Received: 26 July 2023 | Revised: 14 May 2024 | Accepted: 16 May 2024
DOI: 10.1111/jasp.13049

ORIGINAL ARTICLE

Journal of Applied Social Psychology WILEY

Misinformation interventions decay rapidly without an immediate posttest

Georgia Capewell[1] | **Rakoen Maertens**[2] | **Miriam Remshard**[1] | **Sander van der Linden**[1] | **Josh Compton**[3] | **Stephan Lewandowsky**[4,5] | **Jon Roozenbeek**[1,6]

[1]Department of Psychology, University of Cambridge, Cambridge, UK
[2]Department of Experimental Psychology, University of Oxford, Oxford, UK
[3]Speech at Dartmouth, Dartmouth College, Hanover, New Hampshire, USA
[4]School of Psychological Science, University of Bristol, Bristol, UK
[5]Department of Psychology, University of Potsdam, Potsdam, Germany
[6]Department of War Studies, King's College London, London, UK

Correspondence
Jon Roozenbeek, Department of Psychology, University of Cambridge, Cambridge, UK.
Email: jjr51@cam.ac.uk

Funding information
British Academy, Grant/Award Number: #PF21\210010; Department of Psychology at the University of Cambridge, and the American Psychological Association together with the Centers for Disease Control and Prevention, Grant/Award Number: #6NU87PS004366-03-02; European Research Council (ERC Advanced Grant), Grant/Award Number: 101020961 PRODEMINFO; Humboldt Foundation through a research award, and from UK Research and Innovation (EU Horizon replacement funding), Grant/Award Number: 10049415

Abstract

In recent years, many kinds of interventions have been developed that seek to reduce susceptibility to misinformation. In two preregistered longitudinal studies (N_1 = 503, N_2 = 673), we leverage two previously validated "inoculation" interventions (a video and a game) to address two important questions in misinformation interventions research: (1) whether displaying additional stimuli (such as videos unrelated to misinformation) alongside an intervention interferes with its effectiveness, and (2) whether administering an immediate posttest (in the form of a social media post evaluation task after the intervention) plays a role in the longevity of the intervention. We find no evidence that other stimuli interfere with intervention efficacy, but strong evidence that immediate posttests strengthen the learnings from the intervention. In study 1, we find that 48 h after watching a video, participants who received an immediate posttest continued to be significantly better at discerning untrustworthy social media posts from neutral ones than the control group (d = 0.416, p = .007), whereas participants who only received a posttest 48 h later showed no differences with a control (d = 0.010, p = .854). In study 2, we observe highly similar results for a gamified intervention, and provide evidence for a causal mechanism: immediate posttests help strengthen people's memory of the lessons learned in the intervention. We argue that the active rehearsal and application of relevant information are therefore requirements for the longevity of learning-based misinformation interventions, which has substantial implications for their scalability.

Georgia Capewell and Rakoen Maertens contributed equally to this study.

1 | INTRODUCTION

The spread of and belief in misinformation is a substantial societal problem (Lewandowsky et al., 2017; Roozenbeek & van der Linden, 2024), having been linked to a range of negative public health outcomes (Schmid et al., 2023), including reduced vaccine uptake (Loomba et al., 2021, 2024). Proposed approaches to countering misinformation include both system-level solutions (such as introducing new legislation or changing recommender algorithms) and individual-level solutions (boosting competences, nudging people to promote better sharing behavior, debunking/fact-checking, and content labeling); see Kozyreva et al. (2024) and Ecker et al. (2022) for discussions. In recent years, *prebunking* (pre-emptive debunking) and psychological inoculation have gained prominence as a pre-emptive individual-level method of tackling misinformation (Lewandowsky & van der Linden, 2021).

Prebunking is an umbrella term for various approaches that seek to protect people against misinformation. Inoculation theory is a commonly used form of prebunking. Grounded in a medical analogy, inoculation interventions commonly consist of two components, which when combined should induce increased psychological resistance against future unwanted persuasion attempts: (1) a warning of an impending attack on one's beliefs (i.e., a forewarning), and (2) a pre-emptive refutation of the false or misleading argument, narrative, or manipulation technique that participants are "inoculated" against (Compton, 2013; Compton et al., 2021; McGuire, 1964; McGuire & Papageorgis, 1962). Inoculation interventions such as infographics, informative videos, and online games have shown promise as a way to build resistance against misinformation at scale (Basol et al., 2021; Cook et al., 2022; Lewandowsky & Yesilada, 2021).

Recent inoculation studies share a similar study design, with a posttest (usually in the form of an item rating task, e.g., rating the perceived reliability of a series of true and false/misleading social media posts) administered immediately after the treatment or control intervention.[1] In fact, the majority of contemporary studies testing the efficacy of misinformation interventions (not just using inoculation) involve immediate posttesting (Johansson et al., 2022; Kozyreva et al., 2024; Ziemer & Rothmund, 2022). And although some longitudinal studies are available (Basol et al., 2021; Brashier et al., 2021; Maertens et al., 2021; Swire et al., 2017),[2] most of these studies have involved participants going through an intervention, then immediately taking a posttest, and then taking another posttest some time after.

This study design is potentially problematic. In a recent study using three types of inoculation interventions (text, video, and game), Maertens et al. (2024) tested whether administering a posttest immediately after the intervention yields similar outcomes (in terms of the strength and longevity of the inoculation effect) as only administering the posttest after ~10, ~12, or ~30 days. They found that participants who watched an inoculation video and did a posttest immediately after were significantly better than a control group at discerning misinformation from nonmisinformation ~10 days later, with no evidence for decay. However, participants who *only* received a posttest after ~12 days (and not immediately) showed evidence for decay and did not perform significantly better than control group participants. Similar results were found for the game-based intervention, which was no longer effective after ~10 or ~30 days if no immediate posttest was administered, despite a previous study showing good longevity (up to 2 months) with regular testing (Maertens et al., 2021). In other words, there are indications that immediate posttests help strengthen the lessons from the intervention in memory, thus boosting the strength and longevity of the intervention's efficacy. This is in line with the broader educational psychology literature, which has established the importance of retrieval practice for durable learning and memory (Agarwal et al., 2021; Carpenter et al., 2022). Intermittent testing of memory has been found to be more effective for overall learning than providing additional study opportunities across more than 200 studies (Carpenter et al., 2022). The testing effect constitutes one of the most robust findings in the psychology of learning and education, arising in nearly all conditions in school and classroom settings (Agarwal et al., 2021). And indeed, Bird et al. (2015) found that people who actively rehearsed the contents of a video were substantially better at recall 8 and 18 days after viewing compared to a group that did no rehearsal.[3] Even in the context of debunking approaches to misinformation, retrieval practice has been shown to improve memory of corrective information, although this improved memory does not necessarily translate to more accurate beliefs (Collier et al., 2023; Fazio et al., 2023).

Within the inoculation theory literature, Compton and Ivanov (2012) found that interrupting an inoculation treatment (in their study, after the forewarning but before the preemptive refutational content) to measure threat seemed to enhance the resistance conferred by the inoculation treatment, when compared to interventions that did not have this interruption. They note that "the interruption of the experimental design and the embedding of a threat scale in the inoculation message contributed to greater levels of threat and ultimately greater resistance (i.e., more negative attitudes toward the counterattitudinal attack)" (Compton & Ivanov, 2012, p. 10). More broadly, these findings raise the question of whether posttests (or other memory strengthening exercises) should be seen as a component of many types of learning-based antimisinformation interventions, which has substantial implications for their scalability.

In addition, the vast majority of studies into the efficacy of misinformation interventions have looked at how these interventions perform in isolation (Johansson et al., 2022). We therefore know very little about the extent to which distractions (such as watching one or more irrelevant videos alongside a video designed to help people better identify misinformation) reduce people's performance on subsequent relevant outcome measures. In their YouTube field study, Roozenbeek et al. (2022) administered their outcome measure a median of 18 h after participants watched their intervention as a YouTube ad; it is likely that participants were distracted by other stimuli during this time period, but they did not measure distraction explicitly. Previous research has generally shown that distractions are detrimental to memory (Blasiman et al., 2018; Craik, 2014; Middlebrooks et al., 2017). It is therefore possible that results from studies which only look at the efficacy of (learning-based) misinformation interventions in isolation overestimate this efficacy when distracting stimuli are taken into account, for instance in social media environments where distractions are omnipresent (Pennycook & Rand, 2019).

We test both of these possibilities in two preregistered experiments ($N_1 = 503$ and $N_2 = 673$), one with a video-based and one with a game-based intervention. All materials including data sets, cleaning, analysis, and visualization scripts (written in R), Qualtrics files, and other supplementary information can be found on our OSF page: https://osf.io/9frch/.

2 | STUDY 1

In study 1 ($N = 503$, preregistration link: https://aspredicted.org/W92_MDY), we look at whether a previously validated video-based inoculation intervention (Roozenbeek et al., 2022) retains its effectiveness with and without an immediate posttest after 48 h. In addition, we test if viewing unrelated content alongside the inoculation intervention (two other videos about random topics) interferes with the intervention's efficacy. Briefly put, while we find that there is no interference with unrelated content (which is good news as in the real world people are exposed to a large number of stimuli alongside any potential intervention), administering an immediate posttest matters a great deal. Participants who only received a posttest 48 h after viewing an inoculation video performed no better than a control group, whereas participants who did get an immediate posttest continued to substantially outperform the control group after 48 h.

