emotional news, supporting the notion that pre-emptive strategies can protect individuals against the persuasive tactics employed in misinformation dissemination. The significant decrease in perceived reliability, increased confidence in reliability, and improved veracity discernment among inoculated participants underscores the robustness of the emotional fallacy inoculation approach. That is, inoculating news consumers against emotionally manipulative news increases their ability to resist judging emotionally deceptive headlines as reliable and accurate, and increases their confidence in this assessment.

Of course, emotions are part and parcel of human communication and play a role in all types of discourse (Juez & Mackenzie, 2019), including factual news that adheres to strict journalistic practices. The inoculation message in this study specifically noted this caveat, and it is therefore positive that inoculated participants were better at discerning between misleading and neutral news compared to the control condition. As such, contrary to some reports that media literacy interventions could make people too sceptical of the news (Hoes et al., 2024; van der Meer et al., 2023), we find that inoculation messages which help people look out for specific emotional manipulation attempts (rather than all emotional news) improve veracity discernment, consistent with other recent work on logic and fallacy-based inoculation (Banas et al., 2023; Hruschka & Appel, 2023).

Interestingly, the study also reveals a noteworthy influence of social cues on misinformation susceptibility. Even among inoculated individuals, those who were exposed to persuasive social cues judged misleading headlines as more reliable than those who only saw the headlines. This is in line with research showing that when social cues signal that a majority judges misinformation to be reliable, this increases

misinformation susceptibility (Traberg, Harjani, et al., 2024). This suggests that while inoculation can reduce misinformation susceptibility in general, it does not protect news consumers against the effects of social proof.

It is interesting to note that the social cue effect (increasing misinformation susceptibility) was descriptively larger than the inoculation effect (reducing misinformation susceptibility). This highlights both the potency of social influence in shaping individuals and underscores the importance of social context in considering what makes people vulnerable to misinformation, including the social dynamics that spread and sustain it. While crowds have been shown to be wise when their judgements are aggregated at scale (Martel et al., 2024), social media users are rarely exposed to the aggregated judgements of sufficiently large, independent and diverse groups – group attributes that seminal research has identified as necessary for the emergence of ‘collective intelligence’ (Malone & Bernstein, 2022). In fact, due to the existence of online echo-chambers (Törnberg, 2018), it is likely that news consumers do not wish to call out misinformation shared within their own networks (Allen et al., 2022).

Given that current interventions against misinformation focus on content rather than context, the need to explore the role of social context more thoroughly in misinformation interventions seems imperative (Traberg, 2024). That said, the lack of a significant interaction between inoculation and social cues suggests that while social influences can be problematic, the protective benefits of inoculation remain consistent in the face of this. This underscores the potential robustness and versatility of inoculation interventions in addressing the challenges posed by social dynamics in the digital age (Compton et al., 2021). The results here speak to the notion that inoculation interventions more generally may be useful for reducing online misinformation susceptibility - even in social contexts.

However, the current research also suggests a necessity for developing inoculation messages that inoculate news consumers against social influence effects, in particular as the study also uncovers what we term a '*third-person consensus effect*'. This is an effect wherein participants acknowledge the influence of social consensus cues more on others than on themselves. This highlights that while individuals are aware of the impact of social influence on others, they do not tend to acknowledge the effect of social influence on themselves. This finding highlights the potential need for further interventions that can address this meta-perception. For example, research has highlighted that meta-cognitive judgements - how citizens reason about their own reasoning - can predict important tendencies such as likelihood of polarizing science (Said et al., 2022).

This research is of course not without limitations. Firstly, while the study finds significant inoculation effects, the effect size were smaller ($d = 0.23$) than those generally found in game-based interventions, which tend to range from $d = 0.35$ $d = 0.60$ (Basol et al., 2020, 2021; Roozenbeek, Traberg, et al., 2022; Roozenbeek & van der Linden, 2019) though still in line with the average meta-analytic effect for veracity discernment (Lu et al., 2023). This may be due to the fact that game-based interventions are often longer (e.g., 7–15-minute gameplay) and require more cognitive involvement and reflection than reading text on a screen. However, text-based interventions may be easier to develop, and employ by practitioners or in public communication campaigns. So, despite their more modest effects, these may still be practically useful.

Second, the study focused on inoculating individuals against emotionally misleading news. Naturally, there are more deception strategies employed in misinformation (Roozenbeek, van der Linden, et al., 2022). This study cannot speak to whether inoculation against more polarising, extreme, or identity-based

misinformation (Pereira et al., 2023; Van Bavel & Pereira, 2018) is robust in the presence of persuasive social cues. Promisingly though, recent work has shown that inoculation successfully reduces susceptibility to misinformation that is published by political in-group news outlets (Traberg, Roozenbeek & van der Linden, 2024), suggesting that inoculation can offer protection in the face of some identity-based source effects.

