

Effects of fact-checking warning labels and social endorsement cues on climate change fake news credibility and engagement on social media

Timo K. Koch^{1,2}  | Lena Frischlich³  | Eva Lermer^{2,4} 

¹Department of Psychology, LMU Munich, Munich, Germany

²LMU Center for Leadership and People Management, LMU Munich, Munich, Germany

³Department of Communication, University of Münster, Münster, Germany

⁴Augsburg University of Applied Sciences, Augsburg, Germany

Correspondence

Timo K. Koch, Department of Psychology, LMU Munich, Munich, Germany.
Email: timo.koch@psy.lmu.de

Funding information

Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen

Abstract

Online fake news can have noxious consequences. Social media platforms are experimenting with different interventions to curb fake news' spread, often employing them simultaneously. However, research investigating the interaction of these interventions is limited. Here, we use the heuristic-systematic model of information processing (HSM) as a theoretical framework to jointly test two interventions against fake news that are implemented at scale by social media platforms: (1) adding warning labels from fact checkers to initiate systematic processing and (2) removing social endorsement cues (e.g., engagement counts) to reduce the influence of this heuristic cue. Moreover, we accounted for dispositions previously found to affect a person's response to fake news through motivated reasoning or cognitive style. An online experiment in Germany ($N = 571$) confirmed that warning labels reduced the perceived credibility of a fake news post exaggerating the consequences of climate change. Warning labels also lowered the (self-reported) likelihood to amplify fake news. Removing social endorsement cues did not have an effect. In line with research on motivated reasoning, left-leaning individuals perceived the climate fake news to be more credible and reported a higher likelihood to amplify it. Supporting research on cognitive style, participants with lower educational levels and a less analytic thinking style also reported a higher likelihood of amplification. Elaboration likelihood was associated only with age, involvement, and political leaning, but not affected by warning labels. Our findings contribute to the mounting evidence for the effectiveness of warning labels while questioning their relevance for systematic processing.

KEYWORDS

fake news interventions, heuristic-systematic model, social media

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Applied Social Psychology* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Since the presidential election in the United States in 2016, so-called fake news—“fabricated information that mimics news media content in form but not in organizational process or intent” (Lazer et al., 2018; p. 1094)—has raised global concerns, since its dissemination can have severe consequences for individual and collective well-being. Different regulatory attempts, such as the European Commission's code of practice on disinformation, try to curb the spread of fake news (European Commission, 2018). Under increasing pressure to manage the “information disorder” on its platform (Donovan, 2020), Facebook, the largest social network website in many countries, partners with different fact-checking institutions to combat mis- and disinformation, for instance via warning labels. Once a piece of content is found to be false, its algorithmic recommendation is reduced, and the misinformation receives a warning label stating that its content is disputed by fact checkers (Facebook, 2020). Similar warning labels are also used by Twitter (Roth & Pickles, 2020), but they exclude misinformation related to COVID-19 since the takeover by Elon Musk in 2022 (Reuters, 2022). Warning labels on both platforms directly link the fact checks to the warning labels, thus aiming at fostering a more systematic processing of the material.

Beyond warning labels, Meta, the company owning Facebook, also started to allow users to remove social endorsement cues (i.e., the like count) (for a media report, see Constantine, 2019; Rizioiu, 2019). Prior research had shown that such endorsement cues increase the perceived credibility of both factual and fake news and motivate their amplification (i.e., liking or sharing) (Luo et al., 2022). The effect of this heuristic endorsement cue was stronger for fake as compared to factual news (Avram et al., 2020). Thus, removing social endorsement cues could actively contribute to reducing the spread of fake news. However, interaction effects between different countermeasures have gained comparatively little academic interest so far (for a systematic review, see Ziemer & Rothmund, 2022).

In the present work, we studied the interplay of fact-checking warning labels and the removal of social endorsement cues using the heuristic-systematic model of information processing (HSM) (Chaiken, 1980) as a theoretical framework for deriving hypotheses about their effects. We judged the effectiveness of these interventions by examining users' perceived credibility of fake news and their likelihood to amplify fake news posts versus elaborating on them more carefully. Furthermore, we accounted for mounting evidence suggesting that the response to fake news depends on two central psychological characteristics: First, people's cognitive styles, such that those with a more superficial, heuristic processing style are more prone to fall for fake news, and, second, the attitudinal consistency between their worldviews and the fake news, that is, their motivational processing of fake news' content (for a systematic review, see Ziemer & Rothmund, 2022). Thus, we controlled our analyses for a set of socio-demographic and psychological dispositions mirroring these two dimensions based on prior research (e.g., Guess et al., 2019; Pennycook & Rand, 2020).

2 | THEORETICAL BACKGROUND

2.1 | The HSM

From a cognitive perspective, the processing of fake news follows the same mechanics as the processing of information more generally. Following two process models of cognition, these mechanics can be distinguished between a fast, peripheral process (sometimes referred to as system I), and a more systematic, elaborated system II path, where information is elaborated carefully (Kahneman, 2013; Petty & Cacioppo, 1986). The HSM, applied as the central framework in the current study, is a two-process model that describes more specifically how “individuals make judgments in the light of information” (Chaiken & Ledgerwood, 2012; p. 254). In the HSM, the systematic (system II) mode describes situations in which information is elaborated carefully and with the intention to thoroughly understand any available information. The heuristic mode (system I) describes conditions under which salient and easily comprehended cues are used as shortcuts for opinion formation (Chaiken, 1980, 1987). Relevant to the current study, a more systematic elaboration is associated with a lower susceptibility to fake news (Pennycook & Rand, 2021). For instance, nudging users to make accurate decisions has been shown to reduce the sharing of fake news (Pennycook et al., 2020).

Heuristic and systematic processing can co-occur and shape each other (Chen et al., 1999; Zuckerman & Chaiken, 1998). As systematic processing requires mental effort, people engage in systematic processing only when they are able and motivated to invest the necessary resources (comparable to the elaboration-likelihood model, see Petty et al., 1983). Heuristic processing needs much less effort and can be viewed as relatively automatic, as it focusses on easily noted and understood cues, so-called heuristics. Examples for such heuristics include shortcuts, such as “experts know best” or “consensus implies correctness” (Chaiken & Ledgerwood, 2012; p. 247). Other motives such as needs for accuracy or defending one's worldview also steer the employment of heuristic and/or systematic processing (Chen et al., 1999). Following the theory of motivated reasoning (Kunda, 1990), engagement in systematic versus heuristic processing can further depend on peoples' motivational goals. If they are motivated to form accurate opinions, they are more likely to engage in systematic processing to form accurate conclusions. However, if people are motivated by directional goals (e.g., the desire to defend their worldview), they are more likely to use whatever strategy brings them to their desired conclusion (e.g., that their worldview is correct).

The HSM had been suggested as a framework for understanding the effectiveness of product warning labels (Zuckerman & Chaiken, 1998) and has recently gained increased scholarly attention as a theoretical framework to study factual and fake news (Ali et al., 2022; Duncan, 2020; Kaye & Johnson, 2021; Sundar et al., 2021). For instance, Ali et al. (2021) showed that fake news posts but not real news posts accompanied by a high number of likes—a heuristic cue for the social endorsement of the content—were perceived to be more credible.

