nature human behaviour

Article https://doi.org/10.1038/s41562-023-01623-8

A meta-analysis of correction effects in science-relevant misinformation

Received: 9 May 2022

Accepted: 9 May 2023

Published online: 15 June 2023

Check for updates

Man-pui Sally Chan[1] ✉ & Dolores Albarracín[2]

Scientifically relevant misinformation, defined as false claims concerning a scientific measurement procedure or scientific evidence, regardless of the author's intent, is illustrated by the fiction that the coronavirus disease 2019 vaccine contained microchips to track citizens. Updating science-relevant misinformation after a correction can be challenging, and little is known about what theoretical factors can influence the correction. Here this meta-analysis examined 205 effect sizes (that is, k, obtained from 74 reports; $N = 60{,}861$), which showed that attempts to debunk science-relevant misinformation were, on average, not successful ($d = 0.19$, $P = 0.131$, 95% confidence interval −0.06 to 0.43). However, corrections were more successful when the initial science-relevant belief concerned negative topics and domains other than health. Corrections fared better when they were detailed, when recipients were likely familiar with both sides of the issue ahead of the study and when the issue was not politically polarized.

Unfounded misinformation about, for example, the false association between the coronavirus disease 2019 (COVID-19) pandemic and the rollout of 5G cellular tower networks[1,2] requires correction because it misleads citizens and can undermine their wellbeing[3–6]. Therefore, it is important to understand the efficacy of corrections of science-relevant misinformation using different methods, including research synthesis. In this Article, we define misinformation as 'false information' (see also refs. 7–10), including 'information considered incorrect based on the best available evidence from relevant experts at the time'[11], 'fabricated information that mimics news media content in form'[12], fictitious misinformation concerning a scientific measurement procedure or scientific evidence for example[13], and scientific findings that have been proven false[14,15]. Specifically, we examined science-relevant misinformation, defined as false claims concerning a scientific measurement procedure or scientific evidence, regardless of the author's intent, as opposed to, for example, political misinformation. We considered research on media campaigns to correct vaccine misconceptions[16,17], correction of false research reports about the impact of personality traits on the performance of firefighters[18], and correction of fictitious fake news coverage about scientific issues[16,19]. All in all, we analysed science-relevant misinformation, which we define as false claims attributed to scientific methods or scientists in areas such as social science, climate change or health. For example, the claim that the COVID-19 vaccine decreases fertility is pseudo-medical and thus comprises a science-relevant claim. This type of misinformation excludes non-scientific information about the same topics, such as a false claim that a politician made a particular statement about the COVID-19 vaccine or that another politician has refused to vaccinate.

Our meta-analysis was driven by theoretical explanations that suggest moderators of the impact of corrections as well as the initial misinformation. Specifically, we investigated how the nature of the misinformation, the correction and the recipient affect the correction. These factors, which are shown in Fig. 1, involve the valence of misinformation (that is, negative versus neutral) and the use of detailed (versus succinct) corrections. We also considered the attitudinal congeniality of the corrections (that is, congenial versus mixed/uncongenial) and issue polarization (that is, polarizing versus not polarizing) among the recipients.

The innovation of our meta-analysis was to synthesize the impact of science-relevant misinformation and its correction. Specifically, we are interested in two research questions. First, to what degree can the public update science-relevant misinformation after a correction? Second, what theoretical factors (that is, negative misinformation, detailed correction, attitudinal congeniality of the correction, and issue polarization) influence the impact of corrections? To address these questions, we synthesized reports of experiments studying the correction of science-relevant misinformation. We included corrections of

[1]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA, USA. [2]Annenberg School for Communication, Annenberg Public Policy Center, School of Arts and Sciences, School of Nursing, Wharton School, University of Pennsylvania, Philadelphia, PA, USA. ✉e-mail: sallycmp@upenn.edu

https://doi.org/10.1038/s41562-023-01623-8

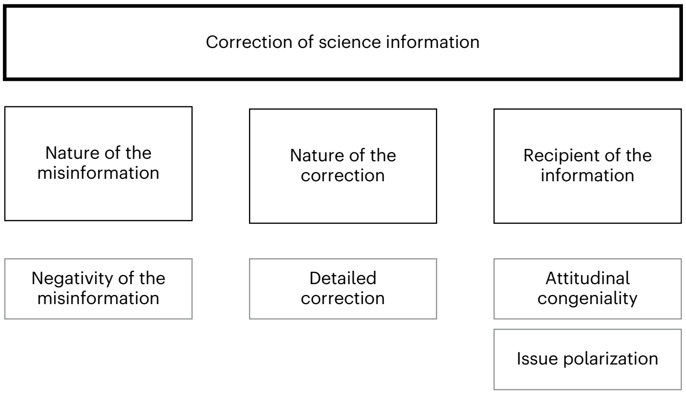

Fig. 1 | Theoretical factors related to the correction of science-relevant misinformation.

misconceptions that circulate in the real world such as misinformation about climate change, genetically modified organisms, COVID-19 and vaccines[20,21]. We also included corrections of fictitious misinformation concerning a scientific measurement procedure or scientific evidence (for example, ref. 13) and corrections of scientific findings that have been proven false (for example, refs. 14,15).

Six prior meta-analyses have examined the persistence of misinformation in news and reports[22–27]. However, none had the same goals as ours. Three of the prior meta-analyses did not concentrate on science-relevant information. Chan et al.[22] meta-analysed eight reports (*k* of effect sizes, 52) of research that used fictitious social and political news as experimental materials. Two other meta-analyses concentrated on social and political news such as restaurant rumours and news about political events[24,25]. The three that did consider correction of science-relevant misinformation[23,26,27] covered either climate change or health but not science-relevant information more generally. The first assessed five effect sizes concerning corrections of vaping misinformation[23], the second synthesized 15 effect sizes reflecting corrections of climate-change or health misinformation[26], and the third examined 24 effect sizes about the impact of a source's reliability ratings on correction of health misinformation on social media[27]. In sum, each of these meta-analyses considered a small number of effect sizes (that is, $k = 5$–24) without examining the moderators we examined or our general question (Fig. 1). These moderators are discussed in turn.

The misinformation

An important consideration about the impact of corrections is the valence of the misinformation, particularly whether the topic can arouse negative emotions in an audience. Given that much of the scientific information disseminated to the public is upsetting[28–31], we first wondered if negative science-relevant misinformation is easier or more difficult to correct than their neutral counterpart. For example, the alleged side effects of infertility or autoimmune diseases following vaccination against human papilloma virus (HPV)[32,33] can elicit fear or sadness in an audience[33–37]. These negative emotional implications may affect corrections of this information, although the direction of influence is debatable. On the one hand, negative information may elicit more attention and more thorough processing[38,39], which may, in turn, increase the persistence of negative (versus neutral) misinformation. On the other hand, people are more likely to hold beliefs that make them feel good about themselves, their future or the world more generally[40–42]. As a result, negative (versus neutral) misinformation may be easier to correct because doing so may improve a person's mood.

The correction

Different correction factors may also affect the correction of science-relevant misinformation. According to the notion of mental models, people's ability to discard a model built on misinformation depends on the strength of that model and the ability of the correction to promote a new model[43,44]. Accordingly, detailed corrections and causal explanations can change prior models and make correcting misinformation (see also refs. 45,46). Yet, detailed corrections and explanations may have the ironic effect of strengthening misinformation persistence[22]. For example, detailed corrections and elaborate explanations of the misinformation may remind the audience of the misinformation[47,48]. As no prior meta-analysis has assessed the influence of detailed (versus succinct) corrections of science-relevant misinformation, we attempted to fill this void by analysing this moderator.

The recipient

Whether a correction is congenial to the recipient's attitudes or beliefs may also affect the success of corrections of science-relevant information[49]. A 2020 meta-analytic review of political misinformation[25] found that corrections were more efficacious when they were congruent with recipients' attitudes than when they were not[50–52]. However, the evidence about the impact of congeniality is not monolithic. For example, receiving misleading headlines congenial to recipients' political ideology does not impair recipients' ability to distinguish true from false information[7,53] or their motivation to share only accurate information with their social networks[10]. Furthermore, we considered if the topic is politically polarized. If people engage in motivated reasoning to protect their political identity[40,54–56], for example, corrections should have lesser effect when the issues are politically polarized. However, corrections have been shown to work for polarized misinformation as well[57]. Our meta-analysis thus estimated the impact of both congeniality and political polarization on corrections of science-relevant information.

The present study

We synthesized a large body of experimental evidence comprising 60 published experiments, 5 working papers, 2 theses and 7 unpublished datasets (total number of reports, 74; *k* of effect sizes, 205). We examined the moderating effects of factors related to the misinformation (that is, negativity), the correction (that is, detailed correction), the recipient (that is, attitudinal congeniality and issue polarization), control factors (that is, domain, fictitious issue, likely familiarity and in-person correction) and report and methodological characteristics (that is, study sample, lab context and method of effect-size calculation). We also assessed the misinformation effect to ensure that our moderators reflected the impact of the correction rather than the impact of the misinformation. All procedures are detailed in Methods.

In addition to the theoretical moderators included in Fig. 1, we controlled for factors that could vary across experiments. We controlled for the domain of the misinformation (that is, political, health, environment and others), whether the misinformation was fictitious[58], whether the audience was likely familiar with the topic and whether the correction was delivered in person[59,60]. We also considered two report characteristics and one methodological characteristic that might affect corrections, including study sample (that is, the United States versus other countries), lab context (that is, laboratory versus online) and method of effect-size calculation (that is, between subjects versus within subjects). We chose these control characteristics on the basis of a review of prior meta-analyses[22,26] and the availability of these details in our sample of reports.