In addition, in relation to social cues, the social proof included in the present study was attributed to an unknown and undefined group. But social identity research highlights that we are more likely to be influenced by others who share with us a salient and meaningful social category – when others are “ingroup members” (Spears, 2021). Attempts at influence by out-group members are instead likely to fail or may even backfire and contribute to polarisation (e.g., Abrams et al., 1990; Mackie et al., 1990; McGarty et al., 1994). Further research is needed to assess whether inoculation can protect individuals against misinformation, even when social consensus (or majority) cues indicate that their in-group supports or endorses the misinformation. Third, although the protective effects of inoculation have been shown to last up to three months (Maertens et al., 2021), this study cannot speak to whether or not the effects of this particular text-based intervention stand the test of time. Finally, the study used a UK-based sample, limiting the generalisability of the findings to a more diverse global audience.

Despite these limitations, this research contributes to the understanding of the intricate relationship between inoculation, social cues, and misinformation susceptibility. The results suggest that while inoculation provides a valuable defence against emotionally misleading news, the role of social influence cannot be ignored. As misinformation continues to evolve in the digital landscape, the findings underscore the need for multifaceted and adaptive

interventions to effectively mitigate its impact on public perception and decision-making. Future interventions may especially need to incorporate strategies to address the influence of social cues to disrupt the real-world dynamics through which misinformation thrives within social networks. For example, injunctive social norms – descriptions of what most people approve or disapprove of – have been shown to increase reportings of fake news (Gimpel et al., 2021). This highlights that while in the current study, social cues play a negative influence role, it may be possible to harness social information to improve the online news environment.

5. REPRODUCIBILITY STATEMENT

The data, scripts and the codebook can be found at <https://osf.io/5vzxf/>.

6. CONFLICTS OF INTEREST

The authors declare no competing interests.

7. AUTHOR CONTRIBUTIONS

CST designed the study, collected and analysed the data, wrote up the results and drafted the paper. TM contributed to the design of the study and the write up of the manuscript. SVDL guided all aspects of the project.

8. ACKNOWLEDGEMENTS

This research was supported by the Harding Distinguished Postgraduate Scholarship, Economic and Social Research Council.

REFERENCES

- Abrams, D., Wetherell, M., Cochran, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology*, 29(2), 97–119. <https://doi.org/10.1111/j.2044-8309.1990.tb00892.x>
- Allen, J., Martel, C., & Rand, D. G. (2022). Birds of a feather don't fact-check each other: Partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3491102.3502040>
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In *Groups, leadership and men; research in human relations* (pp. 177–190). Carnegie Press.
- Auxier, B., & Anderson, M. (2021, April 7). Social Media Use in 2021. *Pew Research Center: Internet, Science & Tech*. <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>
- Banas, J. A., Bessarabova, E., Penkauskas, M. C., & Talbert, N. (2023). Inoculating Against Anti-Vaccination Conspiracies. *Health Communication*, 1–9. <https://doi.org/10.1080/10410236.2023.2235733>
- Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & van der Linden, S. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data & Society*, 8(1), <https://doi.org/10.1177/20539517211013868>
- Basol, M., Roozenbeek, J., & van der Linden, S. (2020). Good News about Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News. *Journal of Cognition*, 3(1), 2. <https://doi.org/10.5334/joc.91>
- Blassnig, S., Büchel, F., Ernst, N., & Engesser, S. (2019). Populism and Informal Fallacies: An Analysis of Right-Wing Populist Rhetoric in Election Campaigns. *Argumentation*, 33(1), 107–136. <https://doi.org/10.1007/s10503-018-9461-2>
- Carrasco-Farré, C. (2022). The fingerprints of misinformation: How deceptive content differs from reliable sources in terms of cognitive effort and appeal to emotions. *Humanities and Social Sciences Communications*, 9(1), 1–18. <https://doi.org/10.1057/s41599-022-01174-9>