2.2 | Perceived credibility of fake news

Perceived credibility is a major factor determining persuasion and one of the central variables by which the effects of fake news are judged (for a review, see Bryanov & Vziatysheva, 2021). Since the seminal work by Hovland et al. (1953), research has repeatedly shown that credible sources are more persuasive than noncredible ones (for a meta-analysis, see Pornpitakpan, 2004). Perceived credibility can also refer to a message itself. Following Appelman and Sundar (2016), message credibility is “an individual's judgment of the veracity of the content of communication” (p. 63) driven by perceiving the message to be authentic, accurate, and believable. Source and message credibility can interact with each other (Flanagin et al., 2020). Although perceived credibility can mediate subsequent (fake) news effects (Halpern et al., 2019), it is an “effect in its own right” (Appelman & Sundar, 2016; p. 63) and thus studied as our first central dependent variable in this work.

The perceived credibility of fake news is likely driven more by heuristics than by systematic evaluations of fake news' content. Perceived message credibility is often the result of heuristics (Chaiken, 1980; Petty & Cacioppo, 1986) and interviews with media users using think-aloud protocols have shown that users seldomly scrutinize fake news carefully (Freiling, 2019). One central heuristic could be the social endorsement of a message, that is, the number of likes or shares on social media (the so-called bandwagon effect, see Sundar et al., 2008). For instance, a high (vs. low) number of likes has been found to increase the perceived credibility of both fake and factual news (Luo et al., 2022). Warning labels that allow for a more careful engagement with online content try to reduce the impact of such heuristics and generally foster more careful engagement and systematic processing to reduce the perceived credibility of fake news (van der Linden, 2018). However, reducing the perceived credibility of fake news alone is not enough, as users might engage with fake news on social media despite their disbelief (Pennycook et al., 2021).

2.3 | Engagement with fake news

Social media platforms allow users to actively respond to the content they see online, for instance by liking or sharing it. Users' sharing in particular is a central factor for fake news' virality (Vosoughi et al., 2018), and some scholars have argued for focusing on interventions reducing sharing in the first place (Pennycook et al., 2021). Algorithmically curated social media platforms, such as Facebook or Twitter, often use user engagement as an indicator for the value of content and further amplify it by recommending it to other users (e.g., in the Twitters Trends). Thus, even though there are several motivations underlying “liking” and “sharing” on social media (Hayes et al., 2016), behavioral engagement with social media content contributes to amplification. Reducing the amplification of fake news is one key effectiveness indicator for fake news interventions.

From the perspective of the HSM and in light of the benefits of systematic processing for individuals' resilience against fake news (Pennycook & Rand, 2019), it is also desirable to foster elaboration. Qualitative research (Freiling, 2019; Tandoc et al., 2018) has indicated that two elaboration strategies are frequently employed by social media users: Exchanging with others about a suspected fake news piece and searching for additional information. Consequentially, countermeasures that increase such elaboration techniques are desirable.

2.4 | The current research focus: Fake news interventions

2.4.1 | Warning labels

Facebook implemented warning labels (indicating “disputed by 3rd party fact-checkers”) for the first time in December 2016 (Mosseri, 2016) and started to roll out a similar strategy on its subsidiary company Instagram 3 years later (Instagram, 2019). Currently, these labels do not only inform people that the content of a post has been disputed but also link to further information about the disputation. The effectiveness of warning labels is often evaluated from the rate of users clicking on the additional information (for a media report, see The Economic Times, 2021). Thus, the underlying theoretical assumption of warning labels is that they can nudge people away from relying primarily on heuristic processing (van der Linden, 2018).

Prior research on the effectiveness of warning labels to reduce the perceived credibility of fake news provided some positive evidence. For instance, Clayton et al. (2020) showed that warning labels reduced the perceived accuracy of fake news about former U.S. president Donald Trump. Similarly, Duncan (2020) found that negative credibility labels decreased the perceived credibility of a news story. In contrast, Oeldorf-Hirsch et al. (2020) found no such effect of warning labels presented together with either news memes (Study 1) or news articles (Study 2), suggesting that the effect of warning labels might be conditional on other factors. Compatible with this idea, Arendt et al. (2019) found warning labels to reduce the perceived accuracy of fake news among participants leaning towards the political left but not among those leaning towards the right.

Warning labels might also reduce the amplification of fake news, although research using behavioral measures is less prevalent. Pennycook et al. (2020) found that fake news with a “disputed” label was perceived to be less accurate and shared less often than posts without such a label. Mena (2020) found that participants were less likely to share fake news when it had been labeled as false. We are currently not aware of a study that explicitly evaluated the effects of warning labels on the elaboration of fake news. We thus formulated one confirmatory hypothesis and one exploratory research question to formalize our expectations regarding the effects of warning labels.

H1: Perceived credibility of a fake news post with a warning label attached is lower compared to one without a warning label.

RQ1: Do users engage (amplification and elaboration) less with a fake news post with a warning label than one without a warning label?

2.4.2 | Social endorsement cues

One heuristic cue that has been typical for social media is the social endorsement of a post, for example, the number of likes, shares, and views a post has received. These cues serve as an indicator for the popularity of the content through other users and thus can trigger a bandwagon effect (Sundar et al., 2008), such that recipients perceive posts with multiple likes as being more valuable and credible (Metzger & Flanagin, 2013).

Social endorsement cues can affect the perceived credibility of fake news. Luo et al. (2022) found that a high number of Facebook likes increased the perceived credibility of both fake and factual news. Fake news might benefit even more from endorsement than real news: A high (vs.) low number of likes increases the perceived credibility of conspiracy theories on Facebook (Shin et al., 2022), and a study by Ali et al. (2021) showed a bandwagon effect for fake but not real news. These and other findings likely motivated Meta to hide like counts on Instagram in 2019.

Social endorsement cues can also affect engagement with fake news. For instance, Winter et al. (2015) found that users were more likely to engage more with Facebook posts with many as compared to few likes. Avram et al. (2020) showed that the number of likes was positively associated with liking and sharing fake news, but negatively with fact-checking it. Relatedly, Molina et al. (2022) found that users were more likely to comment on fake news with many (vs. few) likes. As research using behavioral measures as a dependent variable is generally less prevalent, we formulated one confirmatory hypothesis and one exploratory research question to formalize our expectations.

H2: Perceived credibility of a fake news post with social endorsement cues is higher compared to one without social endorsement cues.

RQ2: Do users engage (amplification and elaboration) more with a fake news post with social engagement cues than one without social engagement cues?

2.4.3 | The interaction of warning labels and social endorsement cues

Social media platforms often combine different strategies to combat fake news, but studies on potential interactions between interventions are rare (Bak-Coleman et al., 2022; Ziemer & Rothmund, 2022). Furthermore, these studies often test either the interplay of different heuristics or the combination of measures targeted towards fostering systematic processing. For instance, Lin et al. (2016) found that multiple heuristics affected the perceived credibility of a tweet. Hameleers (2022) showed that combining news media literacy interventions and fact

checks worked hand-in-hand to reduce users' agreement with fake news. In the present work, we sought to broaden the knowledge about such effects by exploring how the heuristic social endorsement cues would interact with the warning labels aimed to motivate systematic processing, with an eye toward different possible patterns of interaction (i.e., systematic processing triggered by warning labels overruling heuristic processes coming from social endorsement cues or vice versa). We thus formulated the following exploratory research questions:

RQ3: How do social endorsement and warning labels interact with each other in affecting the perceived credibility of a fake news post?

RQ4: How do social endorsement and warning labels interact with each other in affecting the engagement (amplification and elaboration) with a fake news post?

2.5 | Interindividual differences

Research on warning labels has provided some evidence for interindividual differences such as political attitudes shaping their effectiveness (Arendt et al., 2019). There is also a study by Lee et al. (2021) that found that social endorsement cues can have a negative effect on the perceived credibility of Facebook posts, suggesting that interindividual differences might also play a role for social endorsement cues' relevance.

Drawing from research on fake news processing and effects, interindividual differences related to individuals' cognitive style and attitudinal consistency when encountering fake news are central. While the cognitive variables shape the likelihood of engaging in systematic processing (elaboration likelihood model, Petty & Cacioppo, 1986), the attitudinal variables shape motivational processing (Kunda, 1990). Thus, both types of variables should be considered in research about countermeasures (Ziemer & Rothmund, 2022).

Variables associated with individuals' cognitive style include their education, their analytical style, and their involvement with the topic of the post. With regard to educational level, surveys show that those with a lower (vs. higher) educational level have lower news literacy (Meßmer et al., 2021), perform worse in a fake news detection task (Preston et al., 2021), and are more likely to believe in fake news (Allcott & Gentzkow, 2017). One reason for these findings could be that educational level is associated with peoples' cognitive abilities. There is mounting evidence that these abilities, particularly people's analytic thinking, shape responses to fake news (Pennycook & Rand, 2019, 2020; for an overview, see 2021). From the perspective of two-process models, this is due to a higher likelihood of systematic processing when individuals have a greater capacity for systematic elaboration (Cacioppo & Petty, 1982; Petty & Cacioppo, 1986). Another central variable that fosters systematic processing is the interest in the topic, that is, the involvement with an issue (Johnson & Eagly, 1989; Petty et al., 1983), such that higher interest fosters more systematic processing. We thus controlled our analyses for these variables.

One variable that has gained considerable attention in fake news research is people's political attitudes. Fake news often addresses

partisan or politicized topics (Humphrecht, 2019; Marwick & Lewis, 2017) and individuals are more likely to fall for fake news supporting their prior political attitudes (Corbu & Negrea-Busuioc, 2020; Faragó et al., 2019), indicating motivated processing (Kunda, 1990). For instance, ultraconservative US citizens were found to share more fake news than their liberal counterparts (Guess et al., 2019). We thus included a measure of political attitude.

Finally, there is some evidence that older media users might be more susceptible to fake news (Guess et al., 2019; Moore & Hancock, 2022). Thus we also included age as a control variable, although age does not directly correspond to either a specific cognitive style or motivated processing.

3 | METHOD

We tested our hypotheses and answered our research questions by means of an online experiment that realized a 2 (warning labels absent vs. present) \times 2 (social endorsement cues absent vs. present) between-subjects design. All materials, data, and code for this study are available in the project's repository on the Open Science Framework (OSF): <https://osf.io/ugt2v/>.

3.1 | Sample

We recruited participants via social media platforms (e.g., Facebook), university mailing lists, and websites for psychological studies between December 2017 until February 2018. A total of 674 adults completed the study. To ensure high data quality, we excluded participants who indicated that they did not comply with the study rules or showed indicators of survey speeding ($n = 17$) or who failed an attention check ($n = 86$).

The remaining $N = 571$ participants (76% women, 23% men, 1% prefer not to answer) were on average 26.23 years old ($SD = 9.80$) and rather highly educated (95% had graduated from high school). The majority (71%) were students, 25% worked (full- or part-time), 3% were unemployed, and 1% were retired. Overall, our sample resembles the global Facebook population with regard to age and education level, but women were overrepresented and nonbinary individuals underrepresented (Ribeiro et al., 2020; Statista, 2021). Most of the participants were Facebook users (85%); only 6% ($n = 32$) reported that they never used any social media. Participants' political attitudes were skewed to the left ($M = 4.73$, $SD = 1.82$ on an 11-point scale with 1 = "left-leaning" and 11 = "right-leaning").

3.2 | Procedure

Participants gave informed consent and were informed that they would participate in a study about "user behavior on social media." This information served as a cover story.

First, participants reported their social media use and rated their interest in various topics including the topic of interest, the environment. The other topics (e.g., politics, sports) served as filler items. Furthermore, we asked for their political orientation.

Next, participants were informed that they would see a Facebook feed in private and answer questions about the feed's content. We showed them a mock Facebook timeline containing fictitious content (e.g., posts, event invitations, videos, advertisements). To create a typical social media environment for the participants, the feed contained 10 posts (i.e., more than 7 ± 2 items) (Miller, 1956; Tandoc & Kim, 2022). Embedded in the middle of the feed was the fake news post. Participants were randomly assigned to see the post either with or without warning labels and with or without social endorsement cues.

We checked participants' attention by showing them the fake news post and asking specifically whether they recalled seeing or reading it. Participants who failed the attention check (i.e., who did not recall it) were excluded from the analyses as they were unable to evaluate the perceived credibility of the post. Then, we measured belief, perceived credibility, and the self-estimated likelihood to engage with that post. In the last step, we measured participants' analytical thinking and asked for their socio-demographic data. We split the measure of control variables to reduce participants' cognitive load. At the end of the study, all participants were fully debriefed and had the chance to participate in a lottery to win one of three 10€ shopping vouchers.

3.3 | Materials and measurements

3.3.1 | Fake news post

To select a topic for the fake news post that ensures a high variance in perceived credibility, we conducted a pre-test with $N = 173$ participants. During the pretest, participants rated five different fake news posts with varying topics (i.e., sports, politics, and environment; see OSF repository). All posts consisted of a picture, a headline, and a brief caption resembling regular news posts on Facebook. Participants rated each post's perceived credibility (0 = "not credible at all," 10 = "very credible") and whether they believed the content (0 = "no," 1 = "yes"). A fake news post stating that the probability for a "white Christmas" (i.e., snowfall during the Christmas holidays) in Central Europe dropped below 5% because of global warming (which is false due to regional differences) had the highest variance with regard to credibility ($M = 6.34$, $SD = 2.53$) and half of participants (53%) believed it, making an excellent candidate to examine the impact of our central variables of interest, namely warning labels and social endorsement cues. Misleading claims are a central fake news technique (see Wardle, 2018) and most fake news in Germany is not completely false but a mix of half-truths, falsehoods, and true aspects (i.e., the real-world consequences of climate change) (Bader et al., 2020).



FIGURE 1 Fake news post used in the online experiment with warning label and social endorsement cues present. The Facebook post translates to “The effects of climate change can be felt more and more,” the news headline states, “Climate change: White Christmas never again?,” the news text body states “According to researchers, the probability of snowfall around Christmas time in Central Europe drops below 5% due to climate change. Until now the...,” the warning label states: “Challenged by fact-checkers” and “learn more about why the post is contested.”

3.3.2 | Independent variables

Participants saw the fake news post either together with a warning label mimicking the one implemented by Facebook at the time of data collection (see Figure 1) or without such a label. Furthermore, the post was either presented with social endorsement cues or without them. The specific engagement metrics were based on a prestudy. For the prestudy, we crawled the engagement statistics from 16 German Facebook pages that regularly posted false and misleading information (e.g., “Compact Magazin,” “Welcome to reality”; see Boberg et al., 2020; Zimmermann & Kohring, 2018), over a 6-month time period (May 1, 2017, until October 31, 2017) using the Facebook API. The engagement metrics of the most popular posts informed our manipulation. Thus, participants in the social endorsement condition saw a post with 3754 likes, 619 comments, and 6338 shares (Figure 1).

3.3.3 | Dependent variables

Perceived credibility

We asked participants to indicate how credible they found the post on an 11-point scale (0 = “not credible at all,” 10 = “very credible”),

drawing from prior work (Luo et al., 2022). We also asked participants to indicate if they believed that the fake news post was true (0 = “yes,” 1 = “no”). Single items have the advantage of face validity and have been used previously in research on credibility and trust (Slater & Rouner, 1996), and a meta-analysis of meta-analyses recently found that single and multi-item measures perform equally well when it comes to the evaluation of advertisements (Ang & Eisend, 2018). Credibility perceptions and believing the fake news story was true were highly correlated ($r = .61$), corroborating our confidence in the measurement.

Engagement

To measure engagement, we asked participants how likely they would execute the following behaviors: Liking and sharing the post, talking with others about the post, or searching for more information about the topic. All behavioral propensities were rated using a 5-point scale (1 = “not likely at all,” 5 = “very likely”). A parallel factor analysis suggested two factors with behaviors related to amplification (liking and sharing the post) loading on the first and behaviors related to elaboration (talking to others and searching for more information) loading on the second. We thus computed two aggregated mean indices for amplification ($\alpha = .71$) and elaboration likelihood ($\alpha = .75$). Similar approaches to measure engagement using social media's concrete affordances (i.e., the behavioral options on the platform) have been employed successfully in prior work (Amazeen, 2021; Curry & Stroud, 2021).

3.3.4 | Control variables

To account for cognitive variables, we measured participants' educational level on a 5-point ordinal scale (1 = “no formal degree,” 5 = “university degree”). The variable was highly skewed; thus, we created a dummy variable distinguishing those with low education from those with high education (defined as high school or higher). To assess analytic thinking, we used the 12-item subscale of the rational/experiential multimodal inventory (Norris & Epstein, 2011) translated to German. All items were answered on a 5-point scale (1 = “strongly disagree,” 5 = “strongly agree,” $\alpha = .82$). For involvement, we asked participants to indicate on a Likert scale ranging from 1 (= “not interested at all”) to 5 (= “very interested”) how interested they were in environmental topics. Political attitude was measured on an 11-point Likert scale (0 = “left,” 10 = “right”).

3.4 | Analytical approach

We imputed missing values ($n = 47$) on the political attitude measure using predictive mean matching. To test our hypotheses and answer our research questions, we ran multivariate regression models using our dependent variables (perceived credibility, amplification, elaboration) as criteria and our independent variables (presence vs. absence of warning labels and/or social endorsement cues) as predictors.

TABLE 1 Means, standard deviations, and zero-order correlations of assessed variables

Variable	M	SD	1	2	3	4	5	6	7	8	9
1. Age	26.23	9.80									
2. Gender	0.23	0.42	.06								
3. Education	0.46	0.50	.54***	.01							
4. Pol. orientation	4.73	1.82	.08	.1*	.05						
5. Involvement	3.59	1.02	.07	.01	-.07	-.21***					
6. Analytic thinking	3.81	0.58	.08	.16***	.13**	-.02	.14***				
7. Credibility	5.67	2.26	-.07	-.12**	-.05	-.12**	.08	-.06			
8. Belief	0.35	0.50	-.02	-.04	0	-.07	.11**	.06	.61***		
9. Amplification	1.48	0.82	-.14***	-.08	-.19***	-.14***	.07	-.12**	.42***	.3***	
10. Elaboration	2.52	1.14	-.18***	0	-.01*	-.15***	.28***	.08	.26***	.28***	.33***

Note: N = 571. Zero-order Spearman's correlations.

*p < .05; **p < .01; ***p < .001.

To account for interindividual differences, we ran all analyses with and without control variables. Based on an examination of assumptions for different regression models (e.g., normal distribution of errors, homoscedasticity of variance), we used linear regression to examine perceived credibility and ordinal regression to examine amplification and elaboration propensities. All statistical analyses were performed with the statistical software R version 4.0.2 (R Core Team, 2020). Specifically, we used the psych and ordinal R packages (Christensen, 2019; Revelle, 2017).

4 | RESULTS

4.1 | Preliminary analyses

Participants were equally distributed across conditions, $\chi^2(3) = 1.30$, $p > .05$. Table 1 provides an overview of the descriptive statistics and the Pearson's correlation between the central variables in this study.

4.2 | Main analyses

Table 2 provides an overview of the results of all regression models with control variables included. Results of regression models without controls are provided in the project's OSF repository. Excluding the control variables did not change the pattern reported in the following (warning labels only had a marginally significant effect on amplification likelihood without controls though).

4.2.1 | Perceived credibility

Confirming H1, warning labels reduced the perceived credibility of fake news posts. Social endorsement cues had no such effect on

TABLE 2 Results of multivariate linear regression models for credibility perceptions and ordinal regression models for engagement (amplification and elaboration)

Variable	Credibility		Amplification		Elaboration	
	b	SE	b	SE	b	SE
Intercept	6.96***	0.77	-	-	-	-
<i>Intervention measures</i>						
Warning label shown	-.50**	0.19	-.44*	0.18	-.07	0.15
Social cues shown	.06	0.13	.14	0.12	.19	0.11
Warning x cues	-.17	0.19	-.01	0.18	-.29	0.16
<i>Control variables</i>						
Age	-.01	0.01	.01	0.01	-.03***	0.01
Education	-.12	0.20	-.83***	0.20	-.10	0.17
Political orientation	-.13*	0.05	-.16**	0.05	-.09*	0.04
Analytic thinking	-.22	0.17	-.42**	0.16	.19	0.14
Involvement	.18	0.10	.15	0.10	.54***	0.08
R ²	.04		.05		.05	

Note: N = 571. We report multiple R² for credibility perceptions and McFadden's pseudo R² for self-reported engagement likelihood.

*p < .05; **p < .01; ***p < .001.

credibility perceptions, thus rejecting H2. Addressing RQ3, the interaction between the two intervention measures did not reach statistical significance for credibility perceptions. None of the cognitive variables (educational level, analytical thinking, and involvement) significantly contributed to the explanation of the variance in perceived credibility. Indicating motivated processing, political

attitude was a significant predictor such that those with a more right-wing leaning political attitude (traditionally associated with lesser concerns for the environment) perceived the fake news that climate change would dramatically impact the weather across Europe to be less credible. Perceived credibility was not associated with participants' age.

4.2.2 | Amplification likelihood

Answering RQ1, participants who saw the warning label reported a lower likelihood that they would amplify the fake news post. Social endorsement cues had no significant effect, answering RQ2. The interaction of the two intervention measures failed to reach statistical significance, answering RQ4. Cognitive variables explained parts of the variance in amplification likelihood. Individuals with a higher educational level and those scoring higher in analytic thinking reported a lower likelihood that they would like or share the fake news post. Involvement did not have a statistically significant explanatory value beyond the other variables. Political attitude predicted amplification likelihood such that participants with a more right-wing political orientation reported a lower likelihood that they would like or share the fake news post. Amplification likelihood did not vary as a function of age.

4.2.3 | Elaboration likelihood

Neither warning labels nor social endorsement or their interaction predicted participants' elaboration likelihood. In line with the elaboration-likelihood model (Petty et al., 1983), participants who were involved with the topic of the post reported a higher likelihood to talk with others about the post or search for more information. Neither educational level nor analytical thinking had a predictive value. Furthermore, participants with more right-wing political attitudes were less likely to elaborate on the fake news post. Older participants were significantly less likely to report elaboration likelihood than younger participants.

5 | DISCUSSION

There are increasing attempts to reduce the noxious effects of fake news with countermeasures implemented on social media platforms. Despite an increased research interest in these measures, they are seldom tested in concert, embedded in a theoretical background, and studied while accounting for interindividual differences in cognitive and attitudinal variables (Ziemer & Rothmund, 2022). The current study contributed to closing this gap. Informed by the HSM (Chaiken, 1980, 1987; Chaiken & Ledgerwood, 2012), we tested the effectiveness of two measures intended to target fake news: fact-checking warning labels intended to foster systematic processing (van der Linden, 2018) and the removal of social endorsement cues that

otherwise might benefit fake news via a bandwagon heuristic (Sundar et al., 2008). We evaluated these interventions on three central dependent variables, namely their ability to reduce the perceived credibility of fake news, participants' likelihood to amplify fake news, and participants' likelihood to engage in elaboration on fake news. We further controlled for interindividual differences reflecting the central role of cognitive and attitudinal variables for the processing of fake news and examined the role of age. To this end, we included measures for educational level, analytical thinking, involvement, and political attitude.

5.1 | Warning labels diminish fake news' effectiveness, but maybe not as they should

Warning labels were found to be effective in mitigating the perceived credibility (H1) as well as the likelihood to amplify (RQ1) fake news. As such, our findings are consistent with prior work demonstrating the effectiveness of warning labels (Clayton et al., 2020; Pennycook et al., 2020) and contradict studies that do not find such a benign effect (Lee et al., 2021; Oeldorf-Hirsch et al., 2020). Notably, the positive effect of warning labels on individuals' resilience towards fake news remained stable when we adjusted the analyses for both cognitive and attitudinal variables. This is an encouraging observation, as recent research in the highly polarized US context found that both Democrats and Republicans generally support warning labels over other forms of content moderation (Wihbey et al., 2021).

However, warning labels are intended to foster systematic processing and motivate users to click through to the information about the debunking (for a media report, see The Economic Times, 2021). In this study, we did not find a positive effect of warning labels on users' likelihood to elaborate on the fake news post. Consequentially, it might be that warning labels have become a heuristic cue themselves that do not foster careful engagement with social media content but are rather used as another rule of thumb for navigating social media. Although we cannot test whether warning labels are merely a heuristic cue in this study, there is evidence consistent with such a hypothesis. For instance, experimental studies have shown that attaching a warning label to fake news can create an "implied truth effect" among fake news without labels, such that they are perceived to be more credible (Pennycook et al., 2020). If warning labels motivated systematic processing, such an "implied truth effect" should be prevented by people's larger elaboration. Yet rerunning our analyses excluding $n = 32$ participants who stated that they would not use social media and were thus likely less familiar with warning labels did not change the effect of warning labels on credibility substantially (which would be expected if they are a learned heuristic). In any case, the effects of warning labels can diminish over time (Grady et al., 2021) and some studies failed to find an effect at all (e.g., Kreps & Kriner, 2022; Oeldorf-Hirsch et al., 2020), suggesting that warning labels might not trigger a long-lasting learning process. Consequentially, further research on the boundary conditions of the reported effects is needed. For instance, Kreps and

Kirner (2022) found the provision of additional accurate information to be more effective than a mere “false information” tag, possibly since additional information requires a more systematic elaboration than a mere warning label.

5.2 | Social endorsement cues do not matter, but that could be good

Social endorsement cues had no effect on the processing of fake news and did not interact with warning labels on any of our dependent measures (H2, RQ2 - RQ4). This contradicts prior studies that showed that a high (vs. low) number of social endorsement cues increases perceived credibility (Ali et al., 2022; Shin et al., 2022) and engagement intentions (Kim, 2021). As such, our study raises the question of under which conditions the bandwagon heuristic (Sundar et al., 2008) is utilized and contributes to the mounting evidence for the irrelevance of social endorsement cues for the perceived credibility of online content (Kluck et al., 2019; Lee et al., 2021; Mena et al., 2020; Molina et al., 2022). One potential reason for the absence of an effect observed in our study is that social endorsement cues have lost their heuristic value. There was a large election in Germany (parliamentary election 2017) before the time of our data collection that was accompanied by concerns over the influence of automated user accounts, so-called social bots, manipulating online content at scale (Ferrara et al., 2016; Kupferschmidt, 2017). Social warranting theory (for an overview, see DeAndrea, 2014), suggests that only digital cues perceived to be beyond the control of the author of a social media post are perceived as social endorsement. It could be that high like counts lost their warranting potential. Support for this speculation comes from a recent study finding a negative association between the number of likes and perceived credibility of social media news (Lee et al., 2021), as a very high number of likes might trigger feelings of unease (Molina et al., 2022). This assumption is also consistent with a recent representative survey showing that 72% of Germans no longer consider a high number of likes and shares as a good cue for evaluating the trustworthiness of a news item (Meßmer et al., 2021).

5.3 | Interindividual differences in cognitive and attitudinal variables matter

Cognitive variables, namely a higher educational level and more analytic thinking were associated with a lower likelihood to amplify fake news stories by linking or sharing them. However, only involvement was associated with the self-reported likelihood to elaborate on the topic. As such our results are consistent with the assumption that beliefs about fake news and the propensity to share it might be driven by different processes (Pennycook & Rand, 2021) and extend this notion towards the elaboration of fake news (Petty et al., 1983).

In line with studies arguing for the role of motivated processing, particularly a higher susceptibility to attitude-consistent fake news (Corbu & Negrea-Busuioac, 2020; Faragó et al., 2019), participants with more left-wing political attitudes were more likely to rate the fake news post exaggerating the outfall of climate change as being credible, reported a higher likelihood to amplify them, and were more likely to elaborate on them. Thus, our study contradicts prior work speaking for ideological asymmetries and a generally higher susceptibility of right-wing leaning individuals for fake news (Arendt et al., 2019; Jost et al., 2018). Furthermore, our findings underscore the need to account for both cognitive and attitudinal variables in research on fake news and countermeasures (Ziemer & Rothmund, 2022). Future research examining the interplay between these variables and the effectiveness of countermeasures in greater detail thus seems to be a valuable endeavor.

5.4 | Limitations

Our study has several limitations that must be considered when interpreting the results and generalizing our findings. First, we used a single fake news post as stimulus. Although we selected our stimulus based on theoretical considerations as well as a pretest with 173 participants, and single stimuli are common in fake news research (e.g., Lin et al., 2016; Mena et al., 2020), replicating our findings with other fake news topics is needed to evaluate the generalizability of our findings. Relatedly, our selected post had comparatively little emotional valence compared to other fake news stories that are often controversial and emotional (Arendt et al., 2019; Frischlich et al., 2021; Torabi Asr, 2019). This could be one reason why our observed effects are rather small. Although the potential large reach of fake news (Vosoughi et al., 2018) makes even small effects meaningful, participants' engagement intentions were low overall. This observation is in line with prior research indicating that overall sharing of misinformation is rare (Pennycook et al., 2021). Nevertheless, future research is needed to examine how other variables such as a posts' emotionality affect the variables studied here.

Second, we used a convenience sample of German participants. Although our sample reflected the overall Facebook population, women and students were overrepresented. As such, results are not easily generalizable to entire populations nor other countries, particularly because Germans are generally more concerned about climate change than other nationalities (Barasi et al., 2017). Thus, a replication of our findings in more representative and ideally international samples would be desirable.

Third, our mock-up Facebook feed offered a very limited functionality compared to a real Facebook usage scenario. Participants could scroll through our feed but not interact with any of the content. To gain more realistic insights, researchers could use digital traces of real Facebook behavior or observe participants in a laboratory (e.g., accompanied by eye tracking) while they interact with their personal timeline.

Finally, we used a single-item measure to assess credibility. Although single item measures can reliably capture homogenous constructs (Ang & Eisend, 2018; Loo, 2002; Slater & Rouner, 1996), research in the related field of media trust has identified different facets of news credibility (Kohring & Matthes, 2007), such as accuracy and completeness perceptions. Future research exploring these facets in greater detail would provide meaningful insights into the nuanced effects of fake news interventions.

6 | CONCLUSION AND PRACTICAL IMPLICATIONS

In the present study, we demonstrated the value of employing the HSM as a theoretical framework to study fake news interventions and tested the interaction between warning labels intended to foster systematic processing and the removal of social endorsement cues. Furthermore, we included several variables related to cognitive and attitudinal interindividual differences in our analyses. Of practical relevance, we demonstrated that current efforts by Facebook to allow users to remove social endorsement cues might have a limited effect on fighting fake news (Rizoiu, 2019), whereas warning labels were effective even when we adjusted for these interindividual differences. Notably, this effectiveness did not seem to be driven by warning labels' ability to motivate systematic processing. Thus, warning labels should be accompanied by preventive measures targeting systematic processing more directly: for example, by fostering users' media literacy (Moore & Hancock, 2022) and "inoculating" them against common fake news techniques (Lewandowsky & van der Linden, 2021). These "prebunking" techniques have been found to be highly effective in fostering individual resilience in the short- and long-term (Maertens et al., 2021; Roozenbeek & van der Linden, 2019).

ACKNOWLEDGMENTS

We thank Carina Kemmer and Susanne Grundler for their substantial support in the early stage of this project, particularly in conceptualization and data collection. We thank Regina Rockinger for her support with literature research. We thank David Goretzko and Florian Pargent for their statistical advice. The second author was supported by the Digital Society research program funded by the Ministry of Culture and Science of the German State of North Rhine-Westphalia under the number 005-1709-001, project: 1706dgn001. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data and code that support the findings of this study are openly available in the Open Science Framework at <https://osf.io/ugt2v/>.

ORCID

Timo K. Koch  <http://orcid.org/0000-0001-6728-2063>

Lena Frischlich  <https://orcid.org/0000-0001-5039-5301>

Eva Lerner  <https://orcid.org/0000-0002-6600-9580>

REFERENCES

- Ali, K., Li, C., Zain-ul-abdin, K., & Muqtadir, S. A. (2022). The effects of emotions, individual attitudes towards vaccination, and social endorsements on perceived fake news credibility and sharing motivations. *Computers in Human Behavior*, 134, 107307. <https://doi.org/10.1016/j.chb.2022.107307>
- Ali, K., Li, C., Zain-ul-abdin, K., & Zaffar, M. A. (2021). Fake news on facebook: Examining the impact of heuristic cues on perceived credibility and sharing intention. *Internet Research*, 32(1), 379–397. <https://doi.org/10.1108/INTR-10-2019-0442>
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Amazeen, M. A. (2021). Resisting covert persuasion in digital news: Comparing inoculation and reactance in the processing of native advertising disclosures and in article engagement intentions. *Journalism & Mass Communication Quarterly*, 98(4), 1129–1156. <https://doi.org/10.1177/1077699020952131>
- Ang, L., & Eisend, M. (2018). Single versus multiple measurement of attitudes: A meta-analysis of advertising studies validates the single-item measure approach. *Journal of Advertising Research*, 58(2), 218–227. <https://doi.org/10.2501/JAR-2017-001>
- Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, 93(1), 59–79. <https://doi.org/10.1177/1077699015606057>
- Arendt, F., Haim, M., & Beck, J. (2019). Fake news, Warnhinweise und perzipierter Wahrheitsgehalt: Zur unterschiedlichen Anfälligkeit für Falschmeldungen in Abhängigkeit von der politischen Orientierung. [Fake news, warnings and perceptual truth: On the differential susceptibility for fake news depending on recipients' political orientations]. *Publizistik*, 64, 181–204. <https://doi.org/10.1007/s11616-019-00484-4>
- Avram, M., Micallef, N., Patil, S., & Menczer, F. (2020). Exposure to social engagement metrics increases vulnerability to misinformation. *Harvard Kennedy School Misinformation Review*, 1(5). <https://doi.org/10.37016/mr-2020-033>
- Bader, K., Jansen, C., & Rinsdorf, L. (2020). Jenseits der Fakten: Deutschsprachige Fake News aus Sicht der Journalistik [Beyond facts: Journalist perspectives on fake news in German language]. *Desinformationen aufdecken und bekämpfen. Interdisziplinäre Ansätze gegen Desinformationskampagnen und für Meinungspluralität*, 33–76. <https://doi.org/10.5771/9783748904816-33>
- Bak-Coleman, J. B., Kennedy, I., Wack, M., Beers, A., Schafer, J. S., Spiro, E. S., Starbird, K., & West, J. D. (2022). Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour*, 6, 1372–1380. <https://doi.org/10.1038/s41562-022-01388-6>
- Barasi, L., Harding, R., & Dunatchik, A. (2017). Climate concern and pessimism: Examining public attitudes across Europe. *National Centre for Social Research*. <http://Natcen.Ac.Uk/Media/1513272/ESS-Climate-Change-Report-FINAL.Pdf>
- Boberg, S., Quandt, T., Schatto-Eckrodt, T., & Frischlich, L. (2020). Pandemic populism: Facebook pages of alternative news media and the corona crisis—a computational content analysis. *ArXiv:2004.02566 [Cs]*. <http://arxiv.org/abs/2004.02566>
- Bryanov, K., & Vziatysheva, V. (2021). Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news. *PLoS One*, 16(6), e0253717. <https://doi.org/10.1371/journal.pone.0253717>

- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>
- Chaiken, S. (1987). The heuristic model of persuasion. In C. P. Zanna, M. P. Olson, & J. M. Herman (Eds.), *Social influence: The ontario symposium* (5th ed., pp. 3–39). Erlbaum.
- Chaiken, S., & Ledgerwood, A. (2012). A theory of heuristic and systematic information processing. In P. Van Lange, A. Kruglanski, & E. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 246–266). SAGE Publications Ltd. <https://doi.org/10.4135/9781446249215>
- Chen, S., Duckworth, K., & Chaiken, S. (1999). Motivated heuristic and systematic processing. *Psychological Inquiry*, 10(1), 44–49. https://doi.org/10.1207/s15327965pli1001_6
- Christensen, R. H. B. (2019). Ordinal—regression models for ordinal data. *R-Package Version 2019.12-10*.
- Clayton, K., Blair, S., Busam, J. A., Forstner, S., Gance, J., Green, G., Kawata, A., Kovvuri, A., Martin, J., Morgan, E., Sandhu, M., Sang, R., Scholz-Bright, R., Welch, A. T., Wolff, A. G., Zhou, A., & Nyhan, B. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior*, 42, 1073–1095. <https://doi.org/10.1007/s11109-019-09533-0>
- Constantine, J. (2019). Now facebook says it may remove like counts. TechCrunch. <https://social.techcrunch.com/2019/09/02/facebook-hidden-likes/>
- Corbu, N., & Negrea-Busiuc, E. (2020). Populism meets fake news: Social media, stereotypes and emotions. In B. Krämer, & C. Holtz-Bacha (Eds.), *Perspectives on Populism and the Media* (pp. 181–200). Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783845297392-181>
- Curry, A. L., & Stroud, N. J. (2021). The effects of journalistic transparency on credibility assessments and engagement intentions. *Journalism*, 22(4), 901–918. <https://doi.org/10.1177/1464884919850387>
- DeAndrea, D. C. (2014). Advancing warranting theory: Advancing warranting theory. *Communication Theory*, 24(2), 186–204. <https://doi.org/10.1111/comt.12033>
- Donovan, J. (2020). Social-media companies must flatten the curve of misinformation. *Nature*. <https://doi.org/10.1038/d41586-020-01107-z>.
- Duncan, M. (2020). What's in a label? Negative credibility labels in partisan news. *Journalism & Mass Communication Quarterly*, 99(2), 390–413. <https://doi.org/10.1177/1077699020961856>
- European Commission. (2018). *Tackling online disinformation: Commission proposes an EU-wide Code of Practice* [Text]. European Commission. https://ec.europa.eu/commission/presscorner/detail/en/IP_18_3370
- Facebook. (2020). *An update on our work to keep people informed and limit misinformation about covid-19*. About Facebook. <https://about.fb.com/news/2020/04/covid-19-misinfo-update/>
- Faragó, L., Kende, A., & Kreko, P. (2019). We only believe in news that we doctored ourselves: The connection between partisanship and political fake news. *Social Psychology*, 51, 1–14. <https://doi.org/10.1027/1864-9335/a000391>
- Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7), 96–104. <https://doi.org/10.1145/2818717>
- Flanagin, A. J., Winter, S., & Metzger, M. J. (2020). Making sense of credibility in complex information environments: The role of message sidedness, information source, and thinking styles in credibility evaluation online. *Information, Communication & Society*, 23(7), 1038–1056. <https://doi.org/10.1080/1369118X.2018.1547411>
- Freiling, I. (2019). Detecting misinformation in online social networks: A think-aloud study on user strategies. *Studies in Communication and Media*, 8(4), 471–496. <https://doi.org/10.5771/2192-4007-2019-4-471>
- Frischlich, L., Hellmann, J. H., Brinkschulte, F., Becker, M., & Back, M. D. (2021). Right-wing authoritarianism, conspiracy mentality, and susceptibility to distorted alternative news. *Social Influence*, 16(1), 24–64. <https://doi.org/10.1080/15534510.2021.1966499>
- Grady, R. H., Ditto, P. H., & Loftus, E. F. (2021). Nevertheless, partisanship persisted: Fake news warnings help briefly, but bias returns with time. *Cognitive Research: Principles and Implications*, 6(1), 52. <https://doi.org/10.1186/s41235-021-00315-z>
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science Advances*, 5(1), eaau4586. <https://doi.org/10.1126/sciadv.aau4586>
- Halpern, D., Valenzuela, S., Katz, J., & Miranda, J. P. (2019). From belief in conspiracy theories to trust in others: Which factors influence exposure, believing and sharing fake news. In G. Meiselwitz (Ed.), *Social computing and social media. design, human behavior and analytics* (Vol. 11578, pp. 217–232). Springer International Publishing. https://doi.org/10.1007/978-3-030-21902-4_16
- Hameleers, M. (2022). Separating truth from lies: Comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands. *Information, Communication & Society*, 25(1), 110–126.
- Hayes, R. A., Carr, C. T., & Wahn, D. Y. (2016). One click, many meanings: Interpreting paralinguistic digital affordances in social media. *Journal of Broadcasting & Electronic Media*, 60(1), 171–187. <https://doi.org/10.1080/08838151.2015.1127248>
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.
- Humprecht, E. (2019). Where 'fake news' flourishes: A comparison across four Western democracies. *Information, Communication & Society*, 22(13), 1973–1988. <https://doi.org/10.1080/1369118X.2018.1474241>
- Instagram. (2019). *Combating Misinformation on Instagram*. <https://about.instagram.com/blog/announcements/combating-misinformation-on-instagram>
- Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, 106(2), 290–314. <https://doi.org/10.1037/0033-2909.106.2.290>
- Jost, J. T., van der Linden, S., Panagopoulos, C., & Hardin, C. D. (2018). Ideological asymmetries in conformity, desire for shared reality, and the spread of misinformation. *Current Opinion in Psychology*, 23, 77–83. <https://doi.org/10.1016/j.copsyc.2018.01.003>
- Kahneman, D. (2013). *Thinking fast and slow*. Farrar, Strauß, and Giroux.
- Kaye, B. K., & Johnson, T. J. (2021). “Newstrusting” or “newsbusting?” Heuristic and systematic information processing and trust in media. *Atlantic Journal of Communication*, 29(5), 312–327. <https://doi.org/10.1080/15456870.2021.1869004>
- Kim, J. (2021). The meaning of numbers: Effect of social media engagement metrics in risk communication. *Communication Studies*, 72(2), 195–213. <https://doi.org/10.1080/10510974.2020.1819842>
- Kluck, J. P., Schaewitz, L., & Krämer, N. C. (2019). Doubters are more convincing than advocates. *The impact of user comments and ratings on credibility perceptions of false news stories on social media*. *Studies in Communication and Media*, 8(4), 446–470. <https://doi.org/10.5771/2192-4007-2019-4-446>
- Kohring, M., & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research*, 34(2), 231–252. <https://doi.org/10.1177/0093650206298071>
- Kreps, S. E., & Kriner, D. L. (2022). The COVID-19 infodemic and the efficacy of interventions intended to reduce misinformation. *Public Opinion Quarterly*, 86(1), 162–175. <https://doi.org/10.1093/poq/nfab075>

- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- Kupferschmidt, K. (2017). Social media “bots” tried to influence the U.S. election. Germany may be next. *Science*. <https://doi.org/10.1126/science.aap9514>
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lee, S., Liang, F., Hahn, L., Lane, D. S., Weeks, B. E., & Kwak, N. (2021). The impact of social endorsement cues and manipulability concerns on perceptions of news credibility. *Cyberpsychology, Behavior and Social Networking*, 24, 384–389. <https://doi.org/10.1089/cyber.2020.0566>
- Lewandowsky, S., & van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, 32, 348–384. <https://doi.org/10.1080/10463283.2021.1876983>
- Lin, X., Spence, P. R., & Lachlan, K. A. (2016). Social media and credibility indicators: The effect of influence cues. *Computers in Human Behavior*, 63, 264–271. <https://doi.org/10.1016/j.chb.2016.05.002>
- van der Linden, S. (2018). The future of behavioral insights: On the importance of socially situated nudges. *Behavioural Public Policy*, 2(2), 207–217. <https://doi.org/10.1017/bpp.2018.22>
- Loo, R. (2002). A caveat on using single-item versus multiple-item scales. *Journal of Managerial Psychology*, 17(1), 68–75. <https://doi.org/10.1108/02683940210415933>
- Luo, M., Hancock, J. T., & Markowitz, D. M. (2022). Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research*, 49(2), 171–195. <https://doi.org/10.1177/0093650220921321>
- Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology: Applied*, 27(1), 1–16. <https://doi.org/10.1037/xap0000315>
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data & Society Research Institute. <https://datasociety.net/library/media-manipulation-and-disinfo-online/>
- Mena, P. (2020). Cleaning up social media: The effect of warning labels on likelihood of sharing false news on facebook. *Policy & Internet*, 12(2), 165–183. <https://doi.org/10.1002/poi3.214>
- Mena, P., Barbe, D., & Chan-Olmsted, S. (2020). Misinformation on instagram: The impact of trusted endorsements on message credibility. *Social Media + Society*, 6(2), 2056305120935102. <https://doi.org/10.1177/2056305120935102>
- Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220. <https://doi.org/10.1016/j.pragma.2013.07.012>
- Meißner, A.-K., Sänglerlaub, A., & Schulz, L. (2021). *Quelle: Internet? Digitale Nachrichten- und Informationskompetenzen der deutschen Bevölkerung im Test*. Stiftung neue Verantwortung. https://www.stiftung-nv.de/sites/default/files/studie_quelleinternet.pdf
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <https://doi.org/10.1037/h0043158>
- Molina, M. D., Wang, J., Sundar, S. S., Le, T., & DiRusso, C. (2022). Reading, commenting and sharing of fake news: How online bandwagons and bots dictate user engagement. *Communication Research*, 00936502211073398. <https://doi.org/10.1177/00936502211073398>
- Moore, R. C., & Hancock, J. T. (2022). A digital media literacy intervention for older adults improves resilience to fake news. *Scientific Reports*, 12(1), 6008. <https://doi.org/10.1038/s41598-022-08437-0>
- Mosseri, A. (2016). Addressing hoaxes and fake news. *About Facebook*. <https://about.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/>
- Norris, P., & Epstein, S. (2011). An experiential thinking style: Its facets and relations with objective and subjective criterion measures. *Journal of Personality*, 79(5), 1043–1080. <https://doi.org/10.1111/j.1467-6494.2011.00718.x>
- Oeldorf-Hirsch, A., Schmierbach, M., Appelman, A., & Boyle, M. P. (2020). The ineffectiveness of fact-checking labels on news memes and articles. *Mass Communication and Society*, 23(5), 682–704. <https://doi.org/10.1080/15205436.2020.1733613>
- Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, 66(11), 4944–4957. <https://doi.org/10.1287/mnsc.2019.3478>
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), 590–595. <https://doi.org/10.1038/s41586-021-03344-2>
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting covid-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770–780. <https://doi.org/10.1177/0956797620939054>
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. <https://doi.org/10.1016/j.cognition.2018.06.011>
- Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, 88(2), 185–200. <https://doi.org/10.1111/jopy.12476>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, 10(2), 135–146. <https://doi.org/10.2307/2488919>
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34(2), 243–281. <https://doi.org/10.1111/j.1559-1816.2004.tb02547.x>
- Preston, S., Anderson, A., Robertson, D. J., Shephard, M. P., & Huhe, N. (2021). Detecting fake news on facebook: The role of emotional intelligence. *PLoS One*, 16(3), e0246757. <https://doi.org/10.1371/journal.pone.0246757>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org>
- Reuters. (2022). Twitter rolls back COVID misinformation policy. *Reuters*. <https://www.reuters.com/technology/twitter-rolls-back-covid-misinformation-policy-2022-11-29/>
- Revelle, W. R. (2017). *Psych: Procedures for personality and psychological research*.
- Ribeiro, F. N., Benevenuto, F., & Zagheni, E. (2020). How biased is the population of facebook users? Comparing the demographics of facebook users with census data to generate correction factors. *ArXiv:2005.08065 [Cs]*. <http://arxiv.org/abs/2005.08065>
- Rizoio, M.-A. (2019). *Can hiding likes make Facebook fairer and rein in fake news? The science says maybe*. *The Conversation*. <http://>

- theconversation.com/can-hiding-likes-make-facebook-fairer-and-rein-in-fake-news-the-science-says-maybe-124671
- Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, 5(1), 65. <https://doi.org/10.1057/s41599-019-0279-9>
- Roth, J., & Pickles, N. (2020). Updating our approach to misleading information. May 11. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html
- Shin, I., Wang, L., & Lu, Y.-T. (2022). Twitter and endorsed (fake) news: The influence of endorsement by strong ties, celebrities, and a user majority on credibility of fake news during the covid-19 pandemic. *International Journal of Communication*, 16, 2573-2595.
- Slater, M., & Rouner, D. (1996). How message evaluation and source attributes may influence credibility assessment and belief change. *Journalism & Mass Communication Quarterly*, 73(4), 974-991. <https://doi.org/10.1177/107769909607300415>
- Statista. (2021). Facebook: Users by age and gender. *Distribution of facebook users worldwide as of July 2021, by age and gender*. <https://www.statista.com/statistics/376128/facebook-global-user-age-distribution/>
- Sundar, S. S., Molina, M. D., & Cho, E. (2021). Seeing is believing: Is video modality more powerful in spreading fake news via online messaging apps? *Journal of Computer-Mediated Communication*, 26(6), 301-319. <https://doi.org/10.1093/jcmc/zmab010>
- Sundar, S. S., Oeldorf-Hirsch, A., & Xu, Q. (2008). The bandwagon effect of collaborative filtering technology. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, 3453-3458. <https://doi.org/10.1145/1358628.1358873>
- Tandoc, E. C., & Kim, H. K. (2022). Avoiding real news, believing in fake news? Investigating pathways from information overload to misbelief. *Journalism*. <https://doi.org/10.1177/14648849221090744>
- Tandoc, E. C., Ling, R., Westlund, O., Duffy, A., Goh, D., & Zheng Wei, L. (2018). Audiences' acts of authentication in the age of fake news: A conceptual framework. *New Media & Society*, 20(8), 2745-2763.
- The Economic Times. (2021). Twitter redesigns warning labels to help users identify fake news, misinformation easily. *The Economic Times*. [https://economictimes.indiatimes.com/magazines/panache/](https://economictimes.indiatimes.com/magazines/panache/twitter-redesigns-warning-labels-to-identify-misinformation-fake-news-easily/articleshow/87749680.cms?from=mdr)
- [twitter-redesigns-warning-labels-to-identify-misinformation-fake-news-easily/articleshow/87749680.cms?from=mdr](https://twitter.com/redesigns-warning-labels-to-identify-misinformation-fake-news-easily/articleshow/87749680.cms?from=mdr)
- Torabi Asr, F. (2019). One potential route to flagging fake news at scale: Linguistic analysis. *Nieman Lab*. <https://www.niemanlab.org/2019/08/one-potential-route-to-flagging-fake-news-at-scale-linguistic-analysis/>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Wardle, C. (2018). The need for smarter definitions and practical, timely empirical research on information disorder. *Digital Journalism*, 6(8), 951-963. <https://doi.org/10.1080/21670811.2018.1502047>
- Wihbey, J., Morrow, G., Chung, M., & Peacey, M. (2021). *The bipartisan case for labeling as a content moderation method: Findings from a national survey* (Ethics Institute Working Paper). Northeastern University. <https://www.ssrn.com/abstract=3923905>
- Winter, S., Brückner, C., & Krämer, N. C. (2015). They came, they liked, they commented: Social influence on facebook news channels. *Cyberpsychology, Behavior and Social Networking*, 18(8), 431-436. <https://doi.org/10.1089/cyber.2015.0005>
- Ziemer, C.-T., & Rothmund, T. (2022). Psychological underpinnings of disinformation countermeasures: A systematic scoping review. <https://doi.org/10.31234/osf.io/scq5v>
- Zimmermann, F., & Kohring, M. (2018). "Fake News" als aktuelle Desinformation. Systematische Bestimmung eines heterogenen Begriffs. *Medien & Kommunikationswissenschaft*, 66(4), 526-541. <https://doi.org/10.5771/1615-634X-2018-4-526>
- Zuckerman, A., & Chaiken, S. (1998). A Heuristic-systematic processing analysis of the effectiveness of product warning labels. *Psychology and Marketing*, 15(7), 621-642.

How to cite this article: Koch, T. K., Frischlich, L., & Lermer, E. (2023). Effects of fact-checking warning labels and social endorsement cues on climate change fake news credibility and engagement on social media. *Journal of Applied Social Psychology*, 53, 495-507. <https://doi.org/10.1111/jasp.12959>