Analytic procedures

We estimated Hedges' d for the effect of the correction with adjustments to minimize small sample bias and on the basis of either between-subjects or within-subjects variances depending on the procedures of the included studies. We subtracted the mean belief or attitude rating after a correction was introduced from the mean rating before a correction or in a control condition. We used four add-on packages for the statistical software R version 4.0.5: robumeta version

Table 1 | Descriptive statistics of the included reports

	Mean (s.d.)	*k*
Sample size	174.1 (225.02)	205
Percentage of females	59.09 (11.1)	116
Percentage of males	40.24 (11.64)	105
Age	36.1 (11.26)	86
Country		
United States (%)	69%	141
Other countries (%)	31%	62

2.0 (refs. 61,62), metafor version 3.9.9 (ref. 63), puniform version 0.2.5 (ref. 64) and weightr version 2.0.2 (ref. 65) to assess publication/inclusion bias and analyse the mean effect sizes using robust variance estimation (RVE) methods[66–68]. We also used JASP version 0.16.3 (ref. 69), an open-source statistics program, to conduct bias analyses with Bayesian methods. We calculated the I^2 statistic (that is, the percentage of total variation across experimental conditions due to random heterogeneity), which controls for k and indicates the percentage of total variation across experimental conditions that is due to true heterogeneity rather than sampling error[70,71]. Furthermore, we performed meta-regression analyses of our debunking effect sizes with the moderators in Fig. 1 introduced as predictors and repeated this analysis with the misinformation effect as an outcome and later as a covariate. Because debunking effects involved both correction and the reverse of misinformation persistence, these analyses utilized RVE to account for the statistical dependence between the two, which was estimated at 0.53 (that is, ρ). A separate sensitivity analysis, as recommended by Hedges et al.[68], was performed to confirm that the selected ρ estimate was appropriate ($\tau^2 = 0.40$ for ρ ranging from 0 to 1). The pre-registration materials and data and code repositories are available at https://osf.io/vkygw/ (the pre-registration was done before the last round of database update in mid-2022).

Results

We included 74 research reports and 205 independent effect sizes (for the included data, see Supplementary Table 1). According to a review by two authors (Methods), these conditions met inclusion criteria in that they (1) provided a false claim concerning a scientific measurement procedure or scientific evidence, (2) had measures of participants' beliefs or attitudes consistent with the misinformation addressed by the correction and (3) had a baseline or control group. Also, studies were eligible when (4) the misinformation was initially asserted to be true or was known to participants before the study and was later corrected (Methods). Table 1 reports descriptive statistics of the included reports. The number of participants in the synthesis ranged from 7 to 1,180, their mean age was 36 years (standard deviation (s.d.) 10.73) and about 60% of them were female. Participants included university students, graduate students and adults from the community recruited via online survey platforms, such as Amazon Mechanical Turk and Prolific.

Overall effects, heterogeneity and bias

Although the misinformation effect was large ($d = 1.10$, $P < 0.001$, 95% confidence interval (CI) 0.72 to 1.48), the debunking effect was not significant ($d = 0.19$, $P = 0.131$, 95% CI −0.06 to 0.43). The heterogeneity analyses showed high I^2 statistics (misinformation 99.17% and debunking 98.43%), suggesting systematic variability across conditions in addition to sampling error or multiple populations of effects[71]. We also conducted extensive analyses of inclusion bias, which appear in Table 2 (see details of the bias analyses in Methods). These analyses revealed no consistent evidence of bias in the dataset.

Table 2 | Summary of bias analyses

Bias analysis	Result with the outliers	Result without the outliers	Indication of bias
Contoured-enhanced funnel plot with the trim-and-fill method[118–120]	LO: 0 estimated records filled on the right RO: 0 estimated records filled on the right	LO: 0 estimated records filled on the right RO: 0 estimated records filled on the right	LO: no RO: no
Rank correlation test	LO: $b=0.22$, SE = 0.10, $P=0.021$ RO: $b=0.22$, SE = 0.10, $P=0.021$	LO: $b=0.21$, SE = 0.08, $P=0.010$ RO: $b=0.21$, SE = 0.08, $P=0.010$	LO: yes RO: yes
Precision-effect test—precision-effect estimate with standard error[121]	Small difference between the PEESE and the RVE estimate, $d_{diff}=0.06$	Small difference between the PEESE and the RVE estimate, $d_{diff}=0.06$	No
P-uniform test with the default 'P' method set[122]	L.pb = −11.35, $P=1.000$	L.pb = −11.10, $P=1.000$	No
Three-parameter selection model[123]	$b=0.12$, SE = 0.16, $P=0.459$	$b=0.09$, SE = 0.13, $P=0.487$	No
Meta-regression test of publication type	Working paper: $b=0.26$, SE = 0.23, $P=0.259$ Dissertation/thesis: $b=-0.02$, SE = 0.15, $P=0.901$ Unpublished data: $b=0.22$, SE = 0.18, $P=0.211$	Working paper: $b=0.25$, SE = 0.22, $P=0.263$ Dissertation/thesis: $b=-0.03$, SE = 0.14, $P=0.851$ Unpublished data: $b=0.21$, SE = 0.16, $P=0.185$	No
Weight-function models[65]	Log-likelihood ratio was significant $\chi^2(df = 3)=16.38$, $P=0.001$	Log-likelihood ratio was significant $\chi^2(df = 3)=12.97$, $P=0.005$	Yes
RoBMA[124]	$BF_{10} = 9.668$	$BF_{10} = 9.863$	Yes

b indicates unstandardized coefficients, SE indicates standard error, BF_{10} indicates Bayes factor giving the evidence for H1 over H0, and P values ≥ 0.05 were concluded as no indication of the presence of bias.

Moderator analyses

Table 3 presents the results from meta-regressions, and Table 4 reports the predicted estimated mean effect sizes for each level of the categorical moderators using the identified meta-regression model. First, corrections were more effective for negative (versus neutral) misinformation ($b(60) = 0.80$, $P = 0.011$, 95% CI 0.19 to 1.40). Next, corrections were more successful when they were detailed (versus succinct) ($b(60) = 0.64$, $P = 0.020$, 95% CI 0.10 to 1.18) and concerned a non-polarizing (versus polarizing) issue ($b(60) = 0.76$, $P = 0.046$, 95% CI 0.02 to 1.51). Of note, corrections were successful irrespective of whether they were congenial (versus mixed/uncongenial) ($b(60) = 1.68$, $P = 0.066$, 95% CI −0.12 to 3.47). As for control factors, corrections were more efficacious when the misinformation concerned other (versus health) topics ($b(60) = 1.34$, $P = 0.003$, 95% CI 0.47 to 2.22) and when recipients were likely familiar (versus unfamiliar) with the topic ($b(60) = 1.20$, $P = 0.003$, 95% CI 0.43 to 1.98). A marginal positive effect on correction was found when the issues were fictitious (versus real) ($b(60) = 0.73$, $P = 0.053$, 95% CI −0.01 to 1.47). There was no statistically significant difference in in-person (versus through media) corrections ($b(60) = 0.70$, $P = 0.106$, 95% CI −0.15 to 1.55). As shown in Table 3, these meta-regressions also controlled for whether the samples were from the United States (versus other countries), whether the studies were conducted in the lab (versus online) and whether the effect sizes were computed using the between- or within-subjects methods. Of these variables, only lab (versus online) was a significant factor.

Table 3 | Meta-regression results

Variable	Debunking effect			Misinformation effect			Debunking effect controlling for the misinformation effect		
	k=203			*k*=29			*k*=54[a]		
	b (SE)	*P*	95% CI	*b* (SE)	*P*	95% CI	*b* (SE)	*P*	95% CI
Intercept	−1.1 (0.45)	0.017	−1.99 to −0.21	0.43 (0.80)	0.591	−1.14 to 2	−0.05 (0.55)	0.925	−1.35 to 1.25
Nature of the misinformation									
Negative misinformation^	0.8 (0.3)	0.011	0.19 to 1.4	0.42 (0.58)	0.468	−0.71 to 1.55	0.47 (0.53)	0.412	−0.8 to 1.73
Nature of the correction									
Detailed correction^	0.64 (0.27)	0.02	0.1 to 1.18	0.00 (0.59)	0.997	−1.15 to 1.14	−0.15 (0.25)	0.568	−0.73 to 0.44
Recipients of the misinformation									
Attitudinal congeniality of the correction^	1.68 (0.9)	0.066	−0.12 to 3.47	–	–	–	–	–	–
Issue polarization^	−0.76 (0.37)	0.046	−1.51 to −0.02	−0.35 (0.95)	0.713	−2.21 to 1.51	−0.6 (0.53)	0.295	−1.85 to 0.66
Control factor:									
Domain of misinformation: political^	0.57 (0.76)	0.459	−0.95 to 2.09	–	–	–	–	–	–
Domain of misinformation: health^	−1.34 (0.44)	0.003	−2.22 to −0.47	–	–	–	–	–	–
Domain of misinformation: environment^	−0.46 (0.53)	0.395	−1.52 to 0.61	–	–	–	–	–	–
Fictitious issue^	0.73 (0.37)	0.053	−0.01 to 1.47	0.78 (0.82)	0.342	−0.83 to 2.39	−0.14 (0.65)	0.834	−1.69 to 1.4
Likely familiarity with the topic^	1.2 (0.39)	0.003	0.42 to 1.98	−1.30 (0.96)	0.175	−3.18 to 0.58	0.55 (0.74)	0.480	−1.2 to 2.3
In-person correction^	0.7 (0.43)	0.106	−0.15 to 1.55	0.26 (0.4)	0.522	−0.53 to 1.05	1.14 (0.45)	0.037	0.09 to 2.2
Report and methodological characteristics									
Study sample^	−0.14 (0.27)	0.604	−0.67 to 0.39	1.34 (0.74)	0.073	−0.12 to 2.79	1.82 (0.65)	0.027	0.28 to 3.35
Lab context^	0.89 (0.39)	0.026	0.11 to 1.68	–	–	–	–	–	–
Method of the effect-size calculation^	−0.32 (0.24)	0.185	−0.8 to 0.16	–	–	–	–	–	–
Misinformation effects	–	–	–	–	–	–	−0.13 (0.37)	0.728	−1 to 0.73