- Clarkson, J. J., Tormala, Z. L., Rucker, D. D., & Dugan, R. G. (2013). The malleable influence of social consensus on attitude certainty. *Journal of Experimental Social Psychology*, 49(6), 1019–1022. <https://doi.org/10.1016/j.jesp.2013.07.001>
- Compton, J., van der Linden, S., Cook, J., & Basol, M. (2021). Inoculation theory in the post-truth era: Extant findings and new frontiers for contested science, misinformation, and conspiracy theories. *Social and Personality Psychology Compass*, 15(6), e12602. <https://doi.org/10.1111/spc3.12602>
- Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, 12(5), e0175799. <https://doi.org/10.1371/journal.pone.0175799>
- Douglas, K. M., & Sutton, R. M. (2004). Right about others, wrong about ourselves? Actual and perceived self-other differences in resistance to persuasion. *British Journal of Social Psychology*, 43(4), 585–603. <https://doi.org/10.1348/0144666042565416>
- Dvir-Gvirsman, S. (2019). I like what I see: Studying the influence of popularity cues on attention allocation and news selection. *Information, Communication & Society*, 22(2). <https://doi.org/10.1080/1369118X.2017.1379550>
- Gimpel, H., Heger, S., Olenberger, C., & Utz, L. (2021). The Effectiveness of Social Norms in Fighting Fake News on Social Media. *Journal of Management Information Systems*, 38(1), 196–221. <https://doi.org/10.1080/07421222.2021.1870389>
- Global Risks Report*. (2024). World Economic Forum. <https://www.weforum.org/publications/global-risks-report-2024/>
- Gunther, R., Beck, P. A., & Nisbet, E. C. (2019). “Fake news” and the defection of 2012 Obama voters in the 2016 presidential election. *Electoral Studies*, 61, 102030. <https://doi.org/10.1016/j.electstud.2019.03.006>
- Hamlin, C. L. (1970). *Fallacies*. Methuen. <https://www.abebooks.com/Fallacies-Hamblin-C.L-London-Methuen-Ltd/31169487055/bd>
- Hertz, U., Palminteri, S., Brunetti, S., Olesen, C., Frith, C. D., & Bahrami, B. (2017). Neural computations underpinning the strategic management of influence in advice giving. *Nature Communications*, 8(1), Article 1. <https://doi.org/10.1038/s41467-017-02314-5>
- Hoes, E., Aitken, B., Zhang, J., Gackowski, T., & Wojcieszak, M. (2024). Prominent misinformation interventions reduce misperceptions but increase scepticism. *Nature Human Behaviour*, 1–9. <https://doi.org/10.1038/s41562-024-01884-x>
- Hruschka, T. M. J., & Appel, M. (2023). Learning about informal fallacies and the detection of fake news: An experimental intervention. *PLOS ONE*, 18(3), e0283238. <https://doi.org/10.1371/journal.pone.0283238>
- Juez, L. A., & Mackenzie, J. L. (2019). Emotion, lies, and “bullshit” in journalistic discourse: *Ibérica*, 38, Article 38.
- Kim, J. W. (2018). They liked and shared: Effects of social media virality metrics on perceptions of message influence and behavioral intentions. *Computers in Human Behavior*, 84, 153–161. <https://doi.org/10.1016/j.chb.2018.01.030>
- Kobayashi, K. (2018). The Impact of Perceived Scientific and Social Consensus on Scientific Beliefs. *Science Communication*, 40(1). <https://doi.org/10.1177/1075547017748948>
- Lee, S. S., Liang, F., Hahn, L., Lane, D. S., Weeks, B. E., & Kwak, N. (2021). The Impact of Social Endorsement Cues and Manipulability Concerns on Perceptions of News Credibility. *Cyberpsychology, Behavior and Social Networking*, 24(6), 384–389. <https://doi.org/10.1089/cyber.2020.0566>
- Lee, T. (2019). The global rise of “fake news” and the threat to democratic elections in the USA. *Public Administration and Policy*, 22(1), 15–24. <https://doi.org/10.1108/PAP-04-2019-0008>
- Lewandowsky, S., Cook, J., Fay, N., & Gignac, G. (2019). The effects of social cues on the perception of fake news. *Journal of Experimental Social Psychology*, 81, 1–10. <https://doi.org/10.1016/j.jesp.2019.05.006>