2.1 | Methods

2.1.1 | Measures

We conducted a preregistered randomized controlled trial on Prolific Academic using a 30-s long video-based inoculation intervention about emotionally manipulative language (see: https://inoculation.science/inoculation-videos/emotional-language/) as the intervention of interest. Emotionally manipulative language is a technique commonly used in misinformation (Carrasco-Farré, 2022; Zollo et al., 2015). Thus, in previous studies, this video-based inoculation intervention was shown to substantially improve people's ability to discern emotionally manipulative from neutral social media content, increase confidence in their ability to identify misinformation that utilizes this technique, reduce the perceived trustworthiness of manipulative content, and boost the quality of their news sharing decisions.

The inoculation effect in the present study was measured through a posttest, that is, an item rating task consisting of 10 social media posts, identical to how the intervention was tested in previous studies (Maertens et al., 2024; Roozenbeek et al., 2022); see Supporting Information S1: Table 7 for an overview of the item wordings. Each social media post was randomly selected to be either manipulative (i.e., using emotionally evocative language) or neutral (similar in content, length, and topic as the manipulative post but not making use of emotionally evocative language), so that each participant saw on average five manipulative and five neutral items (although this proportion varied per participant). Participants rated on a Likert scale, from 1 to 7 (1 = *strongly disagree*, 7 = *strongly agree*) the following: (1) Manipulativeness ("This post is manipulative"); (2) Confidence in their judgment ("I am confident about my assessment of this post's manipulativeness"); (3) Trustworthiness ("This post is trustworthy"); and (4) Sharing intentions ("I would share this post with people in my network"). For the manipulativeness, trustworthiness, and sharing intentions measures, our key outcome variable of interest is *discernment*, calculated by taking the difference between the average scores for the neutral social media posts and the average scores for their manipulative alternatives. Positive discernment scores thus mean that a participant is generally able to correctly distinguish manipulative from neutral content, whereas negative discernment scores mean that a participant incorrectly identifies neutral content as more manipulative (or less trustworthy, etc.) than manipulative content. For the confidence measure, we look at ratings separately for the manipulative items and neutral items. See Supporting Information S1: Table 0b for descriptive statistics. For an extensive explanation of the development and validation of this item rating task, we refer to Roozenbeek et al. (2022).

To test for interference and posttest effects, we included four separate experimental conditions, two inoculation groups and two control groups. The two inoculation groups were both shown the inoculation video as well as two out of three possible control videos of similar length and aesthetic but unrelated to emotionally manipulative language (about how to make blueberry muffins; the importance of vaccines[4]; and age-related macular degeneration). The two control groups were shown all three control videos. Video display order was randomized at the participant level.

We administered the same posttest at two different time points: immediately after watching the videos (**T1**) and/or 48 h after (**T2**). Group 1 was shown the inoculation video (alongside 2 control videos), then did an immediate posttest, and then did the same posttest again 48 h after. Group 2 was the same as group 1, except they did not do an immediate posttest, only after 48 h. Group 3 was the same as group 1 except participants saw only three control videos rather than an inoculation video. Participants in group 4 were also only shown control videos but—as with group 2—they did not complete an immediate posttest.

We also administered a series of demographic questions: gender, age, political ideology (1 being "very left-wing" and 7 being "very right-wing"), people's political positions on social and economic issues, education level, ethnicity, how often people check the news, social media use, and their favorite news outlet. Participants received an attention check which assessed if they remembered the videos they watched. They were given six options in total, four of which they might have seen as part of the study, and two not being part of the study. As preregistered, those who failed the attention check (by selecting one or both of the videos that were not in the study) were excluded from the final sample. This resulted in 47 participants being removed from the analysis. Participants who took the survey more than once or who did not complete the survey at **T2** were also excluded. This study received ethical approval by the University of Cambridge Psychology Research Ethics Committee (#PRE.2022.113). See Figure 1 for an overview of the design of studies 1 and 2.

2.1.2 | Sample

We conducted an a priori power analysis to determine sample size, using the pwr package in R (power = 95%, α = .05 and d = 0.400, based on the d = 0.439 effect size found for manipulativeness discernment in Maertens et al., 2024). The minimum sample needed to detect the presence of a main effect of inoculation of that size on manipulativeness discernment was N = 680. As preregistered, we collected a total of 800 initial participants at **T0** (~200 responses per group), to account for an attrition rate of 30%. Participant recruitment was limited to the United States and balanced on sex (400 male, 400 female). Participants were paid 1.00 GBP for a 6–7 min survey (~£9 per hour), followed by £0.46 for a 4–5 min follow-up survey (~£6 per hour). After applying our preregistered exclusion criteria, we ended up with a final sample of N = 503, below the target (see Figure 1 for the sample sizes per group). The attrition rate between **T1** and **T2** was 32.4%. The final sample was 50.7% female,

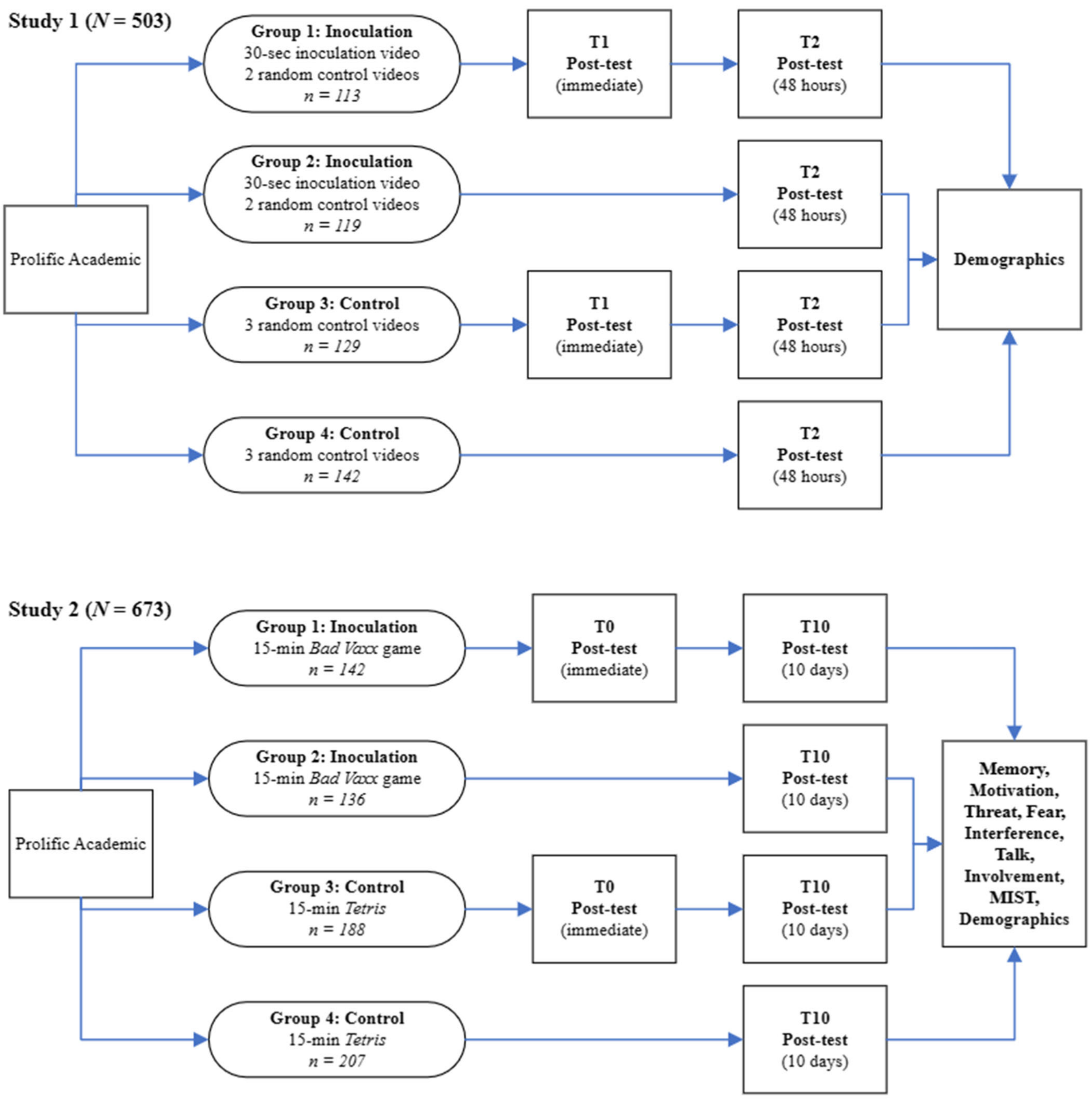

FIGURE 1 Overview of the design of studies 1 and 2.