A caret indicates a categorical variable. Unstandardized coefficients and standard errors in parentheses. RVE was used for the debunking effects, and mixed-effect estimation was used for the misinformation effects. [a]*k*=54 because the same misinformation effect was assigned to the correction effect and the misinformation-persistence effect to the RVE estimation.

An important consideration, however, is whether the effects of negative misinformation, a detailed correction, issue polarization, different domains, fictitious issues and likely familiarity with the topic might impact the misinformation impact rather than the correction per se. Therefore, we next regressed the misinformation effect of the same moderators, an analysis that appears in the second column of Table 3. Additionally, we included the misinformation effect as a covariate and conducted a meta-regression analysis of the debunking effects with the sample available for it, which was only 54 effect sizes (see the third column of Table 3) as well as with imputation for missing misinformation effects (Supplementary Table 5).

Discussion

This meta-analysis assessed two critical questions: To what degree can the public update science-relevant misinformation after a correction? And, what theoretical factors (that is, negative misinformation, detailed correction, attitudinal congeniality of the correction and issue polarization) influence the impact of corrections? We showed that science-relevant misinformation is particularly challenging to eliminate. In fact, the correction effect we identified in this meta-analysis (d = 0.19, P = 0.131, 95% CI −0.06 to 0.43) is smaller than those identified in all other areas (for example, d = 1.14–1.33, 95% CI 0.62 to 2.04 in Chan, Jones, Jamieson & Albarracin[22] and d = 0.40–0.75 in Walter et al.[24] and Walter & Murphy[25,26]). We also identified conditions under which corrections are most effective, including detailed corrections, negative misinformation and issue polarization. Of these, only detailed corrections[22] had been examined in prior meta-analyses.

Our meta-analysis can provide insights into developing evidence-based interventions for science-relevant misinformation. Although there is a growing interest in the development of effective interventions to curb the impact of misinformation[72], the majority of the proposed mechanisms have focused on either the impact of the misinformation or the cognitive processing of corrections. In this context, our work suggests that the correction effects are a joint function of multiple factors concerning the misinformation, the correction and the recipient of the information (Fig. 2). The findings thus provide an integrated model that can better explain the complexity of the processes at hand.

Our findings about attitudinal congeniality and issue polarization are also important in the context of discussions about different reasoning accounts of misinformation[10,50,52,73]. The marginal positive effect of attitudinal congeniality of correction seems to be in line with recent experimental data that false headlines from media sources congenial to recipients' political ideology are perceived to be more accurate[74]. However, as Traberg and van der Linden's work concerned misinformation, future should investigate if political congeniality improves the impact of corrections. Our results concerning issue polarization provide some support for the motivated reasoning[40,56] and identity-based accounts[75,76] of correcting misinformation. Corrections become less efficacious when the issue is polarizing, possibly because

 https://doi.org/10.1038/s41562-023-01623-8

Table 4 | Predicted estimated mean effect sizes for levels of categorical moderators

Variable	df	*d* (SE)	*P*	95% CI
The misinformation				
Negativity of the misinformation				
Negative	18.83	0.74 (0.27)	0.012	0.19 to 1.3
Neutral	31.7	−0.05 (−0.13)	0.688	−0.32 to 0.21
The correction				
Detailed correction				
Detailed	28.29	0.65 (0.2)	0.003	0.23 to 1.06
Succinct	30.81	0.01 (0.16)	0.968	−0.33 to 0.34
The recipient				
Attitudinal congeniality of the correction				
Congenial	2.01	1.85 (1.03)	0.215	−2.57 to 6.26
Mixed/uncongenial	30.48	0.17 (0.11)	0.130	−0.05 to 0.39
Polarizing issue				
Yes	24.09	−0.23 (−0.26)	0.394	−0.77 to 0.31
No	14.74	0.53 (0.18)	0.011	0.15 to 0.92
Control factor				
Domain of the misinformation				
Political	2.71	0.75 (0.88)	0.482	−2.22 to 3.73
Health	16.62	−0.65 (−0.34)	0.071	−1.36 to 0.06
Environment	24.79	−0.18 (−0.5)	0.716	−1.22 to 0.85
Others	17.54	0.71 (0.21)	0.003	0.27 to 1.15
Fictitious issue				
Fictitious	33.21	0.44 (0.14)	0.004	0.15 to 0.72
Real	20.49	−0.29 (−0.31)	0.363	−0.95 to 0.36
Likely familiarity with the topic				
Yes	9.67	0.75 (0.19)	0.003	0.33 to 1.17
No	16.05	−0.45 (−0.28)	0.128	−1.04 to 0.14
In-person correction				
Yes	7.48	0.81 (0.41)	0.090	−0.15 to 1.78
No	44.92	0.11 (0.12)	0.363	−0.14 to 0.36
Report and methodological characteristics				
Study sample				
United States	23.79	0.12 (0.23)	0.592	−0.35 to 0.59
Rest of the world	34.7	0.26 (0.14)	0.070	−0.02 to 0.54
Lab context				
Lab	21.4	0.64 (0.21)	0.007	0.2 to 1.08
Online	21.69	−0.26 (−0.26)	0.338	−0.8 to 0.29
Method of effect-size calculation				
Between subjects	33.88	0.03 (0.15)	0.860	−0.28 to 0.34
Within subjects	30.95	0.35 (0.18)	0.058	−0.01 to 0.71

df indicates degrees of freedom with the small sample correction, *d* indicates the Cohen's predicted *d* of the meta-regression analysis and SE indicates standard error. The predicted *d* were estimated while keeping the covariates of negative misinformation (0.34), detailed correction (0.33), attitudinal congeniality of the correction (0.03), issue polarization (0.41), political domain of misinformation (0.06), health domain of misinformation (0.35), environmental domain of misinformation (0.12), fictitious issues (0.70), likely familiarity with the topic (0.56), in-person correction (0.15), study sample (0.31), lab context (0.53) and method of effect-size calculation (0.40) at their grand means.

recipients defend themselves against identity threats and counterargue the correction. However, these accounts are difficult to separate from the fact that issues one agree with simply appear more valid and need to be validated with a demonstration of the impact of goals.

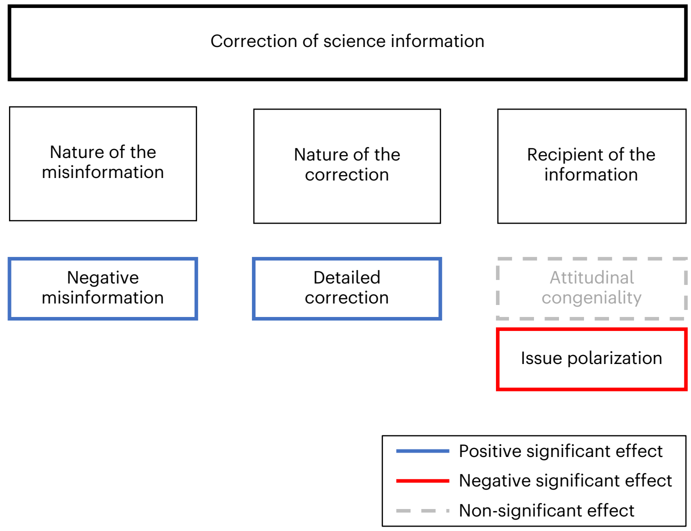

Fig. 2 | Findings of theoretical factors related to the correction of science-relevant misinformation.

Our results suggest practical recommendations for undercutting the influence of science-relevant misinformation. First, to maximize efficacy, corrections should provide detailed arguments rather than simple denials. Second, corrections should be accompanied with methods to reduce polarization around an issue. For example, thinking of a friend with a different political ideology can reduce affective polarization[77,78]. Lastly, corrections are likely to be more effective when recipients are familiar with the topic. Therefore, increasing public exposure to the topics (for example, general information about a subject matter) may also maximize the impact of debunking.

