Received: 17 February 2025 | Accepted: 10 July 2025

DOI: 10.1111/lcrp.70004

the british psychological society
promoting excellence in psychology

ARTICLE

Fore! Does forewarning inoculate people against the false balance effect?

Tianshuang Han | **Brent Snook** | **Martin V. Day**

Department of Psychology, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada

Correspondence
Tianshuang Han, Department of Psychology, Memorial University of Newfoundland, Science Building, St. John's, NL, Canada.
Email: tianshuangh@mun.ca

Abstract

Background & Aims: We examined the effect of falsely balanced messages on perceptions of expert consensus about non-verbal lie detection and whether forewarning inoculates people against the fake debate strategy.

Materials & Methods: Participants (N = 307) read a media report that revealed high consensus among experts (nearly 90%) that non-verbal cues are unreliable indicators of deception and were randomly exposed to (1) no comments from experts, (2) balanced comments (three comments from each expert on opposing sides), (3) evidentiary balanced comments (five comments from a deception detection expert and one comment from a contrarian expert), (4) balanced comments along with a forewarning about the 'fake debate' strategy, or (5) evidentiary balanced comments along with a forewarning about the 'fake debate' strategy.

Results: Results showed that participants intuitively believe that non-verbal cues are reliable indicators of deceit. Although participants were made aware that the consensus from scientists is that non-verbal lie detection is futile, the inclusion of balanced comments alongside the data still decreased perceived scientific consensus. Balanced comments also reduced people's policy support in favour of scientific consensus, and forewarning had minimal effect.

Discussion: We discuss the implications of our findings for efforts to mitigate the fake debate strategy.

KEYWORDS

deception detection, false balance, forewarning, perceived expert consensus, science communication, weight of evidence

BACKGROUND

Misinformation spans various societal issues, from vaccine hesitancy to climate change denial. Although less prominent, the domain of deception detection is also fraught with misinformation. Despite substantial research and a wealth of empirical evidence regarding what works, this area remains riddled with misconceptions about how to catch a liar. One of the most contentious topics in this domain is non-verbal lie detection (Denault, 2020). There is a strong consensus among experts in the scientific literature that no non-verbal cue reliably signals deception (e.g. DePaulo et al., 2003; Luke et al., 2023; Sporer & Schwandt, 2007). Moreover, people's accuracy, including professionals, in detecting lies using non-verbal cues is only around chance levels (54% average, see Bond Jr & DePaulo, 2006), and training programmes that focus on non-verbal cues to deceit offer negligible improvements (Hauch et al., 2016). However, the fringe notion that people can use non-verbal cues to detect lies has permeated popular culture, mass media and consequential areas such as counterintelligence and security, and it has become so deep-rooted that private and public funding is sustaining a multibillion-dollar industry (Denault et al., 2020; Patterson et al., 2023). As Vrij et al. (2019) noted, 'Lively debates about the merits of non-verbal lie detection no longer take place at the scientific conferences that we attend. Yet non-verbal lie detection remains highly popular among practitioners, such as police detectives, and in the media' (p. 302).

The disproportionate attention given to the 'utility' of non-verbal cues to detect deception creates a dangerous illusion of controversy and confusion. The parallel presentation of misinformation and facts imposes a *false balance*, undermining the public's understanding of deception detection. In response to such scientific denialism, researchers have started to focus on testing ways to guard against the aggressive spread of misinformation across various domains. As a consequence, we test the extent to which forewarning people about a fake debate pertaining to non-verbal lie detection can protect the public against harm.

The role of scientists in combating misinformation

The rise of digital platforms has granted the public unprecedented access to information and fuelled an alarming deluge of misinformation. For example, researchers have estimated that 33% of popular mental health content and 52% of ADHD videos on TikTok are misleading (Turuba et al., 2025; Yeung et al., 2022). Google search results and Meta frequently lead individuals to content about fraudulent, alternative medical treatments (e.g. Zenone, Snyder, et al., 2024). Moreover, almost 50% of cancer treatment books on Amazon contain misinformation, and 70% of top Amazon search results contain falsehoods (Zenone, van Schalkwyk, et al., 2024). Online misinformation campaigns reach billions of people with persuasive messages tailored to users' psychological profiles and shield users from information that challenges their cherished beliefs (Ecker et al., 2022). Disproven ideas or denials have become increasingly prevalent in public discourse (e.g. herd immunity rhetoric prevailed during the COVID-19 pandemic; Zenone et al., 2022). This rapid spread of misinformation, or 'infodemic', fuels political division, extremism and a profitable industry based on lies. In this landscape, scientific knowledge mobilization is essential but not enough. When misinformation and facts coexist side-by-side, it misleads the public into believing there is a debate where none exists. As a result, combating misinformation is an urgent issue that requires attention.

Scientists have a unique and important responsibility to fight pseudoscientific claims and science denials while upholding the integrity of empirical-based knowledge (Lilienfeld, 1998; Lilienfeld et al., 2005). The public often relies on experts, that is, people who possess superior skills and abilities for problem-solving and decision-making (Kurz-Milcke & Gigerenzer, 2004), making them an essential source of accurate information. Amid rising populism, fake experts, or 'contrarian voices', are shaping opinions and policies by manufacturing fake debates over scientific and historical facts (Lewandowsky et al., 2017). Countering fake debates protects the public from harm and reduces the associated opportunity costs, redirecting resources to proven, beneficial services (Lilienfeld et al., 2005). Such action also defends public trust in science, preserving its future influence. The defining strength of science lies in empiricism (e.g. conducting research and making data-driven propositions), providing a common understanding of how the world works (Hansson, 2017; Lilienfeld, 1998; McNally, 2003). While legitimate scepticism is critical for generating empirical knowledge, the spread of anti-science and anti-evidence may cast doubt on scientific consensus, erode the empirical foundation of science and risk steering a new generation of researchers away from scientific rigour (Lilienfeld et al., 2005). Therefore, scientists must *effectively* combat misinformation.

False balance and perceived scientific consensus

One way misinformation spreads is via false balance or 'fake debate' strategy. False balance arises when two opposing views of an issue are presented as equally valid despite the overwhelming empirical evidence supporting only one side. Fake debates can lead the public to believe that fringe views or misinformation are on par with facts (Cook, 2022; Dearing, 1995). Although traditionally associated with journalistic practices that exploit the norm of balance (Boykoff & Boykoff, 2004), false balance is increasingly driven by rampant misinformation campaigns through digital infrastructure. The widespread dissemination of falsehoods by vocal and influential science deniers, the use of microtargeting (i.e. targeted advertising based on personal data), social media echo chambers, fragmented news media and ideologically motivated online endorsements have all amplified the issue—giving contrarian views on the fringe outsized exposure in public discussions (Caulfield, 2023).

Empirical evidence has shown that falsely balanced messages reduce public perceptions of scientific consensus on topics such as vaccines and autism (c.f., one-sided scientific fact or unrelated control article; Dixon & Clarke, 2013; Dixon et al., 2015), climate change (Cook et al., 2017; Imundo & Rapp, 2022), health risks of pollution (Kortenkamp & Basten, 2015) and facial recognition (Kohl et al., 2016). Imundo and Rapp (2022) found evidence of a false balance effect even when the science denier's expertise was irrelevant to global warming (i.e. physics) or the science denier did not have a scientific background (i.e. automotive plant manager). Their findings suggest that people focus more on the representativeness of opposing views than on source factors (e.g. expertise, credibility; Chaiken & Maheswaran, 1994).

Empirical studies have also found that false balance undermines public understanding of the strength of scientific support for economic, environmental and legal issues (Han et al., 2024, 2025; Koehler, 2016). Presenting balanced messages alongside numerical weight-of-evidence (WOE; e.g. the percentage of scientists who agree with a research finding; typically measured via surveys of expert opinions) led to less accurate assessments of expert consensus about an issue than when only exposed to the data, particularly for matters where there is a high level of consensus among scientists (Han et al., 2024; Koehler, 2016). These findings suggest that false balance distorts perceived expert consensus even when people know the true level of expert consensus. The juxtaposition of scientific facts and denials appears to cause a binary partition of expert opinions that shifts judgements towards a midpoint, leading people to perceive more controversy than what exists (Han et al., 2025; Koehler, 2016).

Inoculation against fake debates

One way to combat the spread of misinformation from a false balance strategy is to inform people about the WOE information (i.e. data) behind competing views (Dunwoody, 2005). Research has shown that including WOE information alongside a balanced message yields more accurate assessments of the level of expert consensus about an issue than reading a balanced message alone (e.g. false balance + data > false balance; Dixon et al., 2015; Imundo & Rapp, 2022). However, no intervention has addressed the issue of false balance on undermining statistical data about the actual 'weight' on each side (false balance + data < data).