48.1% male, and 1.0% nonbinary. The average age was 44.1 years (SD = 13.5), 78.1% identified as white (21.9% as BAME; Black, Asian, or Minority Ethnic) and the most common political stance was Left (52.5%), followed by right (25.2%) and center (22.3%). The most frequent level of education was a Bachelor's degree (37.8%), followed by a high school degree (28.8%), a Graduate/Professional degree (15.8%), an Associate's degree (13.7%), a Doctorate (3.4%), and no high school degree (0.6%). Participants were somewhat left-leaning on social issues ($M = 3.22$, $SD = 1.81$ on a 7-point scale) but more centrist on economic issues ($M = 3.65$, $SD = 1.87$). 38.6% of participants reported checking the news once a day, followed by 37.4% who checked it multiple times a day. 2.0% reported never checking the news. 64.8% reported checking social media multiple times a day, followed by 20.7% who check it once a day. 2.8% of participants reported never checking social media. See Supporting Information S1: Table 0a for a full overview.

2.1.3 | Hypotheses and analysis plan

We preregistered (https://aspredicted.org/W92_MDY) the following hypotheses:

> **H1.** People who watch a ~30 s inoculation video are better at differentiating emotionally manipulative social media posts from neutral social media posts, compared to those who watched a ~30 s control video, at **T1** (immediately after the intervention).

> **H1a.** Manipulativeness discernment.

> **H1b.** Confidence in manipulativeness judgment.

> **H1c.** Trustworthiness discernment.

> **H1d.** Sharing discernment.

> **H2.** The same effects can be found at **T2** (48 h after the intervention), independent of whether participants received an immediate posttest at **T1** [**H2a**] or not [**H2b**].

Finally, we preregistered the following exploratory hypotheses:

> **eH1.** People who receive an immediate posttest after watching a ~30 s inoculation [**T1**] are better at differentiating emotionally manipulative social media posts from neutral ones, compared to those who watched the inoculation video without an immediate posttest, when tested 48 h later [**T2**].

> **eH2.** The presentation of the inoculation video with two other unrelated control videos will produce a generally reduced inoculation effect, compared to previous studies which present only the inoculation video in isolation (Maertens et al., 2024; Roozenbeek et al., 2022).

We preregistered using ANOVA contrasts to test for between-group differences at **T1** (**H1a–d**) and **T2** (**H2a** and **H2b**). Specifically, we test the following:

> **H1a–d.** Discernment group 1 (**T1**) > discernment group 3 (**T1**).

> **H2a.** Discernment in group 1 (**T2**) > discernment group 3 (**T2**).

> **H2b.** Discernment in group 2 (**T2**) > discernment group 4 (**T2**).

2.2 | Results

2.2.1 | H1a–H1d: The inoculation effect at T1 (immediate posttest)

First, the inoculation video had a significant positive effect on manipulativeness discernment scores compared to the control video at the immediate posttest (**T1**), $F(1,240) = 17.89$, $p < .001$, $M_{diff} = 0.803$, $t(240) = 4.230$, $d = 0.545$, 95% CI [0.287, 0.804], in support of **H1a**. Second, the effect of inoculation on participants' confidence in their judgments of manipulative posts was significant and positive, $F(1,240) = 4.94$, $p = .027$, $M_{diff} = 0.317$, $t(240) = 2.222$, $d = 0.286$, 95% CI [0.031, 0.541], showing support for **H1b**. Although not preregistered, we found a lack of a significant effect of inoculation on confidence ratings of neutral posts, $F(1,240) = 2.42$, $p = .121$, $M_{diff} = 0.236$, $t(240) = 1.555$, $d = 0.200$, 95% CI [−0.054, 0.455]. Third, we found a significant positive effect of inoculation on trustworthiness discernment, $F(1,240) = 5.40$, $p = .021$ $M_{diff} = 0.377$, $t(240) = 2.324$, $d = 0.299$, 95% CI [0.044, 0.555], supporting **H1c**. Lastly, we found a nonsignificant effect of inoculation on a person's sharing discernment relative to controls, $F(1,240) = 1.02$, $p = .315$, $M_{diff} = −0.131$, $t(240) = 1.008$, $d = 0.130$, 95% CI [−0.124, 0.384]. We thus fail to confirm **H1d**, although this might be due to the study being underpowered to detect sharing effects which tend to be smaller (as our study was powered for $d = .400$). Broadly the same effects are found when participants who failed to complete the survey at **T2** are included in the analysis (see Supporting Information S1: Table 1b). Supplementary results (independent and Bayesian t tests) can be found in Supporting Information S1: Tables 1a and 1b.

2.2.2 | H2a–H2b: The inoculation effect at T2 (48 h) with and without immediate posttest

We tested for a main effect of inoculation on the manipulativeness, confidence, trustworthiness, and sharing measures at **T2** (48 h postintervention). For manipulativeness discernment, the omnibus test was significant, $F(3, 499) = 4.66$, $p = .003$, indicating that further preregistered contrasts can be tested. The Tukey-corrected contrasts show that the immediate posttest inoculation group (group 1) had significantly higher manipulativeness discernment at **T2** than the control group with immediate posttest (group 3), $M_{diff} = 0.628$, $t(240) = 3.226$, $p_{tukey} = .007$, $d = 0.416$, 95% CI [0.161, 0.670]. This supports **H2a**, indicating that participants who receive an immediate posttest continue to outperform a control group 48 h later.

However, when we compare the inoculation group that did not receive an immediate posttest (group 2) and the control group without immediate posttest (group 4), we find no difference between both groups at **T2**, $M_{diff} = 0.150$, $t(259) = 0.802$, $p_{tukey} = .854$, $d = 0.010$, 95% CI [−0.145, 0.344]. Furthermore, a (nonpreregistered) Two One-Sided (TOST) equivalence test with a Smallest Effect Size of Interest (SESOI) of $d = 0.100$ (very small) is not significant, $t(259) = 1.334$, $p = .092$, confirming statistical equivalence to zero, meaning that we can rule out the presence of any effects larger than $d = 0.100$ (Blasiman et al., 2018, 2020).[5] These results show that the group that received an immediate posttest continued to outperform the control group after 48 h, whereas the group without immediate posttest did not. These results speak against **H2b** and lend support to **eH1**. While the inoculation with immediate posttest group (group 1) did not perform significantly better at discerning manipulative content at **T2** than the inoculation group without the immediate posttest (group 2, $M_{diff} = 0.484$, $t(230) = 2.438$, $p_{Tukey} = .071$, $d = 0.320$, 95% CI [0.061, 0.579]) in absolute terms (as stipulated in **eH1**), they did perform better in relative terms, as the inoculation with posttest group displayed significantly stronger manipulativeness discernment at **T2** than the control groups, whereas the inoculation group with no immediate posttest was no better at discerning manipulative content than control groups at **T2**. In addition, an exploratory multiple linear regression for manipulativeness discernment at T2, but with demographic variables included as covariates (see below), did show that the inoculation with immediate posttest group (group 1) performed significantly better at discerning manipulative content than the inoculation group without immediate posttest (group 2), $p = .008$; see Supporting Information S1: Table 6. Thus, we conclude that without an immediate posttest, any inoculation effect disappeared within 48 h.

For trustworthiness discernment at **T2** the results were similar, with the immediate posttest inoculation group (group 1) also displaying significantly higher trustworthiness discernment at **T2** than controls with immediate posttest (group 3), $M_{diff} = 0.489$, $t(240) = 3.099$, $p_{Tukey} = .011$, $d = 0.399$, 95% CI [−0.654, −0.145]. Moreover, when no immediate posttest was presented, there was no significant difference between the control (group 4) and inoculation (group 2) groups at **T2**, $M_{diff} = 0.077$, $t(259) = 0.504$, $p_{Tukey} = .958$, $d = 0.063$, 95% CI [−0.307, 0.182].

However, there were no significant differences between the four groups for sharing discernment and confidence at **T2**. Similar results were reported by Maertens et al. (2024), who also noted a decay of inoculation effects on sharing intent and confidence, while the effect on trustworthiness remained significant over time. However, our results may also be due to the fact that our study was powered specifically to detect effects on manipulativeness discernment. See Figure 2 for boxplots and violin plots with data jitter. See Supporting Information S1: Tables 2–5 for the full results, including nonpreregistered Bonferroni-corrected p values (which do not change the results reported here). See Supporting Information S1: Figure 1 for a series of stacked density plots.

Next, we test exploratory hypothesis **eH2**. At **T1**, the between-group effect size for manipulativeness discernment is $d = 0.545$, descriptively *larger* than the effect size of $d = 0.439$ obtained by Maertens et al. (2024) and in the same realm as the effect sizes found in Roozenbeek et al. (2022) of $d = 0.49$ and $d = 0.67$ (this paper included two separate studies with the "emotional language" inoculation intervention). Effect sizes for the other three outcome measures are also descriptively similar to those found by Maertens et al. (2024), although we note that the effect for sharing discernment was not significant, potentially due to a lack of statistical power (see Supporting Information S1: Table 1). We therefore find no support for **eH2**: there appears to be no evidence that watching two unrelated videos alongside the inoculation video interferes with the inoculation effect. In other words, people retain the relevant information from the video even if they are exposed to stimuli containing entirely different information around the same time.