Even though this meta-analysis, to our knowledge, is the most comprehensive in this area, our conclusions have limitations imposed by the existing literature. First, because little experimental work has assessed the impact of repeating the misinformation or the corrections[79–81], future research should address that problem. Second, no experiment measured or estimated people's understanding of the scientific process, which is one of the key factors in science communication[82,83]. Although the level of education attained by participants can serve as a proxy for this knowledge, only one experiment reported the education level attained for each condition in their results[84]. Note that a separate meta-regression ($k = 107$) with the mean education of a study as a whole included as an additional moderator revealed no statistically significant association with the debunking effect ($P = 0.995$). Future research should report educational attainment, ideally for each experimental condition, to assess whether education moderates misinformation and correction effects. Third, the I^2 statistics still showed a high proportion of random heterogeneity (that is, between-studies variability) even after controlling for our moderators. Other factors that contribute to this unexplained heterogeneity may include variability in the social environment, conditions of study administration and experimental paradigms, which may not be discernible from published results but may nonetheless affect study results. Finally, researchers should pre-register their experiments to increase the transparency of their methodologies and improve reproducibility. Direct replications using shared experimental paradigms may overcome the limitations of single experiments and control for the differences in the studies included in a meta-analysis[85]. Taken together, the meta-analytic and replication efforts should provide complementary evidence about how to best protect populations from the dangers of pseudoscientific misinformation.

https://doi.org/10.1038/s41562-023-01623-8

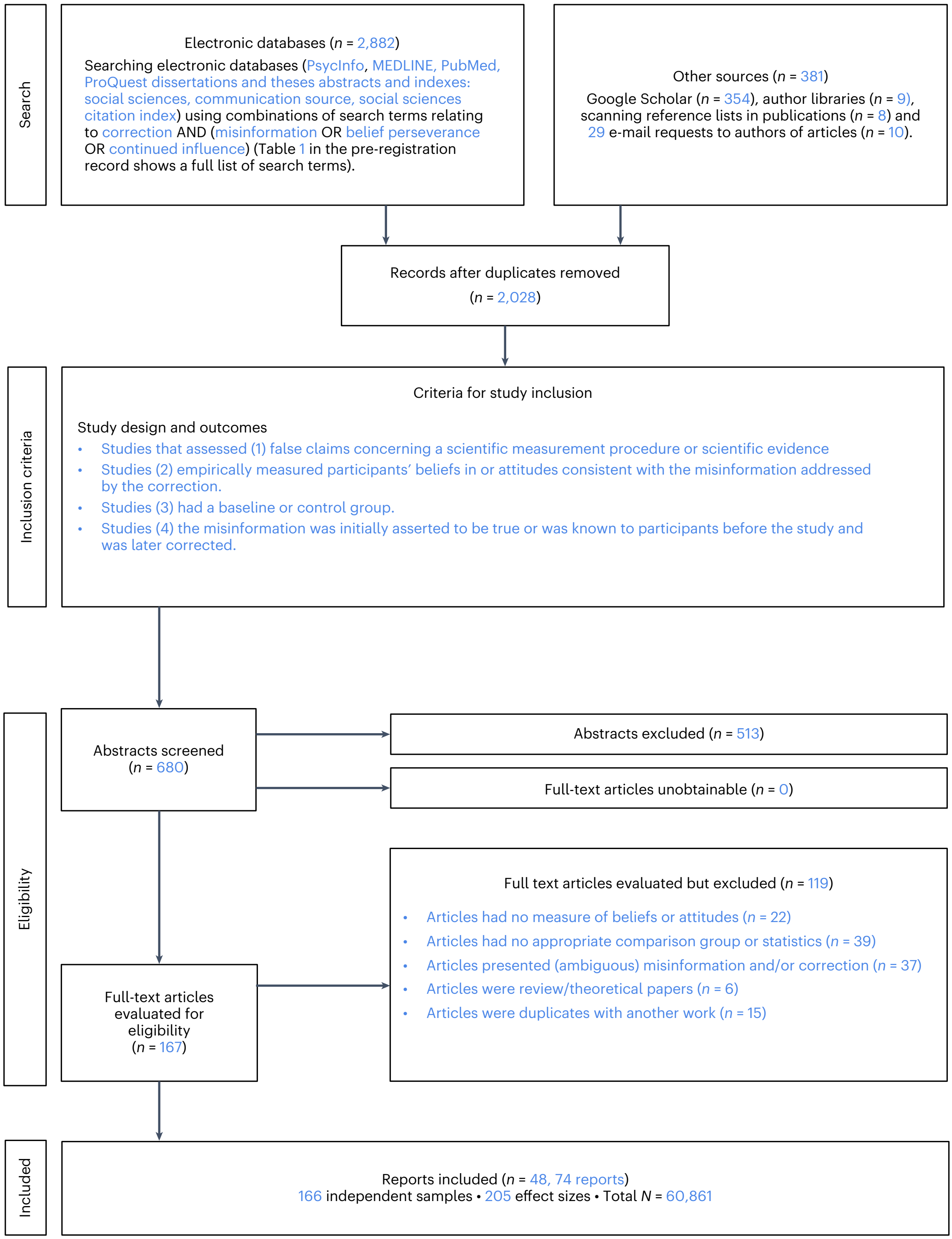

Fig. 3 | Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

Methods

Literature search

We used several search methods to ensure a thorough examination of potential candidate reports. The number of records identified, included and excluded and the reasons for exclusions are shown in Fig. 3. The literature search covered a timeframe up to August 2022.

Multiple-database searches. To obtain relevant articles, we used specific keywords with wildcards, performing a combined search of the following seven online databases: (1) PsycInfo, (2) Google Scholar, (3) MEDLINE, (4) PubMed, (5) ProQuest Dissertations and Theses Abstracts and Indexes: Social Sciences, (6) Communication Source and (7) Social Sciences Citation Index. We paired a series of keywords (that is, misinformation OR misbelief* OR false information OR belief perseverance OR continued influence) with two other series of keywords (that is, [retract* OR correct*] and [label* OR tag* OR flag*]). We expanded the sample of relevant reports by examining the reference lists of a systematic collection of review articles, book chapters and dissertations. The search yielded 2,882 studies.

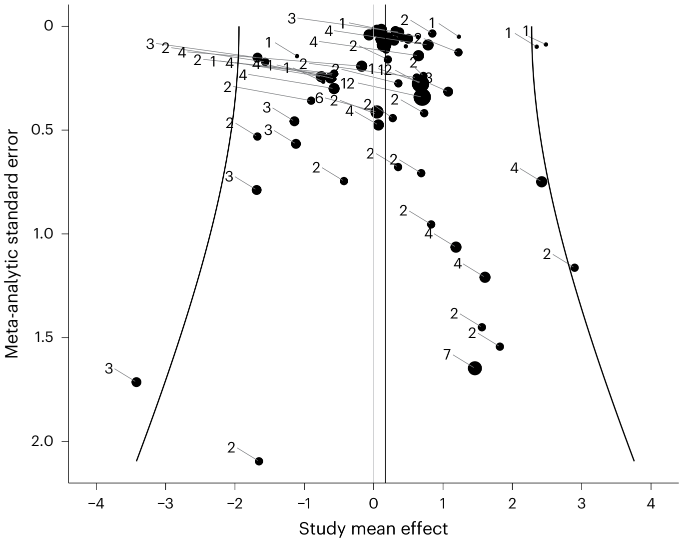

Fig. 4 | Study-funnel plot. Each dot represents a report in an article, and the numbers represent the number of effect sizes included in the estimation.

Other searches, personal contact and electronic platforms. By culling the reference lists of the review papers obtained through the database searches, we were able to identify eight additional articles. We also identified ten studies after contacting a list of researchers who have researched in this area. Additionally, we received materials from additional researchers after posting requests on online forums and e-mail list servers (for example, Society for Personality and Social Psychology). Finally, we searched the Open Science Framework (OSF) using the same set of keywords in early 2021 to obtain 258 unpublished and in-press datasets.

Criteria for inclusion/selection

We included (1) studies that assessed false claims concerning a scientific measurement procedure or scientific evidence (for examples of excluded reports due to measuring non-science-relevant misinformation, see Supplementary Table 2). For example, we included Vraga et al.'s[86] experiments studying logic-based and humour-based corrections for misinformation about climate change and the HPV vaccine. Another example of the included reports was Vijaykumar et al.'s[87] experiments that examined corrections for the inaccurate treatment effect of garlic to cure COVID-19. We also included Anderson et al.'s[88] experiments ostensibly evaluating the relation between firefighter performance and risk-seeking traits. Likewise, we included Greitemeyer's[14] experiments studying the impact of a false link about the relation between the embodiment of height and pro-social behaviour.

Next, we used several eligibility criteria to select reports for inclusion in the meta-analysis, inspecting only studies from reports that were clearly or possibly experimental. Included studies (2) empirically measured participants' beliefs in or attitudes consistent with the misinformation addressed by the correction (for examples of excluded reports after evaluating the full text, see Supplementary Table 3). However, studies that included outcome measures of misinformation sharing (intentions)[10,89–92], the quality judgement of news sources[93], self-efficacy, openness to the messages[94] and whether participants responded to/ignored the corrections were excluded[95]. The included studies (3) had a control or baseline group exposed to either no message, a neutral message or an unrelated message and (4) were eligible when the misinformation was initially asserted to be true or was known to participants before the study and was later corrected. All studies introduced a correction for the misinformation regardless of whether the misinformation was fictitious (for example, ref. 88) and known to be familiar to the public (for example, refs. 16,95). However, studies that described the initial information as hypothetical or uncertain (for example, ref. 96) or as an accusation of scientific misconduct (for example, ref. 97) were excluded. The final dataset of this meta-analysis included 74 research reports (N = 60,861). Reports with multiple experiments and/or experimental groups often contributed more than one effect size.