- G. E. (2019). Science by social media: Attitudes towards climate change are mediated by perceived social consensus. *Memory & Cognition*, 47(8), 1445–1456. <https://doi.org/10.3758/s13421-019-00948-y>
- Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369. <https://doi.org/10.1016/j.jarmac.2017.07.008>
 - Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest, Supplement*, 13(3), 106–131. <https://doi.org/10.1177/1529100612451018>
 - Lewandowsky, S., & van der Linden, S. (2021). Countering Misinformation and Fake News Through Inoculation and Prebunking. *European Review of Social Psychology*, 32(2), 348–384. <https://doi.org/10.1080/10463283.2021.1876983>
 - Lewandowsky, S., & Yesilada, M. (2021). Inoculating against the spread of Islamophobic and radical-Islamist disinformation. *Cognitive Research: Principles and Implications*, 6(1), 57. <https://doi.org/10.1186/s41235-021-00323-z>
 - Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, 5(3), 337–348. <https://doi.org/10.1038/s41562-021-01056-1>
 - Lu, C., Hu, B., Li, Q., Bi, C., & Ju, X.-D. (2023). Psychological inoculation for credibility assessment, sharing intention, and discernment of misinformation: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 25, e49255. <https://doi.org/10.2196/49255>
 - Mackie, D. M., Worth, L. T., & Asuncion, A. G. (1990). Processing of persuasive in-group messages. *Journal of Personality and Social Psychology*, 58(5), 812–822. <https://doi.org/10.1037/0022-3514.58.5.812>
 - Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology: Applied*, 27(1), 1–16. <https://doi.org/10.1037/xap0000315>
 - Maleki, M., Mead, E., Arani, M., & Agarwal, N. (2021, March 22). *Using an Epidemiological Model to Study the Spread of Misinformation during the Black Lives Matter Movement*. arXiv.Org. <https://arxiv.org/abs/2103.12191v1>
 - Malone, T. W., & Bernstein, M. S. (2022). *Handbook of Collective Intelligence*. MIT Press.
 - Martel, C., Allen, J., Pennycook, G., & Rand, D. G. (2024). Crowds Can Effectively Identify Misinformation at Scale. *Perspectives on Psychological Science*, 19(2), 477–488. <https://doi.org/10.1177/17456916231190388>
 - Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 47. <https://doi.org/10.1186/s41235-020-00252-3>
 - Mason, A. M., Compton, J., Tice, E., Peterson, B., Lewis, I., Glenn, T., & Combs, T. (2023). Analyzing the Prophylactic and Therapeutic Role of Inoculation to Facilitate Resistance to Conspiracy Theory Beliefs. *Communication Reports*, 0(0), 1–15. <https://doi.org/10.1080/08934215.2023.2256803>
 - McGarty, C., Haslam, S. A., Hutchinson, K. J., & Turner, J. C. (1994). The effects of salient group memberships on persuasion. *Small Group Research*, 25(2), 267–293. <https://doi.org/10.1177/1046496494252007>
 - McGuire, W. J. (1961). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. *The Journal of Abnormal and Social Psychology*, 63(2), 326–332. <https://doi.org/10.1037/h0048344>
 - McGuire, W. J. (1964). Inducing resistance to

persuasion: Some contemporary approaches. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 1, pp. 191–229). Academic Press.

[https://doi.org/10.1016/S0065-2601\(08\)60052-0](https://doi.org/10.1016/S0065-2601(08)60052-0)

- Moussaïd, M., Kämmer, J. E., Analytis, P. P., & Neth, H. (2013). Social Influence and the Collective Dynamics of Opinion Formation. *PLOS ONE*, 8(11), e78433. <https://doi.org/10.1371/journal.pone.0078433>
- Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative Social Influence is Underdetected. *Personality and Social Psychology Bulletin*, 34(7), 913–923. <https://doi.org/10.1177/0146167208316691>
- Paschen, J. (2019). Investigating the emotional appeal of fake news using artificial intelligence and human contributions. *Journal of Product & Brand Management*, 29(2), 223–233. <https://doi.org/10.1108/JPBM-12-2018-2179>
- Peng, W., Lim, S., & Meng, J. (2023). Persuasive strategies in online health misinformation: A systematic review. *Information, Communication & Society*, 26(11), 2131–2148. <https://doi.org/10.1080/1369118X.2022.2085615>
- Pereira, A., Harris, E., & Van Bavel, J. J. (2023). Identity concerns drive belief: The impact of partisan identity on the belief and dissemination of true and false news. *Group Processes & Intergroup Relations*, 26(1), 24–47. <https://doi.org/10.1177/13684302211030004>
- Pertwee, E., Simas, C., & Larson, H. J. (2022). An epidemic of uncertainty: Rumors, conspiracy theories and vaccine hesitancy. *Nature Medicine*, 28(3), Article 3. <https://doi.org/10.1038/s41591-022-01728-z>
- Roozenbeek, J., & van der Linden, S. (2020). Breaking Harmony Square: A game that “inoculates” against political misinformation. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-47>
- Roozenbeek, J., Traberg, C. S., & van der Linden, S. (2022). Technique-based inoculation against real-world misinformation. *Royal Society Open Science*, 9(5), 211719. <https://doi.org/10.1098/rsos.211719>
- Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online mis-information. *Palgrave Communications*, 5(1), Article 1. <https://doi.org/10.1057/s41599-019-0279-9>
- Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, 8(34), eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- Said, N., Fischer, H., & Anders, G. (2022). Contested science: Individuals with higher metacognitive insight into interpretation of evidence are less likely to polarize. *Psychonomic Bulletin & Review*, 29(2), 668–680. <https://doi.org/10.3758/s13423-021-01993-y>
- Schulster, J. R. (2006). Things we talk about, how frequently, and to whom: Frequency of topics in everyday conversation as a function of gender, age, and marital status. *The American Journal of Psychology* 1 October 2006; 119 (3), 407–432. <https://doi.org/10.2307/20445351>
- Sherrick, B. (2016). The Effects of Media Effects: Third-Person Effects, the Influence of Presumed Media Influence, and Evaluations of Media Companies. *Journalism & Mass Communication Quarterly*, 93(4), 906–922. <https://doi.org/10.1177/1077699016637108>
- Spampatti, T., Hahnel, U. J. J., Trutnevyte, E., & Brosch, T. (2023). Psychological inoculation strategies to fight climate disinformation across 12 countries. *Nature Human Behaviour*, 1–19. <https://doi.org/10.1038/s41562-023-01736-0>
- Spears, R. (2021). Social Influence and Group Identity. *Annual Review of Psychology*, 72(1),

- 367–390. <https://doi.org/10.1146/annurev-psych-070620-111818>
- Tormala, Z. L., DeSensi, V. L., Clarkson, J. J., & Rucker, D. D. (2009). Beyond attitude consensus: The social context of persuasion and resistance. *Journal of Experimental Social Psychology*, 45(1), 149–154. <https://doi.org/10.1016/j.jesp.2008.07.020>
 - Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLOS ONE*, 13(9), e0203958. <https://doi.org/10.1371/journal.pone.0203958>
 - Traberg, C. S. (2024). Coercion by Misinformation: Challenges and Solutions. In *Coercion and Trust: A Multi-Disciplinary Dialogue* (Vol. 1). Routledge.
 - Traberg, C. S., Harjani, T., Basol, M., Biddlestone, M., Maertens, R., Roozenbeek, J., & van der Linden, S. (2023). Prebunking Against Misinformation in the Modern Digital Age. In T. D. Purnat, T. Nguyen, & S. Briand (Eds.), *Managing Infodemics in the 21st Century: Addressing New Public Health Challenges in the Information Ecosystem*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-27789-4>
 - Traberg, C. S., Harjani, T., Roozenbeek, J., & van der Linden, S. (2024). The persuasive effects of social cues and source effects on misinformation susceptibility. *Scientific Reports*, 14(1), Article 1. <https://doi.org/10.1038/s41598-024-54030-y>
 - Traberg, C. S., Roozenbeek, J., & van der Linden, S. (2022). Psychological Inoculation against Misinformation: Current Evidence and Future Directions. *The ANNALS of the American Academy of Political and Social Science*, 700(1). <https://doi.org/10.1177/00027162221087936>
 - Traberg, C. S., Roozenbeek, J., & van der Linden, S. (2024). Gamified inoculation reduces susceptibility to misinformation from political ingroups. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-141>
 - Treen, K. M. d. I., Williams, H. T. P., & O'Neill, S. J. (2020). Online misinformation about climate change. *WIREs Climate Change*, 11(5), e665. <https://doi.org/https://doi.org/10.1002/wcc.665>
 - Van Bavel, J. J., & Pereira, A. (2018). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences*, 22(3), 213–224. <https://doi.org/https://doi.org/10.1016/j.tics.2018.01.004>
 - van der Linden, S. (2022). Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, 28(3), 460–467. <https://doi.org/10.1038/s41591-022-01713-6>
 - van der Linden, S., Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the Public against Misinformation about Climate Change. *Global Challenges*, 1(2), 1600008. <https://doi.org/https://doi.org/10.1002/gch2.201600008>
 - van der Meer, T. G. L. A., Hameleers, M., & Ohme, J. (2023). Can Fighting Misinformation Have a Negative Spillover Effect? How Warnings for the Threat of Misinformation Can Decrease General News Credibility. *Journalism Studies*, 24(6), 803–823. <https://doi.org/10.1080/1461670X.2023.2187652>
 - Walker, M., & Matsa, K. E. (2021). *News Consumption Across Social Media in 2021: More than half of Twitter users get news on the site regularly*. Pew Research Center. <http://www.jstor.org/stable/resrep63516>