Finally, we conduct a series of exploratory analyses with our demographic variables. Using a multiple linear regression with manipulativeness discernment at **T2** as the outcome measure, experimental condition as the independent variable, and age, gender, education level, political ideology, ethnicity, how often people check the news, and social media use as covariates, we find that the main effect of inoculation with immediate posttest on manipulativeness discernment remains significant (comparing the inoculation with immediate posttest group—group 1—against the control group with immediate posttest—group 3), standardized $\beta = -.424$, $p < .001$). In addition, we find a significant effect of ethnicity ($\beta = .263$, $p = .017$) such that in comparison with people who identify as white, BAME (Black, Asian, and Minority Ethnic) people have lower manipulativeness discernment; political ideology ($\beta = -.180$, $p < .001$) so that Conservatives have lower manipulativeness discernment than Liberals; and a small but significant positive effect of education ($\beta = .090$, $p = .045$), with people with higher levels of education also having higher manipulativeness discernment. None of the other covariates show significant associations with manipulativeness discernment at **T2** (all $p > .131$). See Supporting Information S1: Table 6.

2.3 | Discussion

Using a previously validated video-based inoculation intervention, this study sought to address two open questions in

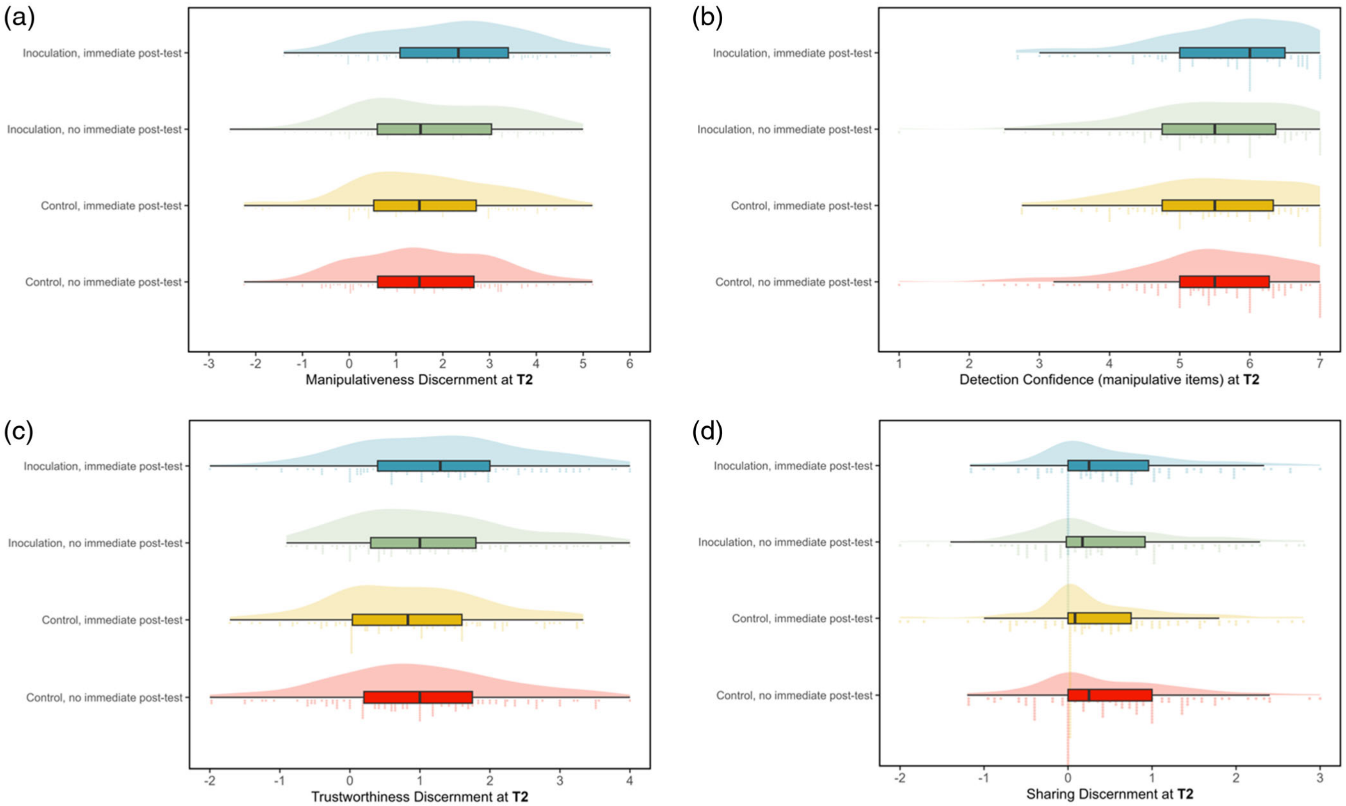

FIGURE 2 Boxplots with distribution and data jitter for the four outcome measures at **T2** (48 h after the intervention): manipulativeness discernment (a), confidence (b), trustworthiness discernment (c), and sharing discernment (d).

misinformation interventions research: (1) whether providing additional, irrelevant stimuli alongside an intervention interferes with its effectiveness, and (2) whether posttests administered immediately after the intervention play a role in the longevity of these effects. We find no evidence that additional stimuli interfere with the inoculation effect conferred by the video, with the inoculation effect not only being highly significant when interference is present, but also with a similar (and descriptively even higher) effect size.

However, in line with earlier work by Maertens et al. (2024), we also find strong evidence that testing effects play a substantial role in even a highly effective intervention's longevity: participants who were not given an immediate posttest just after watching an inoculation video showed no improved discernment of manipulativeness compared to a control group a mere 48 h after viewing. As the relevant p value for this difference was about $p = .854$ and we were able to rule out the possibility of any effects larger than $d = 0.10$, it is unlikely that this null finding is due to lack of power, but rather reflects a true "forgetting" of the lessons from the video. Nonetheless, this study also had several limitations: we only looked at a single intervention (a short video), which limits the generalizability of our findings, and more importantly, we were not able to establish a causal mechanism behind *why* decay sets in more rapidly without an immediate posttest. We address these limitations in study 2.

3 | STUDY 2

For study 2 (preregistration link: https://aspredicted.org/blind.php?x=Q6P_C88), 1066 US participants were recruited through Prolific Academic with representative interlocking hard quota for sex, age, and ethnicity (based on simplified US Census Bureau population group estimates from 2015), of which 964 completed the survey at **T0**, with 673 responses remaining after applying all preregistered exclusion criteria. Study 2 was similar to study 1 in scope and design (see Figure 1), but with two major changes: (1) the time delay was 10 days instead of 48 h, and (2) we used a game-based instead of a video-based inoculation intervention. Specifically, inoculation group participants played *Bad Vaxx*, a 10-min inoculation game about vaccine misinformation. Over the course of four levels, players are tasked with "defeating" four characters (Ann McDotal, Dr Forge, Ali Natural, and Mystic Mac), each of which represents a different manipulation technique that previous studies have identified as being commonly used in vaccine misinformation (Kata, 2010, 2012): emotional storytelling, fake expertise/pseudoscience, the naturalistic fallacy, and conspiratorial reasoning. Control group participants played *Tetris*. In a different study (Appel et al., 2024), *Bad Vaxx* was shown to substantially improve people's ability to discern vaccine misinformation from neutral social media content about vaccines; we consider the present study in part a replication of this work (with minor design changes, explained below). In the interest of brevity, we refer to Appel et al. (2024) for more details about the game's design and testing.

3.1 | Measures, sample, and hypotheses

Similar to study 1, we administered a posttest either immediately postintervention (**T0**) or 10 days after (**T10**), or both, resulting in four conditions: inoculation and control, with and without an immediate posttest. Based on a power calculation with 90% power, $\alpha = .05$, and $d = 0.31$ (based on findings by Appel et al., 2024), we sought to recruit a total of 960 participants, or 240 per group. 964 complete responses were collected (which required recruiting 1066 initial participants). They were paid 2.41 GBP (~£6 per h) for a 24-min survey, followed by ~1.61 GBP (~£6 per h) for the ~16-min follow-up survey. After applying our exclusion criteria, we ended up with a final sample of $N = 842$ at **T0** and $N = 673$ participants who completed the study at both **T0** and **T10**, a total attrition rate of 20.1% ($n = 207$ for control without immediate posttest, $n = 188$ for control with posttest, $n = 136$ for inoculation without posttest, and $n = 142$ for inoculation with posttest). Attrition rates were higher in both inoculation groups, likely because these participants were required to provide a password at the end of the game to ascertain that they had played the game all the way through (unlike participants who played *Tetris*); some participants did complete the entire study but did not provide the correct password and were thus excluded from the main analysis. Study 2 received ethical approval by the University of Cambridge Psychology Research Ethics Committee (#PRE.2022.120).

The posttest was identical to the one used by Appel et al. (2024) and consisted of 12 simulated social media posts, each of which was randomly either manipulative (i.e., using one of the four techniques learned in the game) or neutral (similar in content, length, and topic as the manipulative post but not making use of any manipulation techniques from the game), so that each participant sees on average six manipulative and six neutral items; this procedure is highly similar to the one used in study 1. We included three outcome measures per item: manipulativeness, confidence, and sharing intentions, assessed on the same 1–7 Likert scale used in study 1. See Supporting Information S1: Table 13 for the item wordings. Again following Appel et al. (2024), we also included a separate "manipulation technique recognition" task consisting of 16 fictitious vaccine-related headlines, balanced on two dimensions: true versus false, and manipulative versus not manipulative. For an extensive discussion on the differences between these dimensions (and susceptibility to *false vs. misleading/manipulative* news), we refer to Maertens, Said, et al. (2023). Four of the headlines in this task were both false and manipulative, four were false and not manipulative (i.e., merely containing a false claim without further markers or cues of manipulativeness, e.g., "the flu vaccine doesn't work"), four were true and manipulative, and four were true and not manipulative. Participants were asked to rate each item as either true or false and as manipulative or nonmanipulative, and indicated which (if any) manipulation technique they believed the item contained from a drop-down menu, with only one option being correct. See Supporting Information S1: Table 8b for descriptive statistics (means and standard deviations per experimental group) and Supporting Information S1: Table 14 for the item wordings.

In addition, to investigate *why* decay may set in more quickly if no immediate posttest is administered, we included a battery of questions at the end of the survey (shown to all participants), assessing how well people remembered relevant information from the *Bad Vaxx* game ("memory"; e,g., "How many different misinformation techniques did you learn about in the game?"; control group participants are expected to perform at or near chance level as they are randomly guessing), how much they had come across content about vaccine misinformation in the past 2 weeks ("interference"; e.g., "In the past two weeks, I have heard conflicting arguments about vaccine misinformation [1 *strongly disagree*—7 *strongly agree*]"), how afraid they were of vaccine misinformation ("fear"; e.g., "Thinking about vaccine misinformation I feel fearful [1 *strongly disagree*—7 *strongly agree*]"), how motivated they were to resist vaccine misinformation ("motivation"; e.g., "Thinking about vaccine misinformation motivates me to resist misinformation [1 *strongly disagree*—7 *strongly agree*]"), how much they talked about vaccine misinformation with others ("talk"; e.g., "In the past two weeks, how often did you talk about or discuss vaccine misinformation with other people?"), how much vaccine misinformation means to them ("involvement"; e.g., which [term] best describes how much misinformation means to you?"), and how much they felt threatened by vaccine misinformation ("threat"; a composite measure consisting of the combined scores on the "fear," "motivation," "involvement," "talk," "apprehensive threat," and "accessibility"[6] scales)[7]; see Supporting Information S1: Table 12 for the question wordings. These questions were the same as those used by Maertens et al. (2024); the motivational threat scale was taken from Banas & Richards (2017), and the apprehensive threat scale from Compton & Ivanov (2012). These measures allow us to explore the relative importance of memory, motivation to resist misinformation, and the extent to which people feel threatened by misinformation in explaining the intervention effect's decay over time. Finally, we administered the 20-item Misinformation Susceptibility Test (Maertens, Götz, et al., 2023), COVID-19 vaccine uptake intentions, vaccine safety perceptions (see Supporting Information S1: Table 12), and a set of demographic questions: age, gender, education level, US state (and region), political ideology and US political party affiliation, social media use, and news consumption. For the sample descriptives, we refer to Supporting Information S1: Table 8a. We preregistered the following hypotheses (main effects tested using ANOVA contrasts, as in study 1):

> **H1.** Playing a short online game about vaccine misinformation improves people's ability to detect manipulative social media content (discernment and technique identification) about vaccinations.

H2. The intervention effects decay faster over time when no immediate posttest is administered compared to when an immediate posttest is administered.

H3. The longevity of the intervention effects is predicted by memory and motivation.

3.2 | Results

We show the results for our preregistered analyses for **H1** and **H2** in Table 1.

To summarize, we find support for **H1**: immediately after the intervention (**T0**), inoculation group participants were significantly better

TABLE 1 Study 2: Between-group comparisons (ANOVA contrasts) at **T0** (immediately postintervention) and **T10** (10 days after).

	Omnibus test		Post hoc comparisons						
	F	*p* Value	M_{diff}	*df*	*t*	*p* Value	p_{Tukey}	Cohen's *d*	95% CI
H1: INOCULATION vs. CONTROL|**T0 (IMMEDIATELY POSTINTERVENTION)**									
Inoculation effect									
Manipulativeness discernment	15.00	**<.001**	0.722	328	3.880	**<.001**	**<.001**	0.431	[0.210, 0.652]
Confidence (misinformation items)	14.20	**<.001**	0.467	328	3.760	**<.001**	**<.001**	0.419	[0.197, 0.640]
Confidence (nonmisinformation items)	5.75	**.017**	0.317	328	2.400	**.017**	**.017**	0.267	[0.047, 0.486]
Sharing discernment	18.40	**<.001**	0.630	328	4.290	**<.001**	**<.001**	0.477	[0.255, 0.698]
Manipulation technique recognition									
Manipulative vs. not manipulative	7.15	**.008**	0.035	328	2.670	**.008**	**.008**	0.297	[0.076, 0.517]
True vs. false	5.61	**.018**	0.040	328	2.370	**.018**	**.018**	0.263	[0.044, 0.483]
Technique identification	5.06	**.025**	0.038	328	2.250	**.025**	**.025**	0.250	[0.031, 0.470]
H2: INOCULATION WITHOUT POSTTEST AT T0 vs. CONTROL WITHOUT POSTTEST AT T0|**T10 (~10 DAYS LATER)**									
Inoculation effect									
Manipulativeness discernment	3.82	**.010**	0.342	669	1.945	.052	.210	0.215	[−0.002, 0.432]
Confidence (misinformation items)	3.61	**.013**	0.154	669	1.312	.190	.555	0.145	[−0.072, 0.362]
Confidence (nonmisinformation items)	2.60	.051	0.039	669	.324	.746	.988	0.036	[−0.181, 0.252]
Sharing discernment	1.80	.145	0.153	669	1.021	.308	.737	0.113	[−0.104, 0.330]
Manipulation technique recognition									
Manipulative vs. not manipulative	1.53	.206	0.021	669	1.648	.100	.352	0.182	[−0.035, 0.399]
True vs. false	3.45	**.016**	0.025	669	1.643	.101	.355	0.181	[−0.036, 0.398]
Technique identification	2.75	.042	0.011	669	.660	.509	.912	0.073	[−0.144, 0.290]
H2: INOCULATION WITH POSTTEST AT T0 vs. CONTROL WITH POSTTEST AT T0|**T10 (~10 DAYS LATER)**									
Inoculation effect									
Manipulativeness discernment	3.82	**.010**	0.414	669	2.335	**.020**	.091	0.260	[0.041, 0.478]
Confidence (misinformation items)	3.61	**.013**	0.355	669	3.014	**.003**	**.014**	0.335	[0.116, 0.554]
Confidence (nonmisinformation items)	2.60	.051	0.324	669	2.683	**.007**	**.037**	0.298	[0.079, 0.517]
Sharing discernment	1.80	.145	−0.301	669	1.994	**.047**	.191	0.222	[0.003, 0.440]
Manipulation technique recognition									
Manipulative vs. not manipulative	1.53	.206	0.014	669	1.108	.268	.685	0.123	[−0.095, 0.342]
True vs. false	3.45	**.016**	0.041	669	2.596	**.010**	**.047**	0.289	[0.070, 0.507]
Technique identification	2.75	**.042**	0.047	669	2.708	**.007**	**.035**	0.301	[0.082, 0.520]

Note: Significant differences at $p < .05$ are marked in bold. *t* values are consistently positive, indicating that inoculation group scores are (descriptively or significantly) higher than control group scores on all outcome measures in the table. Note that all statistical comparisons are significant for **H1**, indicating that the Bad Vaxx game yielded an inoculation effect at **T0**. For **H2**, we find no significant between-group differences at **T10** when no immediate posttest was administered, whereas 6/7 effects were significant at **T10** when participants did receive an immediate posttest (4/7 with Tukey-corrected *p* Values).

than the control group at distinguishing manipulative social media content about vaccines from neutral content, were more confident in their assessment, and had better sharing discernment. In addition, there were significant between-group differences on people's ability to distinguish true from false and manipulative from neutral vaccine-related headlines, and were better at correctly identifying which manipulation techniques were used in these headlines (all $p < .025$). All effect sizes are between $d = 0.250$ and $d = 0.477$, indicating a small to near-medium effect. These results are in line with those reported by Appel et al. (2024).

We also find partial support for **H2**: the inoculation group that did not receive a posttest at **T0** did not perform better than the control group at **T10** (all uncorrected $p > .052$, all Tukey-corrected $p > .210$). In contrast, with the exception of the "manipulative vs non-manipulative" part of the manipulation technique recognition task ($p = .268$), the inoculation group that *did* receive a posttest at **T0** performed much better at **T10** than the control group that also received a posttest at **T0** on all outcome measures (all $p < .047$). However, preregistered Tukey corrections render two of these comparisons no longer significant: manipulativeness discernment ($p_{Tukey} = .091$) and sharing discernment ($p_{Tukey} = .191$), potentially due to insufficient power; see Supporting Information S1: Table 9 for a full overview of post hoc comparisons. It should also be noted that a significant difference in performance at **T10** between the inoculation group with immediate posttest at **T0** and the inoculation group without immediate posttest would provide even stronger support of **H2**. However, as in study 1, participants who played the inoculation game and received an immediate posttest did not display significantly greater manipulativeness discernment, confidence, or sharing discernment than those in the inoculation group without immediate posttest. Nonetheless, the pattern of results is clear and highly similar to study 1: 10 days postintervention, inoculation group participants who had received an immediate posttest performed substantially better at a battery of relevant outcome measures than a control group, whereas participants who had not received an immediate posttest did not.

With respect to **H3**, we find that inoculation group participants who received an immediate posttest had significantly better memory of the relevant content and lessons learned in the *Bad Vaxx* game compared to the equivalent control group both at **T0** ($p_{Tukey} < .001$, $d = 2.730$) and at **T10** ($p_{Tukey} < .001$, $d = 1.356$), indicating that participants both learned something from the game and still remembered these lessons 10 days later; see Supporting Information S1: Table 10. However, while we find that both motivation ($p_{Tukey} = .028$) and generalized threat (a composite measure of our threat indices; $p_{Tukey} = .032$) were significantly higher in the inoculation group at **T0**, this effect disappeared at **T10**. We find no significant between-group differences at **T10** for the composite threat index, apprehensive threat, motivation to resist misinformation, fear (of vaccine misinformation), talk (how much people talked about or discussed vaccine misinformation), or issue involvement (how much vaccine misinformation means to people); all Tukey-corrected $p > .062$ with the sole exception of "fear" (how afraid people are of vaccine misinformation), which was significantly higher in the inoculation group without immediate posttest than the control group without immediate posttest at **T10** ($p_{Tukey} = .021$). This indicates that threat, the role of which is to motivate careful (central) processing of the inoculation intervention material and to motivate the accumulation of additional information (e.g., continuing to internally counterargue or "think through" the issue; Banas & Richards, 2017; Compton, 2021; Compton & Ivanov, 2012), indeed plays a role immediately after the persuasive attack, but dissipates relatively quickly afterwards. See Supporting Information S1: Table 11 for a full overview.[8]

When we compare both inoculation groups (with and without immediate posttest), we find that participants who had received an immediate posttest were significantly better at remembering the lessons from the game than those who had not ($p_{Tukey} = .003$, $d = 0.416$) at **T10**; see Supporting Information S1: Table 10 and Figure 3. However, both inoculation groups did not differ significantly from each other at **T10** for any of the motivation or threat measures (all Tukey-corrected $p > .062$, see Supporting Information S1: Table 11). This indicates that "forgetting" may play a role in the reduced effectiveness of the intervention at **T10** when no immediate posttest is administered (i.e., hypothesis **H2**). These results provide partial support for **H3** in that memory matters a great deal for the longevity of the inoculation effect, whereas threat and motivation appear to play a less important role at this later stage This is in line with the memory-motivation model of inoculation longevity proposed by Maertens et al. (2024), who also find that memory is by far the most important predictor of intervention decay and that threat plays a role early on (i.e., at the inoculation stage) in improving learning and memory (i.e., the role of threat is to motivate careful processing of the inoculation intervention). See Figure 3.

3.3 | General discussion and conclusion

Across two preregistered studies, we investigated (1) whether administering additional, irrelevant stimuli alongside a misinformation intervention interferes with its effectiveness, and (2) whether administering a posttest immediately after a misinformation intervention plays a role in the longevity of the intervention effect. With respect to the former, we find no evidence that interference is a significant problem, at least in the short term. This is a positive finding, as in the real world (such as when watching YouTube videos) people encounter all kinds of stimuli alongside any potential antimisinformation intervention, and it is helpful to know that the (immediate) inoculation effect conferred by the video is strong enough to withstand such (relatively minor) distractions (Blasiman et al., 2018; Craik, 2014; Middlebrooks et al., 2017). Although Roozenbeek et al. (2022) included a field test on YouTube where distraction was assumed to be implicit, here we validate this assumption more explicitly.

With respect to the latter, both studies show strong support for the notion that immediate posttests help strengthen the effect of the intervention and stave off decay for a substantial amount of time. We find these effects for two very different types of inoculation interventions, a video and a game, both of which have previously been extensively validated using the same outcome measures used in this study (Appel et al., 2024; Maertens et al., 2024; Roozenbeek et al., 2022). This shows that, in the absence of a further boost to memory, rapid decay can be expected even with highly robust and replicable intervention-outcome measure pairs.

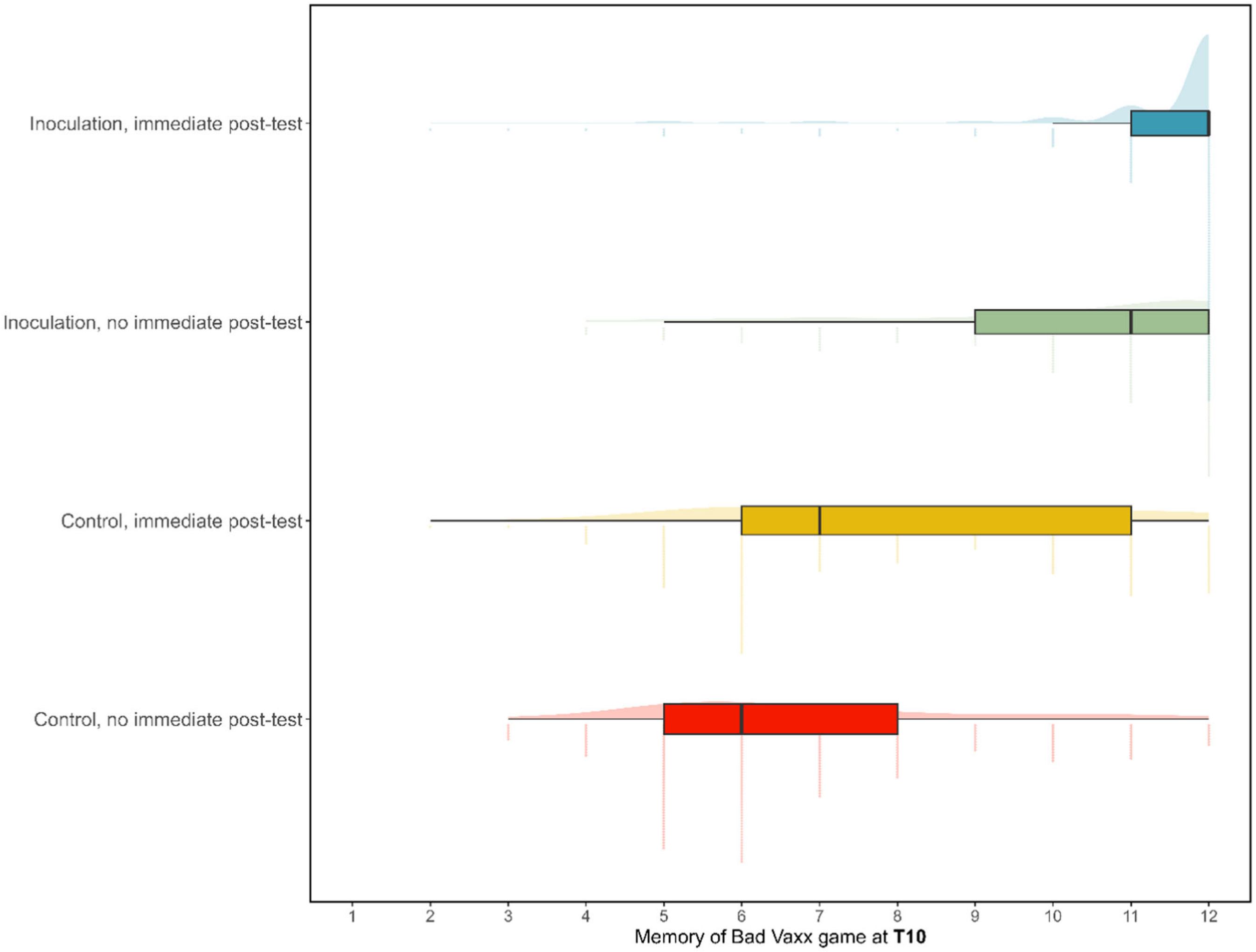

FIGURE 3 Boxplots with distribution and data jitter for "memory" (how well people remembered relevant information from the *Bad Vaxx* game at **T10** (10 days after the intervention). A higher score indicates better performance on the battery of memory tasks. Note that inoculated participants who received an immediate posttest (at **T0**) and again at **T10** scored significantly higher on the memory tasks than inoculated participants who only received a posttest at **T10** ($p_{\text{Tukey}} = 0.003$, $d = 0.416$). Control group participants who received no immediate posttest performed at chance level (as they were given no information whatsoever about vaccines or vaccine misinformation). We expect that control group participants who received an immediate posttest at **T0** performed better at **T10** than control group participants with no immediate posttest (despite not receiving the intervention) due to greater familiarity with the posttest demands and stimuli used. All between-group differences shown in the figure are significant at $p_{\text{Tukey}} < .003$; see Supporting Information S1: Table 10.

Maertens et al. (2024) propose a "memory-motivation" model of intervention effect retention, providing evidence for an exponential decay in effects if memory is not strengthened or "boosted." They also report that how well people remember the content of the intervention is by far the most important factor in how long the effect remains over time. Our study provides further support for this model. Importantly, our study 2 shows that inoculation group participants who received an immediate posttest were significantly more likely than control participants to correctly remember the lessons from the intervention 10 days later, whereas inoculated participants who only went through the intervention without doing a posttest immediately afterwards performed no differently than controls. While it is important to note that we did not observe a statistically significant difference in performance between the two inoculation groups, even after a 10-day delay, this relative difference in performance nevertheless provides support for the idea that the inoculation effect decays quicker *because* participants forgot the relevant lessons more quickly if they did not receive an immediate posttest. Indeed, this model is in accordance with research on debunking approaches to misinformation, which has also established the importance of memory. Specifically, memory for a correction was shown to be strongly associated with reduced belief in the corrected misinformation (Fazio et al., 2023; Swire-Thompson et al., 2023). Thus, in line with previous findings on memory retention after watching a video (Bird et al., 2015) and broader research on the learning benefits of retrieval practice (Agarwal et al., 2021; Carpenter et al., 2022), we posit that a test of newly acquired information is a critical component for strengthening memory. And indeed, Leder et al. (2024) find preliminary evidence that memory-strengthening exercises administered at the end of an inoculation game boost memory and hence item task performance 1 week postintervention.

That said, there is evidence that the intervention effect persists at least for a modest duration even if no posttest is administered: Roozenbeek et al. (2022) conducted a randomized controlled field trial on YouTube with the same "emotional language" video used in the present study, where they administered a single-item posttest (in the form of a single multiple-choice question) a median 18 h after watching the video as a YouTube ad. They report a significant (if relatively small) improvement in the correct identification of manipulative content, indicating that some of the inoculation effect still remains about 18 h after viewing.

Our findings also complement a line of inoculation studies that seek to optimize inoculation treatments with "booster shots," much like booster doses are administered to sustain vaccinations against disease, like COVID-19, something suggested in the earliest inoculation theory research (see McGuire, 1961). To date, most booster research in inoculation theory uses a second full inoculation treatment message some time after the initial intervention (e.g., Pfau et al., 2006). Whereas some have found enhanced resistance from using boosters (e.g., Pfau et al., 2006), others have not (e.g., Pfau et al., 1990). Our findings suggest that tests can boost resistance, and perhaps, test-focused boosters will show more efficacy than repeated inoculation interventions have, much like Compton and Ivanov (2012) found that incorporating a threat measure into the inoculation intervention message boosted resistance.

In our view, our findings have substantial implications for the scalability of misinformation interventions that rely on participants learning and remembering new information (i.e., "boosts," see Hertwig & Grüne-Yanoff, 2017; Kozyreva et al., 2024): if merely watching a video, playing a game, or reading a piece of text (without an opportunity to practise what is learned) is not enough to prevent immediate forgetting of manipulativeness discernment skills, this requires a careful rethinking of how interventions are deployed (e.g., on social media). For instance, it may be very difficult to persuade social media users to not only watch an inoculation video (or other intervention) as an ad on YouTube, Facebook, or another platform, but also to then voluntarily engage in some kind of exercise that helps practice the lessons from the intervention. This, we believe, is a tall order: it is difficult enough to get people to watch ads or engage with (digital) literacy interventions to begin with; how many people will then also voluntarily agree to practice what they learned? We argue that these findings should prompt misinformation researchers to carefully think about intervention design and implementation. For instance, future research may investigate how including exercises with real-time feedback or inductive learning in media literacy interventions or "fake news" games impacts their efficacy and longevity (Motz et al., 2022). One study by Leder et al. (2024) showed promising results for the potential of such feedback exercises to boost the longevity of *Bad News*, a different inoculation game to the one tested in the present study. Nonetheless, we also note that more research is needed to explore whether our findings hold up for text-based "boosting" interventions (such as media literacy tips or infographics).

ACKNOWLEDGMENTS

We thank Cecilie Steenbuch Traberg for her help designing the social media post item rating task used as the posttest in study 1. We also would like to thank Ruth Elisabeth Appel and Rebecca Rayburn-Reeves for their help with the brainstorming for this study. The authors are grateful to the British Academy (#PF21\210010), the Department of Psychology at the University of Cambridge, and the American Psychological Association together with the Centers for Disease Control and Prevention (Award #6NU87PS004366-03-02) for providing funding. S. L. acknowledges financial support from the European Research Council (ERC Advanced Grant 101020961 PRODEMINFO), the Humboldt Foundation through a research award, and from UK Research and Innovation (EU Horizon replacement funding grant number 10049415). This study was financed with funding provided by the British Academy (awarded to J. R.; #PF21\210010), the Department of Psychology at the University of Cambridge, and the American Psychological Association together with the Centers for Disease Control and Prevention (Award #6NU87PS004366-03-02). The funding source had no involvement in the study design, data collection, data analysis, interpretation of data, writing, or decision to submit.

DATA AVAILABILITY STATEMENT

We confirm that we have reported all measures, conditions, data exclusions, and how we determined our sample sizes. All analysis scripts and data sets can be found on our OSF page: https://osf.io/9frch/.

ORCID

Rakoen Maertens http://orcid.org/0000-0001-8507-5359
Miriam Remshard http://orcid.org/0000-0002-6121-1341
Sander van der Linden https://orcid.org/0000-0002-0269-1744
Josh Compton http://orcid.org/0000-0003-4625-3823
Stephan Lewandowsky http://orcid.org/0000-0003-1655-2013
Jon Roozenbeek http://orcid.org/0000-0002-8150-9305

ENDNOTES

[1] Some studies (e.g., Basol et al., 2020; Maertens et al., 2021; van der Linden et al., 2017) have also used a pretest, administered before the intervention, alongside the posttest, as has much of the traditional research on inoculation theory (Pfau & Burgoon, 1988; Pfau et al., 2004, 2006).

[2] But see Compton and Pfau (2005), McGuire (1964), and Compton (2013) for earlier reviews on longitudinal inoculation studies.

[3] Bird et al. (2015) also offer a neuroscientific explanation for their findings.

[4] The posttest did not include any items related to vaccination, see Supporting Information S1: Table 7.

[5] The comparisons between groups 2 and 4 at **T2** for confidence (M_{diff} = 0.012, $t(259)$ = 0.084, p_{Tukey} > .999, d = 0.010, 95% CI [−0.234, 0.255]), trustworthiness discernment (M_{diff} = 0.077, $t(259)$ = 0.504, p_{Tukey} = .958, d = 0.063, 95% CI [−0.307, 0.182]) and sharing discernment (M_{diff} = −0.037, $t(259)$ = −0.310, p_{Tukey} = 0.990, d = −0.039, 95% CI [−0.282, 0.206]) show similar conclusions. See Supporting Information S1: Tables 3–5.

[6] An example of a question measuring apprehensive threat is "The idea that there is the possibility you may come in to contact with deception techniques that are so effective that they might cause you to be misled is [dangerous]". The accessibility question was worded as follows: "Compared to other issues, how often do you think about vaccine misinformation?"; see Supporting Information S1: Table 12.

[7] There are various conceptualizations of "threat" in inoculation theory; in some cases, it can refer to the recognition that one's position on an issue (e.g., the efficacy of vaccines) is vulnerable to future persuasive attacks (e.g., Banas & Richards, 2017; Compton, 2021; Compton & Ivanov, 2012), whereas in the present study and other studies (e.g., Basol et al., 2021; Maertens et al., 2024) threat refers to the extent to which the persuasive attack (e.g., vaccine misinformation) is perceived as threatening.

[8] We preregistered that we would conduct subgroup analyses for our covariates. In the interest of brevity, we will not do so (as this is not the main focus of the present study).

REFERENCES

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review*, *33*(4), 1409–1453. https://doi.org/10.1007/s10648-021-09595-9

Appel, R. E., Roozenbeek, J., Rayburn-Reeves, R., Basol, M., Corbin, J., Compton, J., & van der Linden, S. (2024). Psychological inoculation improves resilience to and reduces willingness to share vaccine misinformation. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/ek5pu

Banas, J., & 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

Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & Linden, S. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data & Society*, *8*(1), 205395172110138. 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), 1–9. https://doi.org/10.5334/joc.91

Bird, C. M., Keidel, J. L., Ing, L. P., Horner, A. J., & Burgess, N. (2015). Consolidation of complex events via reinstatement in posterior cingulate cortex. *The Journal of Neuroscience*, *35*(43), 14426–14434. https://doi.org/10.1523/JNEUROSCI.1774-15.2015

Blasiman, R. N., Larabee, D., & Fabry, D. (2018). Distracted students: A comparison of multiple types of distractions on learning in online lectures. *Scholarship of Teaching and Learning in Psychology*, *4*(4), 222–230. https://doi.org/10.1037/stl0000122

Brashier, N. M., Pennycook, G., Berinsky, A. J., & Rand, D. G. (2021). Timing matters when correcting fake news. *Proceedings of the National Academy of Sciences*, *118*(5), e2020043118. https://doi.org/10.1073/pnas.2020043118

Carpenter, S. K., Pan, S. C., & Butler, A. C. (2022). The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology*, *1*(9), 496–511. https://doi.org/10.1038/s44159-022-00089-1

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*, 1–18. https://doi.org/10.1057/s41599-022-01174-9

Collier, J. R., Pillai, R. M., & Fazio, L. K. (2023). Multiple-choice quizzes improve memory for misinformation debunks, but do not reduce belief in misinformation. *Cognitive Research: Principles and Implications*, *8*(1), 37. https://doi.org/10.1186/s41235-023-00488-9

Compton, J. (2013). Inoculation theory. In J. P. Dillard, & L. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed., pp. 220–236). SAGE Publications, Inc. https://doi.org/10.4135/9781452218410

Compton, J. (2021). Threat and/in inoculation theory. *International Journal of Communication*, *15*, 4294–4306.

Compton, J., & Ivanov, B. (2012). Untangling threat during inoculation-conferred resistance to influence. *Communication Reports*, *25*(1), 1–13. https://doi.org/10.1080/08934215.2012.661018

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

Compton, J. A., & Pfau, M. (2005). Inoculation theory of resistance to influence at maturity: Recent progress in theory development and application and suggestions for future research. *Annals of the International Communication Association*, *29*(1), 97–146. https://doi.org/10.1207/s15567419cy2901_4

Cook, J., Ecker, U. K. H., Trecek-King, M., Schade, G., Jeffers-Tracy, K., Fessmann, J., Kim, S. C., Kinkead, D., Orr, M., Vraga, E. K., Roberts, K., & McDowell, J. (2022). The Cranky Uncle game—Combining humor and gamification to build student resilience against climate misinformation. *Environmental Education Research*, *29*(4), 607–623. https://doi.org/10.1080/13504622.2022.2085671

Craik, F. I. M. (2014). Effects of distraction on memory and cognition: A commentary. *Frontiers in Psychology*, *5*, 841. https://doi.org/10.3389/fpsyg.2014.00841

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Fazio, L. K., Hong, M. K., & Pillai, R. M. (2023). Combatting rumors around the French election: The memorability and effectiveness of fact-checking articles. *Cognitive Research: Principles and Implications*, *8*(1), 44. https://doi.org/10.1186/s41235-023-00500-2

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, *12*(6), 973–986. https://doi.org/10.1177/1745691617702496

Johansson, P., Enock, F., Hale, S., Vidgen, B., Bereskin, C., Margetts, H., & Bright, J. (2022). How can we combat online misinformation? A systematic overview of current interventions and their efficacy. *ArXiv Preprint*. https://doi.org/10.48550/arXiv.2212.11864

Kata, A. (2010). A postmodern pandora's box: Anti-vaccination misinformation on the Internet. *Vaccine*, *28*, 1709–1716. https://doi.org/10.1016/j.vaccine.2009.12.022

Kata, A. (2012). Anti-vaccine activists, web 2.0, and the postmodern paradigm—An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine*, *30*(25), 3778–3789. https://doi.org/10.1016/j.vaccine.2011.11.112

Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., Basol, M., Berinsky, A. J., Betsch, C., Cook, J., Fazio, L. K., Geers, M., Guess, A. M., Maertens, R., Panizza, F., Pennycook, G., Rand, D. J., Rathje, S., Reifler, J., … Wineburg, S. (2024). Toolbox of interventions against online misinformation and manipulation.

Nature Human Behaviour. https://doi.org/10.1038/s41562-024-01881-0

Leder, J., Schellinger, L. V., Maertens, R., Chryst, B., van der Linden, S., & Roozenbeek, J. (2024). Feedback exercises boost discernment for gamified misinformation interventions. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001603

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., & 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*(57), 57. https://doi.org/10.1186/s41235-021-00323-z

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/10.1002/gch2.201600008

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*, 337–348. https://doi.org/10.1038/s41562-021-01056-1

Loomba, S., Götz, F. M., Maertens, R., Roozenbeek, J., de Figueiredo, A., & van der Linden, S. (2024). Ability to detect fake news predicts geographical variation in COVID-19 vaccine uptake. *MedRxiv Preprints*. https://doi.org/10.1101/2023.05.10.23289764

Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., Kerr, J. R., Stieger, S., McClanahan, W. P., Drabot, K., He, J., & van der Linden, S. (2023). The misinformation susceptibility test (MIST): A psychometrically validated measure of news veracity discernment. *Behavior Research Methods*, *56*, 1863–1899. https://doi.org/10.3758/s13428-023-02124-2

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

Maertens, R., Roozenbeek, J., Simons, J., Lewandowsky, S., Maturo, V., Goldberg, B., Xu, R., & van der Linden, S. (2024). Psychological booster shots targeting memory increase long-term resistance against misinformation. *Nature Communications*. https://osf.io/preprints/psyarxiv/6r9as

Maertens, R., Said, N., Buder, J., & Roozenbeek, J. (2023). Misleading but not fake: Measuring the difference between manipulativeness discernment and veracity discernment using psychometrically validated tests. *PsyArxiv Preprints*. https://doi.org/10.31219/osf.io/tckp4

McGuire, W. J. (1961). The effectiveness of supportive and refutational defenses in immunizing and restoring beliefs against persuasion. *Sociometry*, *24*(2), 184. https://doi.org/10.2307/2786067

McGuire, W. J. (1964). Some contemporary approaches. *Advances in Experimental Social Psychology*, *1*(C), 191–229. https://doi.org/10.1016/S0065-2601(08)60052-0

McGuire, W. J., & Papageorgis, D. (1962). Effectiveness of forewarning in developing resistance to persuasion. *Public Opinion Quarterly*, *26*(1), 24–34. https://doi.org/10.1086/267068

Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively distracted: Divided attention and memory for important information. *Psychological Science*, *28*(8), 1103–1115. https://doi.org/10.1177/0956797617702502

Motz, B. A., Fyfe, E. R., & Guba, T. P. (2022). Learning to call bullsh*t via induction: Categorization training improves critical thinking performance. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1037/mac0000053

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pfau, M., & Burgoon, M. (1988). Inoculation in political campaign communication. *Human Communication Research*, *15*(1), 91–111. https://doi.org/10.1111/j.1468-2958.1988.tb00172.x

Pfau, M., Compton, J., Parker, K. A., An, C., Wittenberg, E. M., Ferguson, M., Horton, H., & Malyshev, Y. (2006). The conundrum of the timing of counterarguing effects in resistance: Strategies to boost the persistence of counterarguing output. *Communication Quarterly*, *54*(2), 143–156. https://doi.org/10.1080/01463370600650845

Pfau, M., Compton, J., Parker, K. A., Wittenberg, E. M., An, C., Ferguson, M., Horton, H., & Malyshev, Y. (2004). The traditional explanation for resistance versus attitude accessibility: Do they trigger distinct or overlapping processes of resistance? *Human Communication Research*, *30*(3), 329–360. https://doi.org/10.1093/hcr/30.3.329

Pfau, M., Kenski, H. C., Nitz, M., & Sorenson, J. (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

Roozenbeek, J., & van der Linden, S. (2024). *The psychology of misinformation*. Cambridge University Press.

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). https://doi.org/10.1126/sciadv.abo6254

Schmid, P., Altay, S., & Scherer, L. D. (2023). The psychological impacts and message features of health misinformation: A systematic review of randomized controlled trials. *European Psychologist*, *28*(3), 162–172. https://doi.org/10.1027/1016-9040/a000494

Swire, B., Berinsky, A. J., Lewandowsky, S., & Ecker, U. K. H. (2017). Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society Open Science*, *4*(3), 160802. https://doi.org/10.1098/rsos.160802

Swire-Thompson, B., Dobbs, M., Thomas, A., & DeGutis, J. (2023). Memory failure predicts belief regression after the correction of misinformation. *Cognition*, *230*, 105276. https://doi.org/10.1016/j.cognition.2022.105276

Ziemer, C., & Rothmund, T. (2022). Psychological underpinnings of misinformation counter measures: A systematic scoping review. *Journal of Media Psychology: Theories, Methods, and Applications. Advance online publication*. https://doi.org/10.1027/1864-1105/a000407

Zollo, F., Novak, P. K., Vicario, M. D., Bessi, A., Mozetic, I., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Emotional dynamics in the age of misinformation. *Plos One*, *10*, e0138740. https://doi.org/10.1371/journal.pone.0138740

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Capewell, G., Maertens, R., Remshard, M., van der Linden, S., Compton, J., Lewandowsky, S., & Roozenbeek, J. (2024). Misinformation interventions decay rapidly without an immediate posttest. *Journal of Applied Social Psychology*, 1–14. https://doi.org/10.1111/jasp.13049