Estimation of the effect sizes

We used Hedges' d as the metric for effect-size estimation in our meta-analysis. Hedges' d and Hedges' g pool variances are based on the assumption of equal population variances, and both metrics can be interpreted in the same way[98]. However, Hedges' d includes an adjustment factor j, that is, $1 - 3/(4 \times n - 1)$, for each sample, and in turn, reduces the positive bias for a small sample that is common in experimental studies[99,100]. Two trained raters first decided whether the report was in a between- or within-subjects design and then selected corresponding means and s.d. from different groups or conditions to compute Hedges' d in accordance with the formulas outlined by Borenstein et al.[100]. As only two reports used a mixed-subjects design and provided sufficient data (that is, means and s.d. of the control groups at timepoints 1 and 2), the raters followed a within-subject effect-size equation to calculate the effect sizes. All d statistics are in the same normalized units regardless of whether they derive from between- or within-subjects designs[100]. If a particular study did not report any of these statistics, the rater recorded other relevant statistics, such as F ratios or t values, and then obtained Hedges' d on the basis of the step-by-step workflow as specified in Lakens's effect-size calculation spreadsheet[101]. Given the inclusion of different experimental designs in this meta-analysis, the d obtained from the reports compared different means as explained presently. We obtained effect sizes for misinformation and debunking. Debunking effects combined correction and the reverse of misinformation-persistence effects. We followed different procedures to calculate the variances of effect sizes. In particular, calculations of the between-subjects variances followed Hedges and Olkin's[98] procedures, and calculations of the within-subjects variances followed Morris's[102] procedures with a correlation set at 0.5 between repeated measures.

Within-subjects design. We first illustrate the effects of correction and misinformation persistence using a within-subjects design. Imagine that participants were recruited for an experiment with a pre-test–post-test design and that they provided a rating on a belief or attitude measure from 0 to 9 to indicate their belief or attitude both before (pre-test) and after the experimental manipulations of misinformation (post-test 1) and correction (post-test 2). The comparisons among the ratings for the pre-test and post-test 1 generate a misinformation effect. The comparisons among the ratings for post-tests 1 and 2 generate a correction effect, whereas the comparisons between the ratings for the pre-test and post-test 2 allow us to calculate a misinformation-persistence effect size. For example, imagine that participants gave a rating of 1 at the pre-test, a rating of 9 after the receipt of the misinformation (post-test 1) and then a rating of 6 after correcting the misinformation (post-test 2). Then, the misinformation effect is the difference between the ratings at post-test 1 and at the pre-test (that is, $9 - 1 = 8$); the correction effect is the difference between the ratings at post-tests 1 and 2 (that is, $9 - 6 = 3$); and the misinformation-persistence effect is the difference between the ratings at post-test 2 and the pre-test (that is, $6 - 1 = 5$). When pre-test ratings were unavailable, we used the ratings obtained from a control group as the baseline for comparisons with the ratings at post-tests 1 and 2 using between-subjects procedures.

Between-subjects design. Imagine now a between-subjects design with three groups of participants. Imagine also that participants in the misinformation group received only misinformation, participants in the correction group received the misinformation and subsequently a correction, and participants in the control group received either no information or information on an unrelated topic. Now consider that all participants provided a rating on a belief or attitude measure from 0 to 9. Participants in the misinformation group gave an average rating of 9, participants in the control group gave an average rating of 1, and participants in the correction group gave an average rating of 6. In these circumstances, the differences between the misinformation group and the control group constitute the misinformation effect (that is, $9 - 1 = 8$); the differences between the misinformation-only group and the correction group constitute the correction effect (that is, $9 - 6 = 3$); and the differences between the correction group and the control group constitute the misinformation-persistence effect (that is, $6 - 1 = 5$). Accordingly, d values greater than 0 for later persistence of the misinformation indicate that recipients of the corrections showed more misinformation persistence than participants in the comparison group (for example, the control group).

Coding of moderators

Two authors and a trained research assistant coded for four theoretical moderators, including the variables of interest: (1) whether the misinformation was about negative or neutral topics, (2) the level of detail of the correction messages, (3) the attitudinal congeniality of the correction and (4) the issue polarization and four other control factors, including (5) whether the misinformation was about politics, health, environment and others, (6) whether the misinformation was about fictitious or real issues, (7) the likelihood of familiarity with the topic, and (8) the use of in-person correction[9,46,103–119]. Two rounds of coding of all variables contained about 14% of the reports, and the coding reached an adequate agreement (Krippendorff's α: mean 0.99, s.d. 0.01 and Cohen's κ: mean 0.95, s.d. 0.09). Further, the coders resolved all disagreements by discussion and consultation with another author. Table 4 summarizes effect sizes for each level of all categorical moderators and the number of experimental conditions coded for each level. We next detailed definitions and examples of all moderators.

Misinformation factors. Negativity of the misinformation. We coded whether the misinformation topic was negative or neutral. As a first example, Guenther and Alicke's[110] misinformation about failure feedback on an alleged test of mental acuity to measure a fundamental aspect of intelligence was coded as negative (that is, a score of 2) because of the potential to induce sadness or anxiety. As another example, Anderson's[18] experiments included misinformation about whether risk-seeking or risk-averse firefighter trainees performed better at their job. This misinformation topic was coded as neutral, receiving a score of 1. Only three reports from two studies contained positive misinformation topics. In two, participants received flattering feedback on their cognitive ability based on a task-performance task[111] and on a word-identification task supposedly linked to intelligence[110]. We thus analysed the data excluding these three reports.

Correction factors. Level of detail of the correction messages. The two raters also coded whether the correction simply labelled the initial information as incorrect (1, succinct) or provided detailed information (2, detailed). For example, a detailed correction message (that is, the author realized that some of the facts in the reading were not true) describing why the initial misinformation (that is, the facts of the reading were 'mixed up' with facts of a fictional story also to-be-published) was incorrect was considered as a detailed correction message (a score of 2) (ref. 112).

Recipient factors. Attitudinal congeniality of the correction. This variable captured whether participants had any pre-existing attitudes relative to the position advocated in the correction message. For example, Ecker and Ang's[113] experiment 1 disseminated different corrections to participants who reported being left-wing. Here the conditions with consistent partisan information (for example, Labour supporters receiving left-wing correction) were coded as 1. The conditions with either inconsistent partisan information (for example, Liberal supporters receiving left-wing correction) or non-partisan information were coded as −1.

Issue polarization. We coded whether the topic was associated with disagreement between opposing groups in the country where the experiment was carried out (polarizing, 1; non-polarizing, −1).

Control factors. Domains of the misinformation. We coded whether the misinformation was about politics, health, environment and others (politics, 1; health, 2; environment, 3; others, −1) on the basis of the misinformation included in the reports, regardless of whether it was politicized in the real world. For example, the alleged measles, mumps and rubella vaccines–autism link was coded as concerning health (2) and misconceptions about climate change were coded as concerning the environment (3). Reports with different misinformation (and possibly different domains) were coded as separate records whenever possible, for example[86]. As only four reports ($k = 10$) had misinformation about multiple domains, we decided not to include all possible combinations of domains as separate coding options.

Fictitious issue. We coded whether the claim was fictitious (1) or real (−1). For example, the alleged link between Zika virus vaccines and epilepsy was never true, receiving no scientific support (1), whereas there was scientific support (even minimal) for hydroxychloroquine to be effective against COVID-19 (−1).

Likely familiarity with the topic. We coded for whether the topic used in the experiment had circulated in the real world. For example, Ecker[108] presented vaccine misinformation concerning the link between the measles, mumps and rubella vaccine and autism to UK participants, a topic of wide dissemination in the United Kingdom. This study was coded as 1. In contrast, Sherman and Kim's[114] experiments used the topic of the associations between Chinese characters and English meanings, which was coded as −1 (that is, likely unfamiliar topic).

In-person correction. We next coded whether the correction was given in person or not. As an example of in-person delivery, Golding et al.'s[115] correction involved an experimenter telling research participants in the lab to disregard initial misinformation, and it was coded as in person (that is, a score of 1). In contrast, Sherman and Kim's[114] experiments presented the experimental materials using computer software and were coded as not in person (that is, a score of −1).

Report and methodological characteristics. Geographical location of the study sample. We coded whether participants were self-reported as from the United States or other countries (United States, 1; other countries, 2).

Lab context. We coded whether the experiment was carried out in the lab (a score of 1) or online (a score of 2).

Methods of effect-size calculation. We recorded whether the effect size of the report stemmed from a between-subjects (a score of 1) or within-subjects (a score of 2) design.

Source type. We coded the report's publication category (published article, 1; working paper, 5; dissertation/thesis, 6; unpublished data, 7).

Bias analysis

Examining variability and bias is critical in meta-analysis because much research is affected by both high variability and bias. We adopted the diagnostic procedures proposed by Viechtbauer and Cheung[116] to detect influential cases, and four debunking effects were identified as outliers, $d < -7.00$ or $d > 4.67$. Because outliers and influential cases may represent random noise or reflect systematic heterogeneity as a function of specific moderators, we performed six bias tests to assess publication/inclusion biases for effect sizes with and without the outliers. Figure 4 shows the study-level funnel plot (for the study-level forest plot, see Supplementary Information).