A promising intervention is *forewarning*, which involves warning individuals about attempts to influence the importance they place on a forthcoming persuasion episode, thereby helping them resist its impact (McGuire & Papageorgis, 1962; Schmid et al., 2020; Wood & Quinn, 2003). Inoculation theory suggests that individuals can activate their own 'immune responses' against persuasion when they preemptively receive a warning about being targeted by persuasive messages and then a refutation exposing the fallacies within those messages (McGuire, 1961a, 1961b). The pre-influence warning presumably motivates people to protect themselves from the threat of inappropriate persuasion. At the same time, the refutation equips them with the tools to do so (McGuire, 1964; Van der Linden, 2024). We use forewarning as an umbrella term for any intervention that forewarns people about misinformation. Meta-analyses revealed medium effects of forewarning, showing that forewarned individuals were less convinced by the persuasive messages delivered to them (i.e. $d=0.43$, Banas & Rains, 2010; $d=0.42$, Wood & Quinn, 2003). Empirical evidence also indicated that forewarnings with detailed refutations generated more resistance against persuasion (e.g. $d=0.75$) than forewarnings with general refutations (e.g. $d=0.33$; McGuire & Papageorgis, 1962; Van der Linden et al., 2017). In addition, research shows that forewarning reduces the biasing effects of various types of misinformation, including science denials (Roozenbeek et al., 2022; Van der Linden et al., 2017; Wood, 2007), conspiracy theories (Banas & Miller, 2013), politically polarized content (Smith et al., 2025), extremism (Braddock, 2022; Lewandowsky & Yesilada, 2021) and misleading information that contaminates eyewitness memory (Gallo et al., 1997; Greene et al., 1982; Neuschatz et al., 2003; Schopen et al., 2022).

In the case of false balance, being warned about the existence of a fake debate before it is encoded could reduce its negative influence on perceived scientific consensus (Cook et al., 2017; Schmid et al., 2020; Van der Linden et al., 2017). Specifically, Van der Linden et al. (2017) investigated several types of climate misinformation about scientific consensus (i.e. denial and false balance) and interventions (i.e. consensus message and forewarnings). Although not directly forewarning about false balance or fake debates, one of their forewarning manipulations involved a general warning about politically motivated actors using misleading tactics to convince the public that there is significant disagreement among scientists. Participants who read general forewarning and falsely balanced messages improved their perceptions of scientific consensus compared to those who only read the balanced messages ($d=0.27$). Cook et al. (2017) replicated this finding in the area of global warming by using forewarnings about the 'false balance' strategy employed by the tobacco industry. They found that forewarning individuals about the potentially biasing effect of falsely balanced messages neutralized the negative impact of such information on perceived expert consensus about global warming ($d=0.26$). Schmid et al. (2020) extended the intervention research to vaccinations, demonstrating that forewarning about the false-balance effect mitigated the negative impact of vaccine denialism messages on the audience's beliefs (e.g. vaccination safety, trust in institutions, intention to vaccinate). In short, forewarning appears to help reduce harm before a falsely balanced discussion is processed.

The current research

One goal of the current study is to replicate the finding that false balance overshadows the numerical strength of evidence (WOE). In other words, does false balance influence people's estimation of expert consensus after they see data on the actual level of consensus? While someone might assume that people would simply reiterate the exact number of experts who agreed (regardless of the comments they read), previous research did not find that to be the case (Han et al., 2024, 2025; Koehler, 2016). A second goal was to test whether forewarning people about the 'fake debate' strategy counteracts the false-balance effect in undermining the WOE about the actual level of expert consensus. Empirical evidence suggests that forewarning about false balance inoculates people against its distorting impact on perceived scientific consensus (e.g. Cook et al., 2017). However, little research examines if forewarning would mitigate the distorting effect of false balance on the perception of expert consensus when the WOE data are available. Further, while previous research employed vignettes in its experimental paradigm, we used a mock media report to improve the ecological validity of our findings. We explore the two aforementioned goals within the context of deception detection. We test a scenario where a scientist aims to educate the public, who likely holds a positive view of non-verbal cues, about how non-verbal cues to deception are ineffectual. As mentioned, the domain of deception detection is subject to misinformation and the scientific consensus that non-verbal cues are unreliable indicators of deception is under constant attack (Denault et al., 2020; Patterson et al., 2023).

METHOD

Participants

While planning and following a Bayesian analytical approach, we conducted a power analysis rooted in NHST assumptions as a compromise for readers interested in this approach and those who view the results of our NHST testing (see https://osf.io/tra9w/). An a priori power analysis using G*Power indicated that 305 participants were required to detect a small effect ($d = 0.40$) with an alpha of .05 and 80% power for a between-subjects design with five conditions. We chose to detect an effect of 0.40 based on previous research on the false balance effect in the interrogation context (Han et al., 2024, 2025). The chosen effect size was in the middle range of effect sizes found in previous research, and using that effect size allowed us to have a relatively larger sample size when attempting to replicate Han et al. (2024, 2025).

We recruited participants ($N = 307$) from the general population in Canada through an online survey platform called Prolific. Participants were pre-screened to have English as their first language and compensated £2 ($3.60 Canadian) for their time, which lasted around 20 minutes. Two participants failed the attention check, and one encountered technical issues on their device, resulting in a final sample size of 304. Of those who answered the demographic questions, 165 participants (55%) were women, 210 (69%) were between the ages of 20 and 39, 189 (62%) were White, and 181 (60%) reported having post-secondary education. See https://osf.io/tra9w/ for a breakdown of participants by demographic variables. There were no differences in demographic characteristics across conditions.

Design

Based on Schmid et al. (2020) and Han et al. (2025), we employed a between-participant design with five conditions. Specifically, the conditions were as follows: (1) only WOE data (i.e. percentage of scientists agreeing with the research finding) with no comments (i.e. the no-balance control condition), (2) WOE data with balanced comments from experts (i.e. three agree vs. three disagree; the balance condition), (3) WOE data with more consensus-supporting comments (i.e. five agree vs. one disagree; the evidentiary balance condition), (4) WOE data with balanced comments and a forewarning about the 'fake debate' strategy (i.e. the balance + forewarning condition), and (5) WOE data with more consensus-supporting comments and a forewarning about the 'fake debate' strategy (i.e. the evidentiary balance + forewarning condition).

Our primary variable of interest was people's estimation of expert consensus, that is, perceived expert consensus after they viewed WOE data regarding the actual level of consensus. We pre-registered the experiment at https://osf.io/ze59v and conducted it online via Qualtrics. While we pre-registered a detailed study design, procedure and hypotheses, we failed to pre-register our analytic plan beyond prescribing it to be Bayesian. All experimental materials are available at https://osf.io/tra9w/.

Materials and procedure

As potential covariates, we first asked participants to indicate the extent to which *they* believe that people can use non-verbal cues (e.g. lack of eye contact and fidgeting) to detect lies accurately and the extent to which they think *experts* believe that people can use non-verbal cues (e.g. lack of eye contact and fidgeting) to detect lies accurately. Participants answered these two questions on a slider scale from 0 to 100.

Subsequently, we randomly assigned participants to read one of five versions of a media report (https://osf.io/tra9w/). Every media report mentioned a survey of expert views on deception detection. In the media report, the journalist revealed that nearly 90% of the surveyed experts agreed that non-verbal cues are unreliable indicators of deception. The journalist also reported that experts believe people cannot use non-verbal cues to detect lies accurately. We adapted the data of expert consensus (i.e. WOE; 90%) from Kassin et al. (2018) and Luke et al. (2023), which examined the level of agreement within the scientific community on issues including non-verbal cues and deception detection.

Following the WOE data, the media report presented participants with a different number of comments from two opposing experts, with or without a forewarning about the 'fake debate' strategy. Participants either read no comments (i.e. no-balance control condition), three comments from an expert who agreed that non-verbal cues cannot accurately detect deception and three comments from an expert who disagreed (i.e. the balance condition), five comments from an agreeing expert and one comment from a disagreeing expert (i.e. the evidentiary balance condition), three comments from an agreeing expert and three from a disagreeing expert with a forewarning of fake debates (i.e. the balance + forewarning condition), or five comments from an agreeing expert and one from a disagreeing expert with a forewarning (i.e. the evidentiary balance + forewarning condition). We used the literature on false balance and WOE (Han et al., 2025; Schmid et al., 2020) to decide the number of expert comments needed to adequately manipulate the level of balanced presentation of expert opinions (i.e. evenly split 3:3, evidentiary proportionate 5:1). We set up equivalent source expertise for the two opposing experts and used the same pair of experts across conditions. We opted for two experts rather than multiple ones (one per comment) to maintain a manageable media report length and to ensure feasibility in presenting equivalent information for each expert. Also, readers of media reports may focus less on the sources themselves and more on the content where the journalist places the bulk of the arguments (Imundo & Rapp, 2022). Both experts were university professors with 20 years of experience researching non-verbal cues and deception detection. We provided the information in the media reports.

We created five *agreeing* comments based on research articles and three *disagreeing* comments based on counterarguments to scientific claims. All comments were 34 words in length and were similar in structure. First, we adapted three pairs of equivalent comments for the balance condition (i.e. three agree vs. three disagree), which served as the 'false balance' condition. Specifically, we based the three agreeing comments on a review by Patterson et al. (2023) debunking the misconception that 'the body never lies'. We developed our three disagreeing comments to match the agreeing comments based on counterarguments supporting the minority view (Matsumoto, 2023). The first pair of comments addressed body language, the second pair focused on microexpressions and the third pair juxtaposed debunking and supporting the misconception that 'the body never lies'. We then rephrased the third agreeing comment, which debunked the misconception in general, to create an additional two agreeing comments for the evidentiary balance condition (i.e. five agree vs. one disagree). While all five agreeing comments were included in this condition, only the third disagreeing comment, which supported the misconception, was included (see https://osf.io/tra9w/ for the wording of all comments).

We based our *forewarning* message on inoculation messages used in previous research (Cook et al., 2017; Schmid et al., 2020), which explained that initiating a 'fake debate' is a science denial strategy and challenged the logical fallacy behind both-siding. Specifically, our forewarning read:

> Dr. Luther also pointed out that a debate about the merits of nonverbal lie detection is no longer taking place within the science community. However, there are spokespeople—who are not experts—and a small minority of scientists who keep trying to challenge the scientific fact that people cannot detect deception accurately by looking at nonverbal cues. These contrarians try to manufacture a 'fake debate'. Luther said that 'creating a fake debate is a widely used science denialist strategy that tries to expose the public to "both sides" of an issue when only one side is supported by an overwhelming amount of empirical evidence'. He also commented that 'Ironically, the journalistic standard of both-siding is distorting the reality of deception detection research, and this practice should stop'.

In the balance condition, we first presented comments debunking the link between body language and deception, using micro-expressions to detect lies, and the idea that 'the body never lies' (i.e. three comments agreed with scientific consensus). We then presented three contrarian comments (i.e. three disagreeing with scientific consensus). In the evidentiary balance condition, we presented the same comments that agreed with scientific consensus and two additional comments that aimed to debunk the proposition that non-verbal cues are useful for detecting lies, followed by one comment supporting the misconception. We presented the comments in the same order for participants within each version of the media report because there was no evidence of order effects in previous research (Han et al., 2024, 2025). In the two forewarning conditions, we presented, in order, comments against the utility of non-verbal cues, a forewarning about the fake-debate strategy and comment(s) against the scientific consensus on the issue. We made this decision recognizing that it is unrealistic to rely on the presence of a moderator or platform consistently willing to forewarn about fake debates—especially given that major platforms like Facebook and Google are rolling back misinformation safeguards (Booth, 2025; Weatherbed, 2025). Therefore, we decided to let the deception detection expert deliver the forewarning before the contrarian voiced their opinions, which was the most natural way to preemptively intervene in the following presentation of a fake debate. We also based this decision on misinformation research showing that forewarning does not mean to be literally before misinformation but rather before the subsequent persuasion (Van der Linden, 2024; Van der Linden et al., 2017).

After reading the media report, we asked participants to respond to seven questions. The first item served as an attention check, asking about the topic of the material participants read (two participants did not select non-verbal deception cues and were excluded from the sample). The seventh item also served as an attention check, asking participants to pick the Number 2 from the choices (no one failed this check). We based Items 2–4, which were about perceived expert consensus, on previous literature (Han et al., 2024; Koehler, 2016). Specifically, the second item measured the extent to which experts agree that people cannot use non-verbal cues to detect deception accurately (i.e. we used a 7-point scale, where 1 = *very little* and 7 = *very much*). The third item asked participants to provide the likelihood of two randomly selected experts from the survey sharing the same opinion that people cannot use non-verbal cues to detect deception accurately (i.e. we used a slider scale ranging from 0% to 100% agreement). The fourth item asked about perceptions of consensus in the expert community: 'Suppose 100 different deception detection experts, with similar qualifications, were surveyed about this same issue. How many of these new experts do you think would agree that nonverbal cues cannot be used to detect deception accurately?'. We used a slider scale from 0 to 100 to record responses. Moreover, we asked participants to rate their level of agreement (1 = *strongly disagree*, 7 = *strongly agree*) with the following two statements about using non-verbal cues to detect deception: 'The police should be allowed to use nonverbal cues to detect deception (e.g. during police stops, when deciding whom to interrogate)', and 'Using nonverbal cues to detect deception should be prohibited for security screening purposes (e.g. airport security, job hiring)'.

Statistical analysis

We used a Bayesian approach to analyse our results to estimate the magnitude of the difference among groups and, most importantly, to assess (un)certainty about the estimated true difference between conditions (Gelman et al., 2014). In particular, the Bayesian approach uses evidence from the data collected to update prior beliefs about the compatible magnitudes of differences between the balance and no balance conditions (Kruschke, 2013). After accounting for the collected data and our prior assumptions, *the degree of belief about the true difference is expressed as a posterior probability distribution* (Etz & Vandekerckhove, 2018). The estimated mode and highest density intervals (i.e. HDI) describe our skewed posterior probability distributions. As a type of posterior interval, HDIs represent a range of parameter values most compatible with the data, summarizing the posterior probability mass rather than accepting or rejecting hypotheses (Gelman et al., 2014; McElreath, 2020).

For example, a 95% HDI summarizes the estimated range of the true difference containing 95% of the posterior probabilities, and likewise, an 80% HDI summarizes the estimated range with 80% of the posterior probabilities.

We carried out Bayesian parameter estimation using R (version 4.1.1; R Core Team, 2020) and the *brms* package (version 2.16.1; Bürkner, 2017). We fitted skew-normal models to estimate differences between the no-balance and balance conditions, as our data on perceptions of expert consensus were left-skewed (skewness ranged from −0.90 to −1.38) due to participants responding near scale boundaries after reading a high-consensus statement. We chose the unstandardized difference as our parameter of interest, because it intuitively captures the magnitude of differences in the dependent measures. We used weakly informative priors to make our inferences as objective as possible. Full details for software, Bayesian model metrics and priors are available at https://osf.io/tra9w/, including data and R codes. Although we report means, standard deviations and Cohen's *d* effect sizes, results from NHST analyses are also provided at the above link for readers who may wish to view our analyses framed in a more familiar light.

RESULTS

Table 1 contains the descriptive statistics for a combination of variables as a function of dependent measures, and Table 2 contains the associated effect sizes (i.e. Cohen's *d*). We found, on average, that participants were 65% certain that people can use non-verbal cues to detect lies accurately and were 70% certain that experts can use non-verbal cues to detect lies accurately.

Perception of expert consensus

The correlations between the three measures of perceived expert consensus (i.e. perceived expert agreement, the likelihood of consensus among two experts randomly selected from the survey and perceived expert consensus in the science community) and participants' personal beliefs about using non-verbal cues for deception detection were negligible ($r = -.06$, $-.01$, $-.14$, respectively). Similarly,

TABLE 1 Means and standard deviations for dependent measures.

	Level of balance				
Dependent measures	**No balance**	**Balance**	**Evidentiary balance**	**Balance + forewarning**	**Evidentiary balance + forewarning**
Perceived expert agreement (1–7)	5.26 (2.36)	5.14 (2.07)	4.71 (2.13)	4.97 (2.08)	5.21 (1.72)
Likelihood of two experts surveyed sharing the same opinion (0–100)	81.66 (15.87)	70.60 (28.42)	68.18 (25.80)	71.42 (23.27)	72.30 (22.26)
Perceived consensus in the science community (0–100)	81.02 (18.93)	71.72 (24.90)	71.08 (22.52)	73.68 (23.29)	73.52 (21.36)
The police should be allowed to use non-verbal cues to detect deception (1–7, not reverse-scored)	2.86 (1.74)	3.38 (1.78)	3.15 (1.67)	3.10 (1.74)	3.17 (1.65)
Using non-verbal cues to detect deception should be prohibited for security screening purposes (1–7)	4.49 (1.91)	4.17 (1.94)	4.61 (1.66)	4.75 (1.71)	4.63 (1.71)

Note: Balance = three agree versus three disagree; evidentiary balance = five agree versus one disagree.

TABLE 2 Effect sizes (Cohen's *d*) for group comparisons by dependent measures.

Cohen's d	$d_{3v3\ v.\ nb}$	$d_{5v1\ v.\ nb}$	$d_{5v1\ v.\ 3v3}$	$d_{3v3+forewarning\ v.\ 3v3}$	$d_{5v1+forewarning\ v.\ 5v1}$
Perceived expert consensus (1–7)	−0.05	−0.25	−0.21	−0.08	0.26
Likelihood of two experts surveyed sharing the same opinion (0–100)	−0.47	−0.62	−0.11	0.04	0.17
Perceived consensus in the science community (0–100)	−0.42	−0.48	−0.03	0.08	0.11
The police should be allowed to use non-verbal cues to detect deception (1–7, not reverse-scored)	0.30	0.17	−0.13	−0.17	0.01
Using non-verbal cues to detect deception should be prohibited for security screening purposes (1–7)	−0.17	0.07	0.24	0.31	0.01

the correlations between the three measures of perceived expert consensus and participants' views on expert beliefs about using non-verbal cues to detect deception were negligible ($r = -.07$, $-.05$ -.08, respectively).

The effect sizes for the difference in the three measures of perceived expert consensus (i.e. perceived expert agreement, likelihood of consensus among two experts randomly selected from the survey, and perceived expert consensus in the science community) between conditions ranged from 0.05 to 0.62 (see Table 2). Compared to reading only WOE data (no balance), reading evenly balanced comments (three agree vs. three disagree) alongside the data decreased perceived likelihood of consensus among surveyed experts ($d = 0.47$) and perceived consensus in the science community ($d = 0.42$). Compared to no balance, reading evidentiary balanced comments (five agree vs. one disagree) decreased all three measures of perceptions of expert consensus (*d*s = 0.25, 0.62, 0.48). Moreover, compared to no forewarning, forewarning readers about a 'fake debate' slightly improved perceptions of expert consensus only when exposed to evidentiary balanced comments (i.e. five agree vs. one agree; *d*s = 0.26, 0.17, 0.11). There was a negligible effect of forewarning when exposed to evenly balanced comments (i.e. three agree vs. three disagree).

Posterior probability distributions

Figure 1 depicts the posterior probability distributions of the difference among all conditions for the direct measure of perceived expert agreement. Of interest, the best-fitting model suggests that there is only a 60% chance that, compared to reading the WOE data (no balance), reading evidentiary balanced comments (five agree vs. one disagree) along with the data decreases perceived expert agreement that non-verbal cues cannot detect deception accurately (HDI = −0.52, 0.04, *posterior Mode* = −0.25; see Figure 1). However, compared to no forewarning, there is an 80% chance that forewarning about a 'fake debate' inoculates people against the negative impact of the evidentiary balanced comments (5:1) on perceived expert agreement (HDI = 0.06, 0.85, *posterior Mode* = 0.43). There is also a possibility that evenly balanced comments (three agree vs. three disagree) will not impact perceived expert agreement (*95%* HDI = −0.52, 0.71, *posterior mode* = 0.03) and that forewarning about a 'fake debate' does not negate its impact (*95%* HDI = −0.73, 0.49, *posterior mode* = −0.08).

Figure 2 shows the posterior probability distributions of the difference among all conditions for the likelihood of two experts surveyed sharing the same opinions. For this measure, the best-fitting model suggests a 95% chance that, compared to reading the WOE data, reading evenly balanced comments

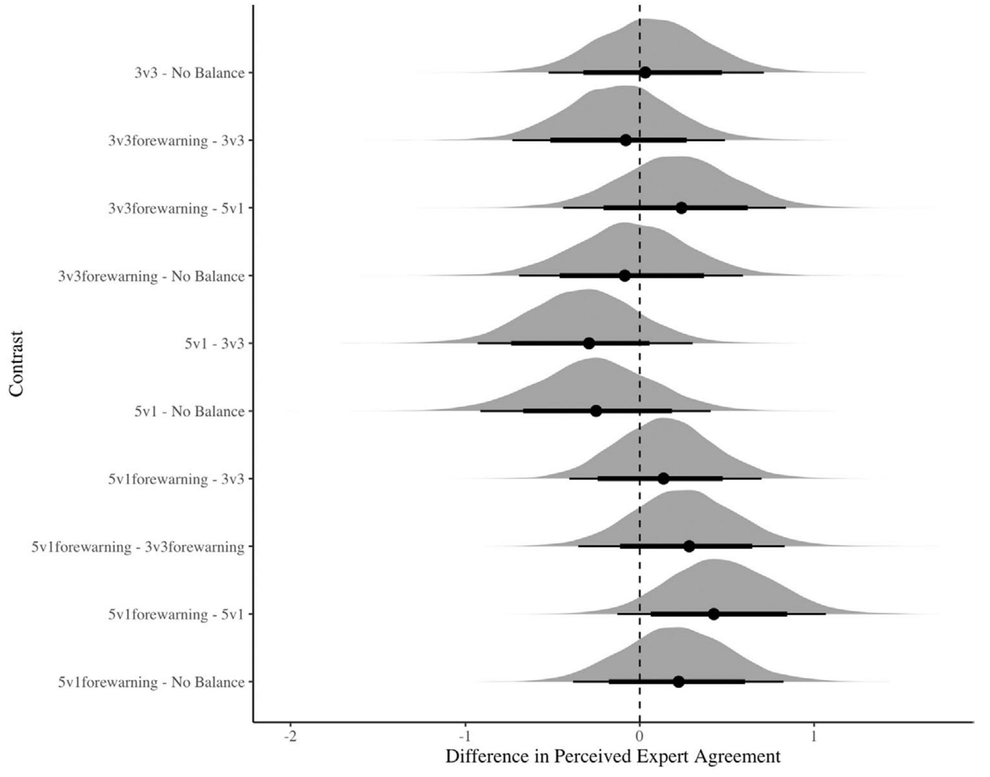

FIGURE 1 The posterior probability distribution of the mean differences among the conditions for perceived expert agreement.

(three agree vs. three disagree) with the data decreases estimations of there being consensus among experts surveyed, even when providing the data about expert consensus (HDI = −19.00, −5.17, *posterior mode* = −11.90; see Figure 2). It is also 95% likely that, compared to reading only the WOE data, reading evidentiary balanced comments (five agree vs. one disagree) with the data decreases the estimated likelihood that the surveyed experts would agree (HDI = −19.40, −5.55, *posterior mode* = −13.00). Exposure to evenly balanced comments (3:3) may similarly impact perceptions of consensus among experts surveyed as an evidentiary balanced message (5:1). Moreover, our model suggests a 75% chance that, compared to no forewarning, forewarning people about a 'fake debate' when reading evidentiary balanced comments (5:1) improves the estimated likelihood of experts sharing the same opinion (HDI = −0.04, 8.85, *posterior mode* = 4.64). However, forewarning might not mitigate the negative impact of two-sided messages. Even with forewarning, reading evenly balanced comments (3:3) or evidentiary balanced comments (5:1) reduces the perceived likelihood of experts agreeing more than only reading the WOE data (95% HDI = −16.10, −3.34, *posterior mode* = −9.02; 95% HDI = −14.30, −2.05, *posterior mode* = −8.00, respectively).

Figure 3 shows the posterior probability distributions of the difference among all conditions for perceived consensus in the science community. As shown in Figure 3, our model suggests a 95% chance that compared to reading the WOE data (no balance), reading evenly balanced comments (three agree vs. three disagree) alongside the data decreases perceptions about expert consensus in the broader scientific community (HDI = −13.00, −1.47, *posterior mode* = −7.28). The figure also shows a 95% chance that, compared to no balance, exposure to evidentiary balanced comments (five agree vs. one disagree) decreases the perceived consensus in the scientific community (HDI = −12.00, −0.25, *posterior mode* = −5.56). However, compared to no forewarning, forewarning people about a 'fake debate' does

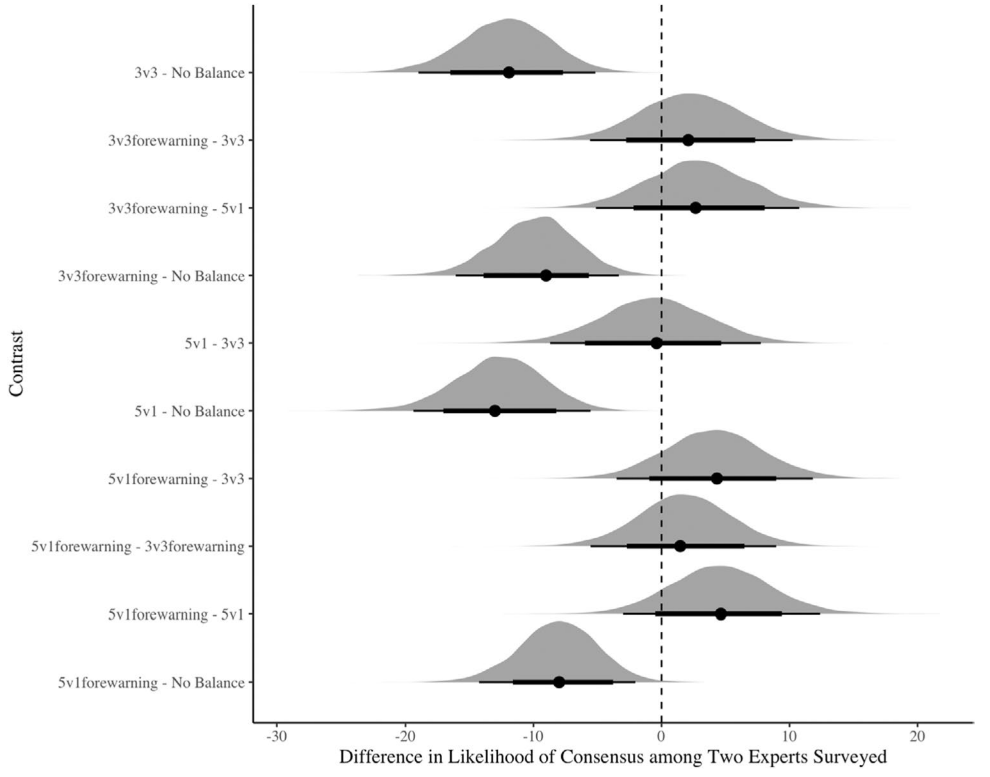

FIGURE 2 The posterior probability distribution of the mean differences among the conditions for likelihood of two experts surveyed sharing the same opinions.

not seem to improve the perceived consensus in the scientific community, neither when they read evenly balanced comments (3:3; 95% HDI = −4.04, 9.49, *posterior mode* = 2.99) nor evidentiary balanced comments (5:1; 95% HDI = −4. 53, 8.36, *posterior mode* = 1.81).

Policy support

There was a small positive correlation between ratings of support for police use of non-verbal cues and participants' personal beliefs about the usefulness of non-verbal cues in detecting deception ($r = .35$) and their views on expert beliefs in non-verbal lie detection ($r = .24$). There was also a slight negative correlation between ratings about prohibiting non-verbal cues in security screening and personal beliefs about non-verbal lie detection ($r = -.20$), and a negligible correlation with their views on expert beliefs ($r = -.08$). That is, the more people believe in non-verbal lie detection, the more likely they are to support anti-science policies. However, participants generally disapproved of using non-verbal cues to detect deception during police interrogations, traffic stops or security screenings after reading a media report that cited the numerical weight regarding expert consensus on the issue (i.e. 90% of experts agreed that non-verbal cues cannot detect deception accurately; see Table 1).

The data in Table 2 showed that, compared to reading the WOE data (no balance), reading evenly balanced comments (three agree vs. three disagree) with the data increased the support for police use of non-verbal cues and decreased the resistance to security screening via non-verbal cues (*d*s = 0.30, 0.17). Compared to no balance, evidentiary balanced comments (five agree vs. one disagree) also increased the support for the police to detect lies using non-verbal cues ($d = 0.17$). While forewarning about a ‘fake

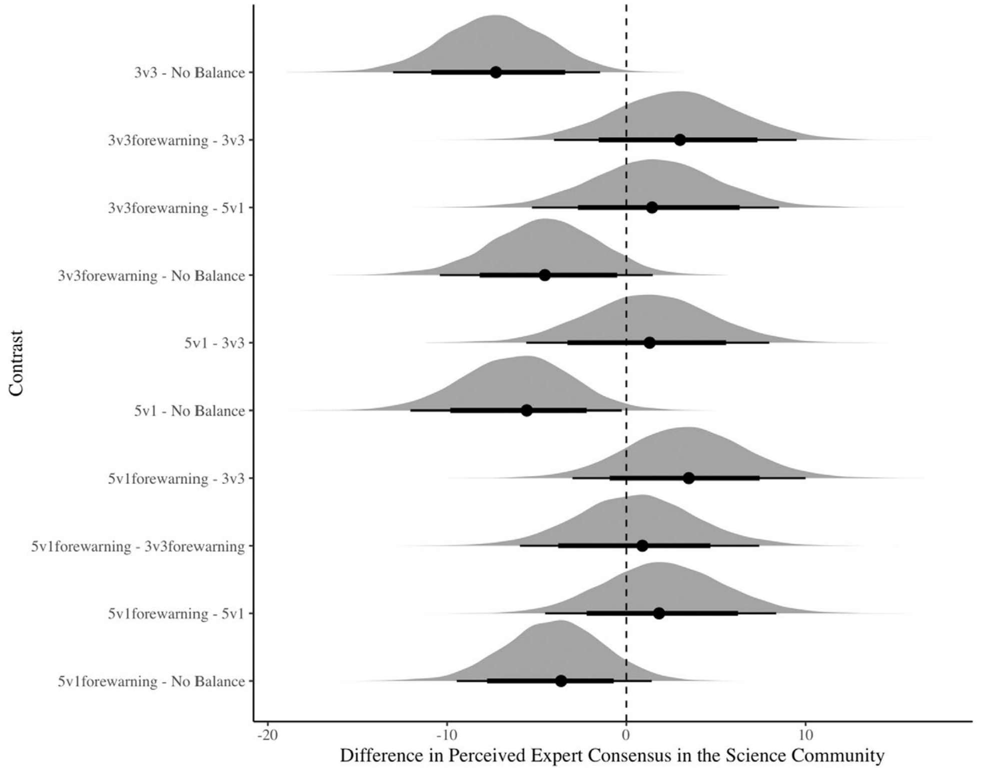

FIGURE 3 The posterior probability distribution of the mean differences among the conditions for perceived consensus in the science community.

debate' did not influence policy support for non-verbal cues when people read evidentiary balanced comments (5:1), it somewhat inoculated the negative impact of the evenly balanced comments on the aforementioned policy measures (3:3; *d*s = 0.17, 0.31).

Posterior probability distributions

Our model suggests an 80% chance that, compared to reading the WOE data (no balance), reading evenly balanced comments (three agree vs. three disagree) with the data increases the support for the police to use non-verbal cues to detect deception compared to only reading the data (HDI = 0.11, 0.91, *posterior mode* = 0.52, see Figure 4). There is roughly a 70% chance that, compared to no balance, reading evidentiary balanced comments (five agree vs. one disagree) increases the support for police use of non-verbal cues (HDI = −0.03, 0.62, *posterior mode* = 0.33). Forewarning might mitigate the impact of evenly balanced comments (3:3; 70% HDI = −0.58, 0.07, *posterior mode* = −0.27) but does not influence people when they view evidentiary balanced comments (5:1).

Figure 5 shows an 80% chance that, compared to no balance, reading evenly balanced comments (3:3) reduces people's support to ban the use of non-verbal cues in security screenings compared to the no balance message (HDI = −0.74, 0.09, *posterior mode* = −0.32). However, it shows an 80% chance that, compared to no forewarning, forewarning people about a 'fake debate' will mitigate the negative impact of evenly balanced comments on views about prohibiting the use of non-verbal cues (HDI = 0.16, 0.98, *posterior mode* = 0.50). It is also possible that, compared to no balance, reading evidentiary balanced

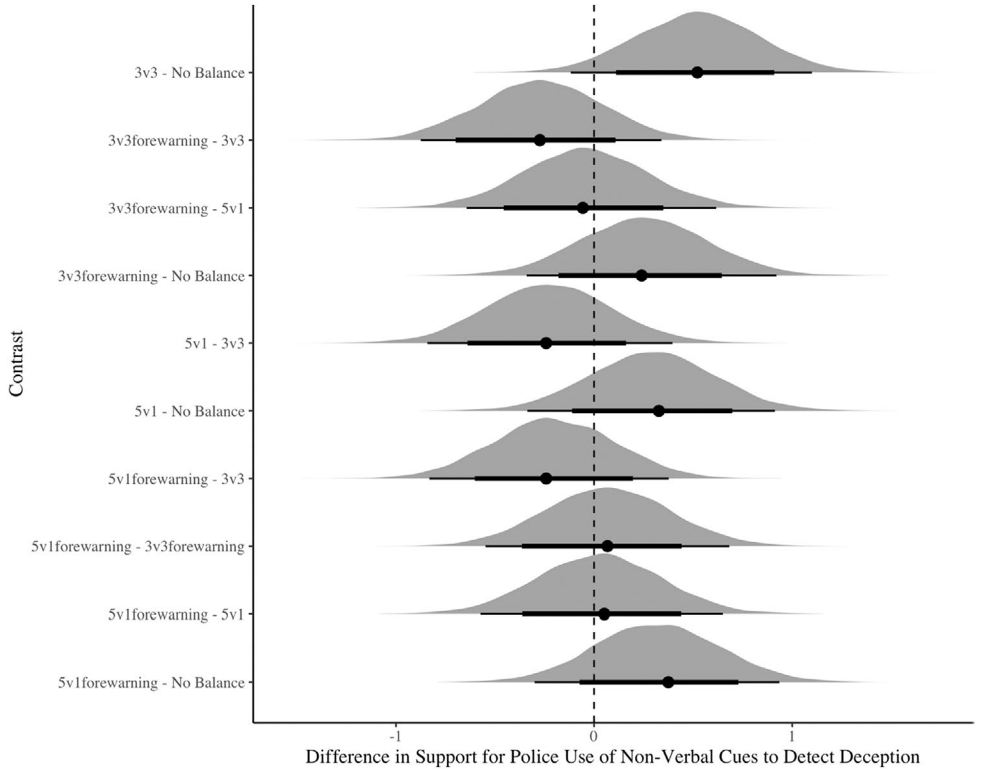

FIGURE 4 The posterior probability distribution of the mean differences among the conditions for support for police use of non-verbal cues to detect deception.

comments (5:1) may not influence the support of a policy banning non-verbal cues for security screening purposes.

DISCUSSION

Consistent with previous research, but using a more ecologically valid experimental paradigm (i.e. reading a media report) we found that viewing any two-sided message after learning that nearly 90% of deception detection experts agreed that non-verbal cues are useless for deception detection (a high consensus issue) decreased perceptions of scientific consensus. In other words, exposing people to an equal number of comments from opposing experts (three agree vs. three disagree; evenly balanced message) or two-sided messages with more consensus comments (five agree vs. one disagree; evidentiary balanced message) can move them further from what the actual data say about a topic (Han et al., 2024, 2025; Koehler, 2016). We also found that presenting evenly balanced comments (3:3) reduced support for a policy (i.e. a ban on using non-verbal cues by practitioners) grounded in scientific consensus. Such findings align with the persuasion literature that personal testimonials are often more influential than objective statistics in shaping risk perceptions and intentions (Borgida & Nisbett, 1977; De Wit et al., 2008; Zebregs et al., 2015). Additionally, we extended the generalisability of the false-balance effect to the topic of deception detection, where people may hold incorrect views (i.e. in our study, participants were 65% certain that non-verbal cues could accurately detect lies and were 70% certain that experts also believed in non-verbal lie detection).

Research has indicated that communicating descriptive norms, such as the fact that 97% of scientists agree humans are causing climate change, enhances perceived scientific consensus compared to no

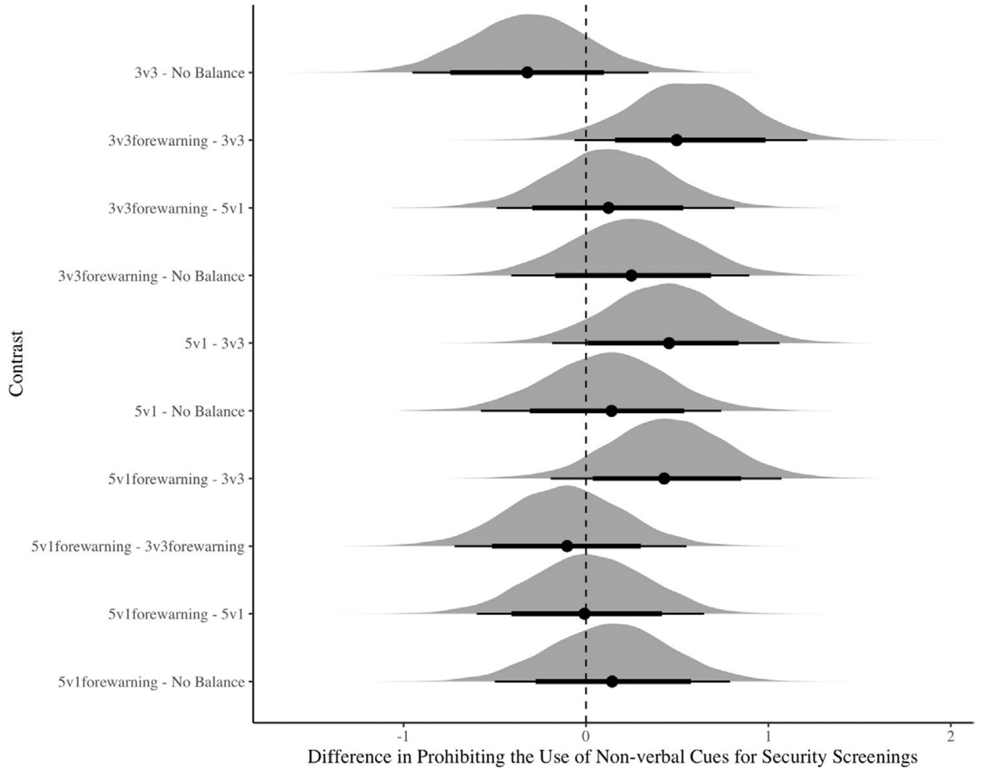

FIGURE 5 The posterior probability distribution of the mean differences among the conditions for prohibiting use of non-verbal cues for security screening purposes.

communication (data > nothing; Maertens et al., 2020; Van der Linden et al., 2017) and can improve perceived consensus when misinformation is present (data + misinformation > misinformation; Cook et al., 2017; Dixon et al., 2015). However, the positive effect of presenting data alone is nullified when misinformation is introduced, resulting in the perceived consensus equivalent to providing no information (data + misinformation = nothing; Maertens et al., 2020; Van der Linden et al., 2017). Similarly, we found that after participants learned that 90% of deception detection experts agreed that non-verbal cues are ineffective for detecting lies, they acknowledged this consensus by estimating expert agreement between 75 and 82 on a scale of 100 (no-balance, data-only control condition, see Table 1). Yet, perceived expert consensus declined when opposing expert comments were presented alongside statistical data, and balanced comments further reduced support for policies based on scientific consensus (e.g. non-verbal lie detection should be prohibited). In line with previous research (Van der Linden et al., 2017), our findings suggest that pairing consensus data with conflicting arguments limits the positive impact of the data, highlighting the challenges of conveying scientific agreement in the presence of misinformation and competing narratives.

We also found that varying the number of comments on each side did not influence perceptions of expert consensus. Namely, exposure to evidentiary balanced messages (five agree vs. one disagree) did not result in a better estimation of expert consensus compared to evenly balanced messages (three agree vs. three disagree)—it might even backfire a little. This finding is inconsistent with previous research that revealed evidentiary balanced comments (5:1) led to more accurate perceptions of expert consensus than evenly balanced comments (3:3; Han et al., 2024, 2025). One possible explanation is rooted in people's pre-existing beliefs contradicting the scientific consensus while those beliefs tend to be neutral or consistent in other studies (Cook et al., 2017; Han et al., 2025). Another possible explanation could be the absence of a multiple-source effect. Unlike previous research that examined partition dependence (by varying the number of comments and the number of sources on each side), we only varied the number of comments; that is, one expert making multiple arguments on each side of the 'debate'. Rather than focusing on the bulk of the content, people may become even less convinced of the advocated position when exposed to the evidentiary balanced comments, since all five agreeing comments come from a single expert. Research suggests that multiple arguments from multiple sources generate more favourable attitudes towards the advocated position than multiple arguments from a single source (Harkins & Petty, 1981, 1987). Moreover, if the false-balance effect is due to the mere frequency of consensus-aligned comments instead of partition account (Han et al., 2024), perceptions of expert consensus should still be distorted even when all of the comments come from a single source on each side (e.g. ½, ⅚, ⅙; Han et al., 2025). Therefore, partition dependence remains a plausible mechanism underlying the false-balance effect. It may be that both the number of arguments and sources, rather than just the number of arguments, contribute to the perceived 'partition'.

While forewarning has proven to be an effective psychological inoculation against the false-balance effect with regard to perceived expert consensus (forewarning + false balance > false balance; Cook et al., 2017; Schmid et al., 2020; Van der Linden et al., 2017), our findings indicate that it may only have a minimal impact on correcting the underestimation of WOE; at least with respect to the use of non-verbal cues to detect deception (data + forewarning + false balance ≈ data + false balance). However, our findings suggest that forewarning people about the fake debate might improve support for policies rooted in the scientific consensus when balanced comments are presented. One potential explanation is that forewarning is ineffective in building resistance against balanced messages, particularly when estimating scientific consensus after viewing the WOE data. This aligns with previous findings that inoculation offers limited additional benefit when consensus information (i.e. WOE) is already presented; for example, Cook et al. (2017) found that combining consensus information with an inoculation message (forewarning) before balanced messages results in similar results to presenting consensus alone.

We also found that forewarning about fake debates might inoculate against evidentiary balanced comments (5:1), but not against evenly balanced comments (3:3), when paired with WOE data. It is possible that our manipulation (i.e. a general forewarning about fake debates) is insufficient to counteract the distorting effect of balanced presentations (1:1, 3:3) on the perceived strength of scientific evidence. Previous research has demonstrated that a detailed forewarning could lead to a larger inoculation effect compared to a general forewarning (Van der Linden et al., 2017). Additionally, in the case of evidentiary balance, where the distribution of expert comments (5:1≈83:17) closely matches actual expert consensus (90:10), our forewarning may override a single dissenting opinion with the help of that extra 'weight' information. Therefore, an adequate forewarning may need to be longer, more specific, and include more details to explain the representative heuristics rather than simply explaining a fake debate denial strategy.

Another explanation for the small effect of forewarning is grounded in the people's personal beliefs and attitudes. When people's beliefs align with scientific consensus, or there is some uncertainty about an issue, forewarning about false-balance misinformation might generate an attitude shift towards their impending position in the face of an impending counterattitudinal appeal (Wood & Quinn, 2003). Nevertheless, when people have beliefs that contradict scientific consensus, as on the topic of non-verbal lie detection, forewarning may fail to override false balance and to shift attitudes back towards the consensus position right after reading WOE data.

CONCLUSION

The field of deception detection has long reached scientific consensus that the relationship between non-verbal cues and deception is unreliable and weak (DePaulo et al., 2003). Even so, misinformation about 'reading lies' through non-verbal communication channels has shaped public opinions on deception detection and continues to attract interest across many domains. What makes matters worse is that exposure to opposing 'expert' opinions obscures the public's understanding of the overwhelming strength of established evidence, leaving people more confused than informed. Although our forewarning about fake debates over the debunked non-verbal lie detection fell short in countering the power of false balance to overshadow consensus data, a glimpse of hope lies in communicating these descriptive norms—which indeed deliver empirical facts to the public.

AUTHOR CONTRIBUTIONS

Martin V. Day: Conceptualization; supervision. **Brent Snook:** Conceptualization; methodology; validation; resources; supervision; writing – original draft; writing – review and editing. **Tianshuang Han:** Conceptualization; methodology; software; data curation; investigation; validation; formal analysis; visualization; project administration; writing – original draft; writing – review and editing.

FUNDING INFORMATION

We did not receive any funding to assist with the submitted work.

CONFLICT OF INTEREST STATEMENT

We have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

The data supporting this study's findings are available at https://osf.io/tra9w/.

ETHICAL APPROVAL

This research has been approved by the Interdisciplinary Committee on Ethics in Human Research (ICEHR) and found to comply with Memorial University's ethics policy. The ICEHR approval number given is 20240801-SC.

ORCID

Tianshuang Han https://orcid.org/0000-0002-3549-6830
Brent Snook https://orcid.org/0000-0002-2739-5474
Martin V. Day https://orcid.org/0000-0001-5274-1626

REFERENCES

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

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

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2

Booth, R. (2025). Meta to get rid of factcheckers and recommend more political content. *The Guardian*. https://www.theguardian.com/technology/2025/jan/07/meta-facebook-instagram-threads-mark-zuckerberg-remove-fact-checkers-recommend-political-content

Borgida, E., & Nisbett, R. E. (1977). The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology*, *7*(3), 258–271. https://doi.org/10.1111/j.1559-1816.1977.tb00750.x

Boykoff, M. T., & Boykoff, J. M. (2004). Balance as bias: Global warming and the US prestige press. *Global Environmental Change*, *14*(2), 125–136. https://doi.org/10.1016/j.gloenvcha.2003.10.001

Braddock, K. (2022). Vaccinating against hate: Using attitudinal inoculation to confer resistance to persuasion by extremist propaganda. *Terrorism and Political Violence*, *34*(2), 240–262. https://doi.org/10.1080/09546553.2019.1693370

Bürkner, P. C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Caulfield, T. (2023, August 28). *Is bothsidesism killing us? (And why scientific consensus matters)*. Healthydebate. https://healthydebate.ca/2023/08/topic/bothsidesism-scientific-consensus-matters/

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, *66*(3), 460–473. https://doi.org/10.1037/0022-3514.66.3.460

Cook, J. (2022). Understanding and avoiding false balance media coverage of climate change. In D. Harwardt, L. Barthel, N. Schumann, R. Brix, & T. Böttiger (Eds.), *Media in the climate crisis* (pp. 65–72). Oekom.

Cook, J., Lewandowsky, S., & Ecker, U. K. (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

De Wit, J. B. F., Das, E., & Vet, R. (2008). What works best: Objective statistics or a personal testimonial? An assessment of the persuasive effects of different types of message evidence on risk perception. *Health Psychology*, *27*(1), 110–115. https://doi.org/10.1037/0278-6133.27.1.110

Dearing, J. W. (1995). Newspaper coverage of maverick science: Creating controversy through balancing. *Public Understanding of Science*, *4*(4), 341–361. https://doi.org/10.1088/0963-6625/4/4/002

Denault, V. (2020). Misconceptions about nonverbal cues to deception: A covert threat to the justice system? *Frontiers in Psychology*, *11*, 573460. https://doi.org/10.3389/fpsyg.2020.573460

Denault, V., Plusquellec, P., Jupe, L. M., St-Yves, M., Dunbar, N. E., Hartwig, M., Sporer, S. L., Rioux-Turcotte, J., Jarry, J., Walsh, D., Otgaar, H., Viziteu, A., Talwar, V., Keatley, D. A., Blandón-Gitlin, I., Townson, C., Deslauriers-Varin, N., Lilienfeld, S. O., Patterson, M. L., … van Koppen, P. J. (2020). The analysis of nonverbal communication: The dangers of pseudoscience in security and justice contexts. *Anuario de Psicología Jurídica*, *30*(1), 1–12. https://doi.org/10.5093/apj2019a9

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74

Dixon, G. N., & Clarke, C. E. (2013). Heightening uncertainty around certain science: Media coverage, false balance, and the autism-vaccine controversy. *Science Communication*, *35*(3), 358–382. https://doi.org/10.1177/1075547012458290

Dixon, G. N., McKeever, B. W., Holton, A. E., Clarke, C., & Eosco, G. (2015). The power of a picture: Overcoming scientific misinformation by communicating weight-of-evidence information with visual exemplars. *Journal of Communication*, *65*(4), 639–659. https://doi.org/10.1111/jcom.12159

Dunwoody, S. (2005). Weight-of-evidence reporting: What is it? Why use it? *Nieman Reports*, *59*(4), 89–91. https://www.proquest.com/scholarly-journals/weight-evidence-reporting-what-is-why-use/docview/216754041/se-2?accountid=12378

Ecker, U. K., 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

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin & Review*, *25*(1), 5–34. https://doi.org/10.3758/s13423-017-1262-3

Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, *4*(2), 271–276. https://doi.org/10.3758/BF03209405

Gelman, A., Carlin, J. C., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Chapman & Hall/CRC.

Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior*, *21*(2), 207–219. https://doi.org/10.1016/S0022-5371(82)90571-0

Han, T., Snook, B., & Day, M. V. (2024). Not every story has two sides: The effect of false balance on perceived scientific consensus about interrogation practices. *Journal of Criminal Psychology*. https://doi.org/10.1108/JCP-09-2024-0091

Han, T., Snook, B., & Day, M. V. (2025). The false balance effect: Exploring partition dependence as a potential explanation. *Applied Cognitive Psychology*, *39*(1), e70025. https://doi.org/10.1002/acp.70025

Hansson, S. O. (2017). Science denial as a form of pseudoscience. *Studies in History and Philosophy of Science Part A*, *63*, 39–47. https://doi.org/10.1016/j.shpsa.2017.05.002

Harkins, S. G., & Petty, R. E. (1981). The multiple source effect in persuasion: The effects of distraction. *Personality and Social Psychology Bulletin*, *7*(4), 627–635. https://doi.org/10.1177/014616728174019

Harkins, S. G., & Petty, R. E. (1987). Information utility and the multiple source effect. *Journal of Personality and Social Psychology*, *52*(2), 260–268. https://doi.org/10.1037/0022-3514.52.2.260

Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2016). Does training improve the detection of deception? A meta-analysis. *Communication Research*, *43*(3), 283–343. https://doi.org/10.1177/0093650214534974

Imundo, M. N., & Rapp, D. N. (2022). When fairness is flawed: Effects of false balance reporting and weight-of-evidence statements on beliefs and perceptions of climate change. *Journal of Applied Research in Memory and Cognition*, *11*(2), 258–271. https://doi.org/10.1016/j.jarmac.2021.10.002

Kassin, S. M., Redlich, A. D., Alceste, F., & Luke, T. J. (2018). On the general acceptance of confessions research: Opinions of the scientific community. *American Psychologist*, *73*(1), 63–80. https://doi.org/10.1037/amp0000141

Koehler, D. J. (2016). Can journalistic "false balance" distort public perception of consensus in expert opinion? *Journal of Experimental Psychology: Applied*, *22*(1), 24–38. https://doi.org/10.1037/xap0000073

Kohl, P. A., Kim, S. Y., Peng, Y., Akin, H., Koh, E. J., Howell, A., & Dunwoody, S. (2016). The influence of weight-of-evidence strategies on audience perceptions of (un) certainty when media cover contested science. *Public Understanding of Science*, *25*(8), 976–991. https://doi.org/10.1177/0963662515615087

Kortenkamp, K. V., & Basten, B. (2015). Environmental science in the media: Effects of opposing viewpoints on risk and uncertainty perceptions. *Science Communication*, *37*(3), 287–313. https://doi.org/10.1177/1075547015574016

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. https://doi.org/10.1037/a0029146

Kurz-Milcke, E., & Gigerenzer, G. (Eds.). (2004). *Experts in science and society*. Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/b105826

Lewandowsky, S., & Yesilada, M. (2021). Inoculating against the spread of Islamophobic and radical-Islamist disinformation. *Cognitive Research: Principles and Implications*, *6*, 1–15. https://doi.org/10.1186/s41235-021-00323-z

Lewandowsky, S., Ecker, U. K., & 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

Lilienfeld, S. O. (1998). Pseudoscience in contemporary clinical psychology: What it is and what we can do about it. *The Clinical Psychologist*, *51*(4), 3–9.

Lilienfeld, S. O., Fowler, K. A., Lohr, J. M., & Lynn, S. J. (2005). Pseudoscience, nonscience, and nonsense in clinical psychology: Dangers and remedies. In R. H. Wright & N. A. Cummings (Eds.), *Destructive trends in mental health: The well-intentioned path to harm* (pp. 187–218). Routledge.

Luke, T. J., Giolla, E. M., Memon, A., Landström, S., Granhag, P. A., & Kassin, S. (2023). What have we learned about cues to deception? A survey of expert opinions. *Psychology, Crime & Law*, *31*, 1–20. https://doi.org/10.1080/1068316X.2023.2292504

Maertens, R., Anseel, F., & van der Linden, S. (2020). Combatting climate change misinformation: Evidence for longevity of inoculation and consensus messaging effects. *Journal of Environmental Psychology*, *70*, 101455. https://doi.org/10.1016/j.jenvp.2020.101455

Matsumoto, D. (2023, October 13). *The latest science about behavioral indicators of deception*. Humanintell. https://www.humintell.com/2023/10/the-latest-science-about-behavioral-indicators-of-deception/

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman & Hall/CRC Press.

McGuire, W. J. (1961a). The effectiveness of supportive and refutational defenses in immunizing and restoring beliefs against persuasion. *Sociometry*, *24*(2), 184–197. https://doi.org/10.2307/2786067

McGuire, W. J. (1961b). 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. *Advances in Experimental Social Psychology*, *1*, 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

McNally, R. J. (2003). The demise of pseudoscience. *The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology, Psychiatry, and Social Work*, *2*(2), 97–101.

Neuschatz, J. S., Benoit, G. E., & Payne, D. G. (2003). Effective warnings in the Deese- Roediger-McDermott false-memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 35–41. https://doi.org/10.1037/0278-7393.29.1.35

Patterson, M. L., Fridlund, A. J., & Crivelli, C. (2023). Four misconceptions about nonverbal communication. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *18*(6), 1388–1411. https://doi.org/10.1177/17456916221148142

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

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

Schmid, P., Schwarzer, M., & Betsch, C. (2020). Weight-of-evidence strategies to mitigate the influence of messages of science denialism in public discussions. *Journal of Cognition*, *3*(1), 36. https://doi.org/10.5334/joc.125

Schopen, K., Otgaar, H., Howe, M. L., & Muris, P. (2022). Effects of forewarnings on children's and adults' spontaneous false memories. *European Journal of Developmental Psychology*, *19*(2), 177–197. https://doi.org/10.1080/17405629.2021.1904877

Smith, F., Simchon, A., Holford, D., & Lewandowsky, S. (2025). Inoculation reduces social media engagement with affectively polarized content in the UK and US. *Communications Psychology*, *3*(1), 11. https://doi.org/10.1038/s44271-025-00189-7

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law*, *13*(1), 1–34. https://doi.org/10.1037/1076-8971.13.1.1

Turuba, R., Zenone, M., Srivastava, R., Stea, J., Quintana, Y., Ow, N., Marchand, K., Kwan, A., Ong, A. J., Ding, X., Warren, C., Marcon, A. R., Henderson, J., Mathias, S., & Barbic, S. (2025). Do you have depression? A summative content analysis of mental health-related content on TikTok. *DIGITAL HEALTH*, *11*, 20552076241297062. https://doi.org/10.1177/20552076241297062

Van der Linden, S. (2024). Countering misinformation through psychological inoculation. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. *69*, pp. 1–58). Elsevier. https://doi.org/10.1016/bs.aesp.2023.11.001

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

Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, *70*, 295–337. https://doi.org/10.1146/annurev-psych-010418-103135

Weatherbed, J. (2025, January 17). Google rejects EU fact-checking commitments for Search and YouTube. *The Verge*. https://www.theverge.com/2025/1/17/24345747/google-eu-dsa-fact-checks-disinformation-code-search-youtube

Wood, M. L. (2007). Rethinking the inoculation analogy: Effects on subjects with differing preexisting attitudes. *Human Communication Research*, *33*(3), 357–378. https://doi.org/10.1111/j.1468-2958.2007.00303.x

Wood, W., & Quinn, J. M. (2003). Forewarned and forearmed? Two meta-analysis syntheses of forewarnings of influence appeals. *Psychological Bulletin*, *129*(1), 119–138. https://doi.org/10.1037/0033-2909.129.1.119

Yeung, A., Ng, E., & Abi-Jaoude, E. (2022). TikTok and attention-deficit/hyperactivity disorder: A cross-sectional study of social media content quality. *The Canadian Journal of Psychiatry*, *67*(12), 899–906. https://doi.org/10.1177/07067437221082854

Zebregs, S., van den Putte, B., Neijens, P., & de Graaf, A. (2015). The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: A meta-analysis. *Health Communication*, *30*(3), 282–289. https://doi.org/10.1080/10410236.2013.842528

Zenone, M., Snyder, J., Marcon, A., & Caulfield, T. (2022). Analyzing natural herd immunity media discourse in the United Kingdom and the United States. *PLOS Global Public Health*, *2*(1), e0000078. https://doi.org/10.1371/journal.pgph.0000078

Zenone, M., Snyder, J., van Schalkwyk, M., Bélisle-Pipon, J. C., Hartwell, G., Caulfield, T., & Maani, N. (2024). Alternative cancer clinics' use of Google listings and reviews to mislead potential patients. *BJC Reports*, *2*(1), 55. https://doi.org/10.1038/s44276-024-00071-9

Zenone, M., van Schalkwyk, M., Hartwell, G., Caulfield, T., & Maani, N. (2024). Selling misleading "cancer cure" books on Amazon: Systematic search on Amazon.com and thematic analysis. *Journal of Medical Internet Research*, *26*, e56354. https://doi.org/10.2196/56354

How to cite this article: Han, T., Snook, B., & Day, M. V. (2025). Fore! Does forewarning inoculate people against the false balance effect? *Legal and Criminological Psychology*, *00*, 1–19. https://doi.org/10.1111/lcrp.70004