We performed bias tests to assess publication/inclusion biases for effect sizes with and without outliers[65,117–124]. Overall, the bias tests showed no consistent results regarding the presence of any bias in the dataset (Table 2). Table 2 presents a consistent pattern of the results of bias analyses between the datasets with and without the outliers. The rank correlation test, the log-likelihood ratio tests method[65], and the robust Bayesian meta-analysis (RoBMA)[124] showed the possibility of bias in the dataset ($P < 0.05$ and $BF_{10} > 1$). In contrast, the trim-and-fill method, the meta-regression analyses of publication type, the PET-PEESE (precision-effect test and precision-effect estimate with standard errors) test, the three-parameter selection method[124] and the *P*-uniform test showed no evidence of the presence of bias.

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
The data that support the findings of this study are openly available in OSF at https://osf.io/vkygw/.

Code availability
All code for data analyses associated with the current submission is available at https://osf.io/vkygw/. Any updates will also be published in OSF.

References

1. Ahmed, W., Downing, J., Tuters, M. & Knight, P. Four experts investigate how the 5G coronavirus conspiracy theory began. *The Conversation* https://theconversation.com/four-experts-investigate-how-the-5g-coronavirus-conspiracy-theory-began-139137 (2020).
2. Heilweil, R. The conspiracy theory about 5G causing coronavirus, explained. *Vox* (2020); https://www.vox.com/recode/2020/4/24/21231085/coronavirus-5g-conspiracy-theory-covid-facebook-youtube
3. Pigliucci, M. & Boudry, M. The dangers of pseudoscience. *The New York Times* (2013); https://opinionator.blogs.nytimes.com/2013/10/10/the-dangers-of-pseudoscience/
4. Gordin, M. D. The problem with pseudoscience: pseudoscience is not the antithesis of professional science but thrives in science's shadow. *EMBO Rep.* **18**, 1482 (2017).
5. Townson, S. Why people fall for pseudoscience (and how academics can fight back). *The Guardian* (2016); https://www.theguardian.com/higher-education-network/2016/jan/26/why-people-fall-for-pseudoscience-and-how-academics-can-fight-back
6. Caulfield, T. Pseudoscience and COVID-19—we've had enough already. *Nature* https://doi.org/10.1038/d41586-020-01266-z (2020).
7. Pennycook, G. & Rand, D. G. Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* **188**, 39–50 (2019).
8. Vraga, E. K. & Bode, L. Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. *Polit. Commun.* https://doi.org/10.1080/10584609.2020.1716500 (2020).
9. Lewandowsky, S. et al. The Debunking Handbook 2020. *Databrary* https://doi.org/10.17910/b7.1182 (2020).
10. Pennycook, G. et al. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021).
11. Garrett, R. K., Weeks, B. E. & Neo, R. L. Driving a wedge between evidence and beliefs: how online ideological news exposure promotes political misperceptions. *J. Comput.-Mediat. Commun.* **21**, 331–348 (2016).
12. Lazer, D. M. J. et al. The science of fake news: addressing fake news requires a multidisciplinary effort. *Science* **359**, 1094–1096 (2018).
13. Wyer, R. S. & Unverzagt, W. H. Effects of instructions to disregard information on its subsequent recall and use in making judgments. *J. Pers. Soc. Psychol.* **48**, 533–549 (1985).
14. Greitemeyer, T. Article retracted, but the message lives on. *Psychon. Bull. Rev.* **21**, 557–561 (2014).
15. McDiarmid, A. D. et al. Psychologists update their beliefs about effect sizes after replication studies. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-021-01220-7 (2021).
16. Yousuf, H. et al. A media intervention applying debunking versus non-debunking content to combat vaccine misinformation in elderly in the Netherlands: a digital randomised trial. *EClinicalMedicine* **35**, 100881 (2021).
17. Kuru, O. et al. The effects of scientific messages and narratives about vaccination. *PLoS ONE* **16**, e0248328 (2021).
18. Anderson, C. A. Inoculation and counterexplanation: debiasing techniques in the perseverance of social theories. *Soc. Cogn.* **1**, 126–139 (1982).
19. Jacobson, N. G. *What Does Climate Change Look Like to You? The Role of Internal and External Representations in Facilitating Conceptual Change about the Weather and Climate Distinction* (Univ. Southern California, 2022).
20. Pluviano, S., Watt, C. & Sala, S. D. Misinformation lingers in memory: failure of three pro-vaccination strategies. *PLoS ONE* **12**, 15 (2017).
21. Maertens, R., Anseel, F. & van der Linden, S. Combatting climate change misinformation: evidence for longevity of inoculation and consensus messaging effects. *J. Environ. Psychol.* **70**, 101455 (2020).
22. Chan, M. S., Jones, C. R., Jamieson, K. H. & Albarracin, D. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. *Psychol. Sci.* **28**, 1531–1546 (2017).
23. Janmohamed, K. et al. Interventions to mitigate vaping misinformation: a meta-analysis. *J. Health Commun.* **27**, 84–92 (2022).
24. Walter, N. & Tukachinsky, R. A meta-analytic examination of the continued influence of misinformation in the face of correction: how powerful is it, why does it happen, and how to stop it? *Commun. Res.* **47**, 155–177 (2020).
25. Walter, N., Cohen, J., Holbert, R. L. & Morag, Y. Fact-checking: a meta-analysis of what works and for whom. *Polit. Commun.* **37**, 350–375 (2020).
26. Walter, N. & Murphy, S. T. How to unring the bell: a meta-analytic approach to correction of misinformation. *Commun. Monogr.* **85**, 423–441 (2018).
27. Walter, N., Brooks, J. J., Saucier, C. J. & Suresh, S. Evaluating the impact of attempts to correct health misinformation on social media: a meta-analysis. *Health Commun.* **36**, 1776–1784 (2021).
28. Chan, M. S., Jamieson, K. H. & Albarracín, D. Prospective associations of regional social media messages with attitudes and actual vaccination: A big data and survey study of the influenza vaccine in the United States. *Vaccine* **38**, 6236–6247 (2020).
29. Lawson, V. Z. & Strange, D. News as (hazardous) entertainment: exaggerated reporting leads to more memory distortion for news stories. *Psychol. Pop. Media Cult.* **4**, 188–198 (2015).

30. Nature Microbiology. Exaggerated headline shock. *Nat. Microbiol.* **4**, 377–377 (2019).
31. Pinker, S. The media exaggerates negative news. This distortion has consequences. *The Guardian* (2018); https://www.theguardian.com/commentisfree/2018/feb/17/steven-pinker-media-negative-news
32. CDC. HPV vaccine safety. *U.S. Department of Health & Human Services* https://www.cdc.gov/hpv/parents/vaccinesafety.html (2021).
33. Jaber, N. Parent concerns about HPV vaccine safety increasing. *National Cancer Institute* https://www.cancer.gov/news-events/cancer-currents-blog/2021/hpv-vaccine-parents-safety-concerns (2021).
34. Brody, J. E. Why more kids aren't getting the HPV vaccine. *The New York Times* https://www.nytimes.com/2021/12/13/well/live/hpv-vaccine-children.html (2021).
35. Walker, K. K., Owens, H. & Zimet, G. 'We fear the unknown': emergence, route and transfer of hesitancy and misinformation among HPV vaccine accepting mothers. *Prev. Med. Rep.* **20**, 101240 (2020).
36. Normile, D. Japan reboots HPV vaccination drive after 9-year gap. *Science* **376**, 14 (2022).
37. Larson, H. J. Japan's HPV vaccine crisis: act now to avert cervical cancer cases and deaths. *Lancet Public Health* **5**, e184–e185 (2020).
38. Soroka, S., Fournier, P. & Nir, L. Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proc. Natl Acad. Sci. USA* **116**, 18888–18892 (2019).
39. Baumeister, R. F., Bratslavsky, E., Finkenauer, C. & Vohs, K. D. Bad is stronger than good. *Rev. Gen. Psychol.* **5**, 323–370 (2001).
40. Kunda, Z. The case for motivated reasoning. *Psychol. Bull.* **108**, 480–498 (1990).
41. Kopko, K. C., Bryner, S. M. K., Budziak, J., Devine, C. J. & Nawara, S. P. In the eye of the beholder? Motivated reasoning in disputed elections. *Polit. Behav.* **33**, 271–290 (2011).
42. Leeper, T. J. & Mullinix, K. J. Motivated reasoning. *Oxford Bibliographies* https://doi.org/10.1093/OBO/9780199756223-0237 (2018).
43. Johnson, H. M. & Seifert, C. M. Sources of the continued influence effect: when misinformation in memory affects later inferences. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 1420–1436 (1994).
44. Wilkes, A. L. & Leatherbarrow, M. Editing episodic memory following the identification of error. *Q. J. Exp. Psychol. Sect. A* **40**, 361–387 (1988).
45. Ecker, U. K. H., Lewandowsky, S. & Apai, J. Terrorists brought down the plane!—No, actually it was a technical fault: processing corrections of emotive information. *Q. J. Exp. Psychol.* **64**, 283–310 (2011).
46. Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N. & Cook, J. Misinformation and its correction: continued influence and successful debiasing. *Psychol. Sci. Public Interest* **13**, 106–131 (2012).
47. Nyhan, B. & Reifler, J. Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine* **33**, 459–464 (2015).
48. Nyhan, B., Reifler, J., Richey, S. & Freed, G. L. Effective messages in vaccine promotion: a randomized trial. *Pediatrics* **133**, e835–e842 (2014).
49. Nyhan, B. & Reifler, J. When corrections fail: the persistence of political misperceptions. *Polit. Behav.* **32**, 303–330 (2010).
50. Rathje, S., Roozenbeek, J., Traberg, C. S., van Bavel, J. J. & van der Linden, S. Meta-analysis reveals that accuracy nudges have little to no effect for U.S. conservatives: regarding Pennycook et al. (2020). *Psychol. Sci.* https://doi.org/10.25384/SAGE.12594110.v2 (2021).
51. Greene, C. M., Nash, R. A. & Murphy, G. Misremembering Brexit: partisan bias and individual predictors of false memories for fake news stories among Brexit voters. *Memory* **29**, 587–604 (2021).
52. Gawronski, B. Partisan bias in the identification of fake news. *Trends Cogn. Sci.* **25**, 723–724 (2021).
53. Pennycook, G. & Rand, D. G. Lack of partisan bias in the identification of fake (versus real) news. *Trends Cogn. Sci.* **25**, 725–726 (2021).
54. Borukhson, D., Lorenz-Spreen, P. & Ragni, M. When does an individual accept misinformation? An extended investigation through cognitive modeling. *Comput. Brain Behav.* **5**, 244–260 (2022).
55. Roozenbeek, J. et al. Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking susceptibility to misinformation. *Judgm. Decis. Mak.* **17**, 547–573 (2022).
56. Bolsen, T., Druckman, J. N. & Cook, F. L. The influence of partisan motivated reasoning on public opinion. *Polit. Behav.* **36**, 235–262 (2014).
57. Hameleers, M. & van der Meer, T. G. L. A. Misinformation and polarization in a high-choice media environment: how effective are political fact-checkers? *Commun. Res.* **47**, 227–250 (2020).
58. Guay, B., Berinsky, A., Pennycook, G. & Rand, D. How to think about whether misinformation interventions work. Preprint at *PsyArXiv* https://doi.org/10.31234/OSF.IO/GV8QX (2022).
59. Hove, M. J. & Risen, J. L. It's all in the timing: interpersonal synchrony increases affiliation. *Soc. Cogn.* **27**, 949–960 (2009).
60. Tesch, F. E. Debriefing research participants: though this be method there is madness to it. *J. Pers. Soc. Psychol.* **35**, 217–224 (1977).
61. Tanner-Smith, E. E. & Tipton, E. Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Res Synth. Methods* **5**, 13–30 (2014).
62. Tanner-Smith, E. E., Tipton, E. & Polanin, J. R. Handling complex meta-analytic data structures using robust variance estimates: a tutorial in R. *J. Dev. Life Course Criminol.* **2**, 85–112 (2016).
63. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.*, https://doi.org/10.18637/jss.v036.i03 (2010).
64. van Aert, R. C. M. CRAN—package puniform. *R Project* https://cran.r-project.org/web/packages/puniform/index.html (2022).
65. Coburn, K. M. & Vevea, J. L. weightr: estimating weight-function models for publication bias. (2021); https://cran.r-project.org/web/packages/weights/index.html
66. Fisher, Z. & Tipton, E. robumeta: an R-package for robust variance estimation in meta-analysis. *ArXiv*. https://doi.org/10.48550/arXiv.1503.02220 (2015).
67. Sidik, K. & Jonkman, J. N. Robust variance estimation for random effects meta-analysis. *Comput. Stat. Data Anal.* **50**, 3681–3701 (2006).
68. Hedges, L. V., Tipton, E. & Johnson, M. C. Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* **1**, 39–65 (2010).
69. JASP Team. *JASP* (2022); https://jasp-stats.org/
70. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *Br. Med. J.* **327**, 557–560 (2003).
71. Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558 (2002).
72. Tay, L. Q., Hurlstone, M. J., Kurz, T. & Ecker, U. K. H. A comparison of prebunking and debunking interventions for implied versus explicit misinformation. *Br. J. Psychol.* **113**, 591–607 (2022).

73. Tappin, B. M., Berinsky, A. J. & Rand, D. G. Partisans' receptivity to persuasive messaging is undiminished by countervailing party leader cues. *Nat. Hum. Behav.*, https://doi.org/10.1038/s41562-023-01551-7 (2023).
74. Traberg, C. S. & van der Linden, S. Birds of a feather are persuaded together: perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Pers. Individ. Dif.* **185**, 111269 (2022).
75. van Bavel, J. J. & Pereira, A. The partisan brain: an identity-based model of political belief. *Trends Cogn. Sci.* **22**, 213–224 (2018).
76. Kahan, D. M. Misconceptions, misinformation, and the logic of identity-protective cognition. *SSRN Electron. J.* https://doi.org/10.2139/SSRN.2973067 (2017).
77. Levendusky, M. *Our Common Bonds: Using What Americans Share to Help Bridge the Partisan Divide* (Univ. Chicago Press, 2023).
78. Voelkel, J. G. et al. Interventions reducing affective polarization do not improve anti-democratic attitudes. *Nature Human Behaviour*, **7**, 55–64 (2023); https://doi.org/10.31219/OSF.IO/7EVMP
79. Ecker, U. K. H., Hogan, J. L. & Lewandowsky, S. Reminders and repetition of misinformation: helping or hindering its retraction? *J. Appl. Res. Mem. Cogn.* **6**, 185–192 (2017).
80. Schwarz, N., Sanna, L. J., Skurnik, I. & Yoon, C. Metacognitive experiences and the intricacies of setting people straight: implications for debiasing and public information campaigns. in. *Adv. Exp. Soc. Psychol.* **39**, 127–161 (2007).
81. Ecker, U. K. H., Lewandowsky, S. & Chadwick, M. Can corrections spread misinformation to new audiences? Testing for the elusive familiarity backfire effect. *Cogn. Res Princ. Implic.* **5**, 41 (2020).
82. Kappel, K. & Holmen, S. J. Why science communication, and does it work? A taxonomy of science communication aims and a survey of the empirical evidence. *Front. Commun.* **4**, 55 (2019).
83. Fischhoff, B. The sciences of science communication. *Proc. Natl Acad. Sci. USA* **110**, 14033–14039 (2013).
84. Winters, M. et al. Debunking highly prevalent health misinformation using audio dramas delivered by WhatsApp: evidence from a randomised controlled trial in Sierra Leone. *BMJ Glob. Health* **6**, 6954 (2021).
85. Registered replication reports. *Association for Psychological Science* http://www.psychologicalscience.org/publications/replication (2017).
86. Vraga, E. K., Kim, S. C. & Cook, J. Testing logic-based and humor-based corrections for science, health, and political misinformation on social media. *J. Broadcast Electron. Media* **63**, 393–414 (2019).
87. Vijaykumar, S. et al. How shades of truth and age affect responses to COVID-19 (mis)information: randomized survey experiment among WhatsApp users in UK and Brazil. *Humanit. Soc. Sci. Commun.* **8**, 1–12 (2021).
88. Anderson, C. A., Lepper, M. R. & Ross, L. Perseverance of social theories: the role of explanation in the persistence of discredited information. *J. Pers. Soc. Psychol.* **39**, 1037–1049 (1980).
89. Sirlin, N., Epstein, Z., Arechar, A. A. & Rand, D. G. Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. *Harv. Kennedy Sch. Misinformation Rev.*, https://doi.org/10.37016/mr-2020-83 (2021).
90. Arechar, A. A. et al. Understanding and reducing online misinformation across 16 countries on six continents. Preprint at *PsyArXiv* https://psyarxiv.com/a9frz/ (2022).
91. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.* **31**, 770–780 (2020).
92. Jahanbakhsh, F. et al. Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. in *Proc. ACM on Human–Computer Interaction* vol. 5, 1--42 (Association for Computing Machinery, 2021); https://doi.org/10.1145/3449092 (2021).
93. Pennycook, G. & Rand, D. G. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc. Natl Acad. Sci. USA* **116**, 2521–2526 (2019).
94. Gesser-Edelsburg, A., Diamant, A., Hijazi, R. & Mesch, G. S. Correcting misinformation by health organizations during measles outbreaks: a controlled experiment. *PLoS ONE* **13**, e0209505 (2018).
95. Mosleh, M., Martel, C., Eckles, D. & Rand, D. Promoting engagement with social fact-checks online. Preprint at *OSF* https://osf.io/rckfy/ (2022).
96. Andrews, E. A. *Combating COVID-19 Vaccine Conspiracy Theories: Debunking Misinformation about Vaccines, Bill Gates, 5G, and Microchips Using Enhanced Correctives*. MSc thesis, State Univ. New York at Buffalo (2021).
97. Koller, M. Rebutting accusations: when does it work, when does it fail? *Eur. J. Soc. Psychol.* **23**, 373–389 (1993).
98. Greitemeyer, T. & Sagioglou, C. Does exonerating an accused researcher restore the researcher's credibility? *PLoS ONE* **10**, e0126316 (2015).
99. Hedges, L. V. & Olkin, I. *Statistical Methods for Meta-analysis* (Academic, 1985).
100. Hedges, L. V. Distribution Theory for Glass's estimator of effect size and related estimators. *J. Educ. Stat.* **6**, 107 (1981).
101. Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. *Introduction to Meta-analysis* (Wiley, 2009).
102. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Front. Psychol.* **4**, 863 (2013).
103. Morris, S. B. Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *Br. J. Math. Stat. Psychol.* **53**, 17–29 (2000).
104. Hart, W. et al. Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychol. Bull.* **135**, 555–588 (2009).
105. Lord, C. G., Ross, L. & Lepper, M. R. Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *J. Pers. Soc. Psychol.* **37**, 2098–2109 (1979).
106. Seifert, C. M. The continued influence of misinformation in memory: what makes a correction effective? *Psychol. Learn. Motiv.* **41**, 265–292 (2002).
107. van der Linden, S., Leiserowitz, A., Rosenthal, S. & Maibach, E. Inoculating the public against misinformation about climate change. *Glob. Chall.* **1**, 1600008 (2017).
108. Ecker, U. K. H. et al. The psychological drivers of misinformation belief and its resistance to correction. *Nat. Rev. Psychol.* **1**, 13–29 (2022).
109. Ecker, U., Sharkey, C. X. M. & Swire-Thompson, B. Correcting vaccine misinformation: A failure to replicate familiarity or fear-driven backfire effects. *PLoS One,* **18**, e0281140 (2023).
110. Gawronski, B., Brannon, S. M. & Ng, N. L. Debunking misinformation about a causal link between vaccines and autism: two preregistered tests of dual-process versus single-process predictions (with conflicting results). *Soc. Cogn.* **40**, 580–599 (2022).
111. Guenther, C. L. & Alicke, M. D. Self-enhancement and belief perseverance. *J. Exp. Soc. Psychol.* **44**, 706–712 (2008).
112. Misra, S. Is conventional debriefing adequate? An ethical issue in consumer research. *J. Acad. Mark. Sci.* **20**, 269–273 (1992).
113. Green, M. C. & Donahue, J. K. Persistence of belief change in the face of deception: the effect of factual stories revealed to be false. *Media Psychol.* **14**, 312–331 (2011).
114. Ecker, U. K. H. & Ang, L. C. Political attitudes and the processing of misinformation corrections. *Polit. Psychol.* **40**, 241–260 (2019).

115. Sherman, D. K. & Kim, H. S. Affective perseverance: the resistance of affect to cognitive invalidation. *Pers. Soc. Psychol. Bull.* **28**, 224–237 (2002).
116. Golding, J. M., Fowler, S. B., Long, D. L. & Latta, H. Instructions to disregard potentially useful information: the effects of pragmatics on evaluative judgments and recall. *J. Mem. Lang.* **29**, 212–227 (1990).
117. Viechtbauer, W. & Cheung, M. W.-L. Outlier and influence diagnostics for meta-analysis. *Res. Synth. Methods* **1**, 112–125 (2010).
118. Borenstein, M. in *Publication Bias in Meta-analysis: Prevention, Assessment, and Adjustments* (eds Rothstein, H. R., Sutton, A. J. & Borenstein, M.) 194–220 (John Wiley & Sons, 2005).
119. Duval, S. in *Publication Bias in Meta-analysis: Prevention, Assessment, and Adjustments* (eds Rothstein, H. R., Sutton, A. J. & Borenstein, M.) 127–144 (John Wiley & Sons, 2005).
120. Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R. & Rushton, L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J. Clin. Epidemiol.* **61**, 991–996 (2008).
121. Stanley, T. D. & Doucouliagos, H. Meta-regression approximations to reduce publication selection bias. *Res. Synth. Methods* **5**, 60–78 (2014).
122. van Assen, M. A. L. M., van Aert, R. C. M. & Wicherts, J. M. Meta-analysis using effect size distributions of only statistically significant studies. *Psychol. Methods* **20**, 293–309 (2015).
123. Pustejovsky, J. E. & Rodgers, M. A. Testing for funnel plot asymmetry of standardized mean differences. *Res. Synth. Methods* **10**, 57–71 (2019).
124. Maier, M., Bartoš, F. & Wagenmakers, E. J. Robust Bayesian meta-analysis: addressing publication bias with model-averaging. *Psychol. Methods*, https://doi.org/10.1037/met0000405 (2022).

Acknowledgements

We thank D. O'Keefe, who assisted in the inter-rater reliability. Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R01MH114847 (D.A.), the National Institute on Drug Abuse of the National Institutes of Health under Award Number DP1 DA048570 (D.A.) and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R01AI147487 (D.A. and M.S.C.) and P30AI045008 (Penn Center for AIDS Research [Penn CFAR] subaward; M.S.C.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research was supported by the Science of Science Communication Endowment from the Annenberg Public Policy Center at the University of Pennsylvania. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

D.A. initiated the project, and M.S.C. supervised the project. Both M.S.C. and D.A. contributed to the theoretical formalism, developed the coding scheme and performed the coding reliability. M.S.C. took the lead in the data curation, preparing the analytical plan and performing the analytic calculations. Both M.S.C. and D.A. discussed the results and contributed to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-023-01623-8.

Correspondence and requests for materials should be addressed to Man-pui Sally Chan.

Peer review information *Nature Human Behaviour* thanks Jon Roozenbeek, Sander van der Linden and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Man-pui Sally Chan

Last updated by author(s): Apr 1, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We did not use any software to collect data for this project.
Data analysis	We used four add-on packages for the statistical software R version 4.0.5: robumeta version 2.0 (Tanner-Smith et al., 2016; Tanner-Smith & Tipton, 2014), metafor version 3.9.9 (Viechtbauer, 2010), puniform version 0.2.5 (van Aert, 2022), and weightr version 2.0.2 (Coburn & Vevea, 2017) to assess publication/inclusion bias and analyze the mean effect sizes using robust variance estimation (RVE) methods (Fisher & Tipton, 2015; Hedges et al., 2010; Sidik & Jonkman, 2006). We also used JASP version 0.16.3 (JASP Team, 2022), an open-source statistics program, to conduct bias analyses with Bayesian methods. We also used Amelia version 1.8.0 1 and mice version 3.14.0 2 to impute the misinformation effect sizes, and the results were reported in the Supplementary Information.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data that support the findings of this study are openly available in OSF at https://osf.io/vkygw/.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study was a meta-analysis of a large body of experimental evidence comprising 60 published experiments, five working papers, two theses, and seven unpublished datasets (total number of reports = 74; k of effect sizes = 205). The data are quantitative.
Research sample	The reports were obtained from searches of online databases, e.g., PsycInfo. We included all the samples as reported in the included reports. The number of participants in the synthesis ranged from 7 to 1180, their mean age was 36 years old (SD = 10.73), and about 60 percent of them were female. Participants included university students, graduate students, and adults from the community recruited via online survey platforms, such as Amazon Mechanical Turk and Prolific.
Sampling strategy	The sample size was the total number of participants reported in the included reports. No additional sampling strategy was used.
Data collection	We used specific keywords with wildcards, performing a combined search of the following seven online databases: (a) PsycInfo, (b) Google Scholar, (c) MEDLINE, (d) PubMed, (e) ProQuest Dissertations and Theses Abstracts and Indexes: Social Sciences, (f) Communication Source, and (g) Social Sciences Citation Index. We paired a series of keywords (i.e., misinformation OR misbelief* OR false information OR belief perseverance OR continued influence) with two other series of keywords (i.e., [retract* OR correct*] and [label* OR tag* OR flag*]). We expanded the sample of relevant reports by examining the reference lists of a systematic collection of review articles, book chapters, and dissertations. By culling the reference lists of the review papers obtained through the database searches, we were able to identify eight additional articles. We also identified ten studies after contacting a list of researchers who have researched in this area but have no awareneess of the research questions and hypotheses. Additionally, we received materials from additional researchers after posting requests on online forums and e-mail list servers (e.g., Society for Personality and Social Psychology). Finally, we searched the Open Science Framework (OSF) using the same set of keywords in early 2021 to obtain 258 unpublished and in-press data sets. Authors and coders were blinded to the experimental conditional and hypotheses of the included reports. They were not blinded to the meta-analysis's research questions as they were preregistered at OSF https://osf.io/vkygw
Timing	We did not set the start date of the literature search and the end date was on August 31 2022.

Data exclusions	We included (a) studies that assessed false claims concerning a scientific measurement procedure or scientific evidence. We used several eligibility criteria to select reports for inclusion in the meta-analysis, inspecting only studies from reports that were clearly or possibly experimental. Included studies (b) empirically measured participants' beliefs in or attitudes consistent with the misinformation addressed by the correction (see Table S3 for examples of excluded reports after evaluating the full text). However, studies that included outcome measures of misinformation sharing (intentions) (Arechar et al., 2022; Jahanbakhsh et al., 2021; Pennycook et al., 2020, 2021; Sirlin et al., 2021), the quality judgment of news sources (Pennycook & Rand, 2019a), self-efficacy, openness to the messages (Gesser-Edelsburg et al., 2018), and whether participants responded to/ignored the corrections were excluded (Mosleh et al., 2022). The included studies (c) had a control or baseline group exposed to either no message, a neutral message, or an unrelated message and (d) were eligible when the misinformation was initially asserted to be true or was known to participants prior to the study and was later corrected. All studies introduced a correction for the misinformation regardless of whether the misinformation was fictitious (e.g., Anderson et al., 1980) and known to be familiar to the public (e.g., Andrews, 2021; Yousuf et al., 2021). However, studies that described the initial information as hypothetical or uncertain (e.g., Koller, 1993) or as an accusation of scientific misconduct (e.g., Greitemeyer & Sagioglou, 2015) were excluded. The coding protocol and exclusion criteria were preregistrated in mid-2022 at OSF https://osf.io/vkygw.
Non-participation	No participants were involved in this study
Randomization	Randomization was not applicable to this study as this is a meta-analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging