EJHC | European Journal of Health Communication

ISSN 2673-5903
ejhc.org

Article *Special Issue on Mis- and Disinformation About Covid-19*

Stop and Think! Exploring the Role of News Truth Discernment, Information Literacy, and Impulsivity in the Effect of Critical Thinking Recommendations on Trust in Fake Covid-19 News

DOI: 10.47368/ejhc.2022.203
2022, Vol. 3(2) 40-63
CC BY 4.0

Jasmijn Kruijt
Graduate School of Communication, University of Amsterdam, The Netherlands

Corine S. Meppelink
Amsterdam School of Communication Research, University of Amsterdam, The Netherlands

Lisa Vandeberg
Behavioral Science Institute, Radboud University, The Netherlands

Abstract

Covid-19-related fake news widely circulates on social media. This is problematic as people commonly do not process information on social media in a very critical manner. Also, when people encounter particular online content several times this tends to increase the content's trustworthiness, sometimes irrespective of the accuracy of the provided information. Our study aims to explore whether, how, and for whom a simple critical thinking recommendation added to a social media newsfeed can aid people to better discern true news from fake news and reduce their trust in fake news. In an online experiment, 220 participants were exposed to a Twitter newsfeed with true and fake Covid-19-related news messages, either with or without critical thinking recommendations. The findings showed that participants who were exposed to the recommendations showed less trust in fake news messages, which was mediated by an increased accuracy in news truth discernment. Results showed no significant moderating effects of information literacy and impulsivity characteristics. Overall, the findings of this study are promising as this scalable, low-cost intervention might potentially help combat the effects of fake news on social media.

Corresponding author:
Jasmijn Kruijt: j.kruijt@uva.nl

Keywords
Fake news, trust, news truth discernment, critical thinking recommendations, information literacy.

Information about the Covid-19 pandemic is omnipresent on social media (Apuke & Omar, 2021; Rosenberg et al., 2020) and many people use social media as an information source for Covid-19-related news (Casero-Ripollés, 2020). News is expected to be accurate, objective, and truthful (Tandoc et al., 2018). This contrasts with *fake news*, which is defined as fabricated information that mimics news media's form and editorial norms but does not have truthful intent (Lazer et al., 2018). Fake news is a subcategory of *disinformation;* both concepts apply to inaccurate information that is intended to mislead. However, the layout of fake news mimics news outlets, which is not necessarily the case for disinformation (Wardle, 2018). Moreover, fake news is different from *misinformation*, which also applies to incorrect information but does not necessarily aim to mislead (Wardle, 2018). Misinformation can therefore be used as an umbrella term for "any information that turns out to be false" (Ecker et al., 2022). A serious share of the Covid-19-related social media messages can be classified as misinformation (Shahi et al., 2021) and many people encounter misleading Covid-19 news on a daily basis (Watson, 2021).

Research demonstrates that online misinformation, with instances of fake news and unproven vaccination standpoints in particular, more effectively reaches people than truthful, evidence-based news (Johnson et al., 2020; Lutkenhaus et al., 2019; Vosoughi et al., 2018). Even brief exposures to misinformation can lead to long-term false memories (Zhu et al., 2012). Moreover, people often consume fake news without recognizing it as such, and they might act upon these falsehoods accordingly (Rapp & Salovich, 2018). Relying on Covid-19 misinformation could therefore increase fear, societal disagreement, and support of conspiracy beliefs (Naeem et al., 2021), or stimulate ineffective treatments (e.g., drinking bleach; Reimann, 2020). On a broader scale, fake news can have severe societal consequences such as institutional distrust, polarization in politics, increased discrimination, or journalistic sensationalism (Dentith, 2016). Moreover, misinformation in general and fake news in particular pose a major risk for public health as they may harm vaccine uptake (Iacobucci, 2019) and compliance to preventive behaviours (Bridgman et al., 2020; Lee et al., 2020). As a consequence, fake news might endanger the health of individuals as well as the containment of the pandemic.

The prevalence and consequences of Covid-19-related fake news emphasize the importance to combat it. Therefore, the present study aims to foster people's ability to accurately distinguish true news from fake news by testing a simple recommendation to process information thoughtfully. Scholars have been searching for ways to limit the influence and spread of fake news. One solution can be offered by technology companies (e.g., Twitter, Facebook), using algorithms to filter fake news (Zhou et al., 2019). However, a solution that is solely dependent on technology might not be ideal, because algorithms have their limitations, such as the inability to define intentions, sarcasm, or metaphors (Cardoso Durier da Silva et al., 2019). Moreover, such a solution on the platform-side does not empower people to better recognize and cope with fake news that potentially reaches them through other channels. Alternatively, professional fact-checkers who publish objective evaluations of widely circulated claims could be effective in countering fake news (Pavleska et al., 2018). However,

fact-checkers need time and funding to verify news and debunking false beliefs in people's minds is difficult (see meta-analysis Chan et al., 2017). The present study therefore contributes to the trend that besides professional and technological solutions, a solution is sought at the audience level (Ha et al., 2021).

Recent studies show that people are less likely to believe and share fake news when they actively reflect on the veracity of news messages (Fazio, 2020; Salovich & Rapp, 2020) – regardless of whether the news is consistent with their ideology (Bronstein et al., 2019; Pennycook & Rand, 2021; Pennycook & Rand, 2019). For example, it has been shown that letting people explain how they determine whether a story is true or not reduces the intention to share false information (Fazio, 2020). Another study showed that when people are exposed to guidelines on how to evaluate online news reduces the likelihood to trust and share fake news (Lutzke et al., 2019). Despite their effectiveness, such evaluation instructions might be less suitable to directly incorporate in the social media feed, as they require explicit and detailed (step-by-step) deliberations by the news receiver. Another line of research focuses on *accuracy nudges*, which are more easily implementable. This work consistently shows that subtly inducing people to reflect on the accuracy of the news they encounter online – without pointing to specific message elements – improves the level of truth discernment and reduces the sharing of inaccurate information (e.g., Pennycook et al., 2021; Pennycook et al., 2020). Our study adds to the research on accuracy nudges by evaluating the effectiveness of a single-sentence recommendation advocating the thoughtful evaluation of *news content* and *source*. These are important indicators of news accuracy but are frequently overlooked or ignored by users (Ecker et al., 2022). Moreover, prior studies mainly focused on the effectiveness of interventions in terms of trustworthiness or sharing intentions (Fazio, 2020; Lutzke et al., 2019; Pennycook et al., 2020). The current study aims to identify a potential mechanism by assessing whether the instruction to process news thoughtfully improves people's ability to discern true news from fake news, resulting in lower trust in fake news. Additionally, this study aims to identify potential boundary conditions by exploring the role of information literacy and trait impulsivity on the effectiveness of interventions fighting fake news. Finally, as most fake news literature focuses on political news in the U.S. (Ha et al., 2021; Mele et al., 2017), this study extends the literature by focusing on the impact of Covid-19-related fake news in the Netherlands.

The following research question is central to our study: *What is the effect of a critical thinking recommendation on an individual's ability to discern true from fake news on social media, and (how) does this relate to distrust in fake news?*

Theoretical Framework

Shallow Versus Deep Information Processing

Dual-process and dual-systems models serve as a basis for the hypotheses (i.e., the dual systems approach by Kahneman, 2011; Elaboration Likelihood Model by Petty & Cacioppo, 1986, Associative-Propositional Evaluation model by Gawronski & Bodenhausen, 2006; Reflective-Impulsive Model by Strack & Deutsch, 2004; Heuristic-Systematic Processing Model as applied in Zuckerman & Chaiken, 1998). These approaches argue that human information processing, as well as attitude formation and decision-making, are affected by two qualitatively different mental processes or systems. Overall, the first system entails more automatic, fast, course-grained, low-effort, associative, peripheral, or subconscious processes that more

heavily rely on, for instance, reflexes, impulses, and shallow information processing. The second system, on the other hand, relies more on intentional, slow, fine-grained, effortful, propositional, central, or conscious processes that are characterized by ratio, elaboration, and deeper information processing. Generally speaking, people tend to avoid demanding cognitive processes and, instead, often process information efficiently using the first system (Kahneman, 2011). Our hypotheses are based on the assumption that news on social media is by default likely processed in this low-effort - and thereby less critical – manner (cf. Boczkowski et al., 2018). However, prompting people to stop and think about the source and content of the news would stimulate them to process the news more deeply and critically. Indeed, using the second, more effortful system while processing news information is positively related to correctly assessing news quality (Pennycook & Rand, 2019).

The Effect of Critical Thinking Recommendations on Trust in Fake News

Different studies have demonstrated the effect of messages inducing critical processing on cognitive, affective, and behavioural outcomes. For example, ample research on disclosures and forewarnings has shown that warnings that explicitly state the persuasive nature of marketing messages enhance people's ability to recognize these messages as such, which negatively impacts people's attitude towards these messages (e.g., Van Reijmersdal et al., 2016). There are many ways in which people can be warned, varying from adding disclosures to televised content such as 'this program contains product placement' (Spielvogel et al., 2021) to instructions on how to evaluate news (Lutzke et al., 2019). In the context of climate change, research has shown that instructions on how to evaluate news reduces trust in fake news as well as people's intentions to like and share fake news (Lutzke et al., 2019). Another study found that people who were asked to reflect on the accuracy of unrelated news (versus people who were not asked) tended to believe and share information of higher quality (Pennycook et al., 2020). This suggests that the effect of warnings or accuracy nudges depends (at least in part) on stimulating people to critically process the presented information.

The Mediating Role of Discernment Accuracy

Research suggests that a recommendation to evaluate the veracity of online news or to reflect on its accuracy can reduce the trust in fake news as well as sharing intentions (e.g., Lutzke et al., 2019; Fazio, 2020; Pennycook et al., 2020). It is possible that an instruction to critically evaluate information prompts people to process a message with more effort, resulting in an improved ability to discriminate between true and fake messages. Once messages are recognized as fake, this will likely reduce people's trust in those messages. This explanation pinpoints a shift in processing mode (from heuristic to systematic) due to the recommendation. Research related to the Dual Processing Model of Credibility Assessment (Metzger, 2007) has shown that people frequently judge the credibility of websites based on heuristic cues, such as consistency (e.g., Metzger et al., 2010). This indicates that people do not necessarily evaluate online information in a critical manner, leaving room for instruction to potentially change the way in which information is processed. A study by Pennycook and Rand (2021) indeed indicates that a lack of careful reasoning is associated with poor truth discernment of news messages. Also, in the context of fake claims, it has been shown that explicit instructions to engage in evaluative thinking (by prompting people to act as fact-checkers) help people identify inaccuracies in information (Brashier et al, 2020). Furthermore, research shows that correcting

misperceptions by exposing people to science communication is most effective when they are motivated to accurately process the presented information (Van Stekelenburg et al., 2020).

Based on the above, we hypothesize that prompting people to critically process news messages on social media enhances their ability to accurately discern true news from fake news, which in turn leads to a greater distrust in fake news. Formally:

> *H1:* People exposed to news messages with critical thinking recommendations have less trust in fake news than people exposed to news messages without such recommendations; this effect is mediated by truth discernment.

The Moderating Role of Personal Characteristics

Little research investigated the boundary conditions of critical thinking recommendations, especially in the context of fake news. Scheufele and Krause (2019) argue that some people find it more difficult to identify misinformation due to lower literacy skills. Therefore, this study explores the conditional effect of *information literacy*. Information literacy is defined as a set of skills to locate, understand, evaluate, and effectively apply (online) information (Williams & Wavell, 2007). Individuals with higher levels of information literacy are better able to critically evaluate online information (Jones-Jang et al., 2021), to verify information (Khan & Idris, 2019), and to identify misinformation (Scheufele & Krause, 2019). However, the literature review by De Paor and Heravi (2020) suggests that information literacy skills are only beneficial in flagging falsehoods when people are reminded of potential falsehoods. A recommendation as used in the present study would therefore prompt people to apply their information literacy skills. This idea is supported by a study demonstrating that a recommendation prompting people to evaluate online health information using established criteria increases their ability to discriminate between high and low quality health websites, but only for people with high levels of health literacy (Diviani & Meppelink, 2017). Hence, the following hypothesis is proposed:

> *H2:* The mediation effect of critical thinking recommendations on distrust in fake news through news truth discernment (H1) is moderated by information literacy, such that the observed effect is stronger for people with higher (versus lower) information literacy levels.

Next to information literacy, this study explores the influence of *impulsivity*, which provides a relevant construct in the fast-paced and superficial context in which social media messages are being consumed and processed. Impulsivity is a personality trait reflecting a cognitive style in which people lack the willingness to think (Coutlee et al., 2014). Impulsive people are characterized by reflexive (as opposed to reflective) thinking, causing them to make fast and sometimes inaccurate decisions (Baron, 2019; Messer, 1976). Reflective people, on the other hand, gather information and evaluate their actions more carefully (Messer, 1976). This active, open-minded thinking allows people to better evaluate information (Baron, 2019). In the context of fake news, it has been shown that people who prefer to reflect critically compared to those who prefer to think impulsively are better able to distinguish true news from fake news (Batailler et al., 2021) and judge fake news as less accurate (Pennycook & Rand, 2020).

Furthermore, impulsivity as a trait affects how people shift between the dual states of information processing. It is argued that people with high levels of impulsivity are less likely to switch to reflective, effortful processing states than people with low impulsivity (Strack & Deutsch, 2006). That is, impulsive people are more likely to evaluate information based on gut

feelings in an impulsive, low-effort processing state. The effectiveness of the critical thinking recommendation in the present study is therefore likely to vary for people with different levels of impulsivity. Specifically, impulsive people are likely less sensitive to critical thinking recommendations, resulting in less accurate news truth discernment and more trust in fake news compared to more reflective people. Altogether, we hypothesize that the effect of the recommendation on news truth discernment may be weaker for more impulsive people. This results in the following hypothesis:

> *H3:* The mediation effect of critical thinking recommendations on distrust in fake news through news truth discernment (H1) is moderated by impulsivity, such that the observed effect is weaker for people with higher (versus lower) impulsivity levels.

Methods

Design and Participants

The experiment consisted of a one factorial between-subjects design (critical thinking recommendation: absent vs present) with two quasi-experimental moderating factors (information literacy and impulsivity). Participants were presented with a newsfeed containing four fake news messages and three true news messages and were randomly assigned to the condition in which the news messages were accompanied by the critical thinking recommendation (experimental condition) or not (control condition). The study was approved by the Ethical Committee of the Amsterdam School of Communication Research, University of Amsterdam (nr: 2021-PC-13364).

Between April 22nd and May 9th, 2021, a total of 220 Dutch speaking participants were recruited using the lab website of a Dutch university as well as the researchers' social networks. Student participants (n = 186) were rewarded with course credits whereas other participants (n = 34) could win a €25.- gift card. Some respondents were excluded from analyses as they did not complete the survey (n = 11), were under the age of 16 (n = 1), or never used social media (n = 1). The final sample consisted of 105 participants in the experimental condition and 102 participants in the control condition (N = 207).[1]

Stimuli

Seven news messages were selected, representing various sub-topics related to Covid-19 (e.g., prominent people, biological processes, vaccinations). Three of these stories were obtained via the fact-check agency Factcheck.Vlaanderen, which is part of Poynter - The International Fact-Checking Network (two fake and one true news message). In addition, four fact-checked messages were derived from a study by Pennycook et al. (2020) about Covid-19 fake news (two fake and two true news messages). All news messages circulated on social media between the end of 2020 and the beginning of 2021. The headlines of the messages are presented in Table 1. Twitter was chosen as the channel because much misinformation about Covid-19 is shared via this platform (Rosenberg et al., 2020). Messages consisted of a headline complemented with the original image and the first sentence of the news message, to reflect the authentic interface of Twitter. This minimal presentation is sufficient for people to misperceive fake news as accurate (Allcott & Gentzkow, 2017). The messages were combined into a Twitter newsfeed optimized for smartphones and computers to assure a realistic experimental setting, benefiting ecological validity. To control the influence of dates, likes,

and shares, these numbers were kept equal among all messages. All study materials are available on OSF (https://osf.io/sbrf6/).

In the experimental condition, the following single-sentence critical thinking recommendation was added to every news message: "Stop and think before trusting and sharing online information: evaluate the content and the source!" (formulated in Dutch). The content of this recommendation was inspired by the intervention guidelines of Lutzke et al. (2019), who successfully stimulated individuals to stop and think before trusting news. In turn, their guidelines reflect the recommendations of the International Federation of Library Associations and Institutions and the Facebook Help Center (Lutzke et al., 2019).

Table 1. Headlines of the News Messages and the Sources Presented with the Messages

	Dutch original version	English translation	Source
True news messages	Trump heeft overheidsinstanties die verantwoordelijk zijn voor het reguleren van de coronavirus uitbraak de afgelopen 2 jaar gekort	Trump spent the past 2 years slashing the government agencies responsible for handling the coronavirus outbreak	Business Insider
	Britse premier Johnson geprikt met Astra Zeneca-vaccin	British prime minister Johnson receives Astra Zeneca jab	RD Opinie
	Waarom vliegveldscreening de verspreiding van het coronavirus niet tegengaat	Why airport screening won't stop the spread of coronavirus	Science Magazine
Fake news messages	Meest gelezen Britse medische auteur geeft ongekende waarschuwing af voor Covid-vaccins	Most-read British medical author appoints an unprecedented warning for Covid vaccines	Xander Nieuws
	Coronavirus: Noord-Korea's eerste positieve patiënt doodgeschoten - Naija Live Tv	Coronavirus: North Korea's first confirmed patient shot dead – Naija Live Tv	Naja TV
	Vitamine C beschermt tegen het Coronavirus	Vitamin C protects against Coronavirus	Health Impact News
	Pfizer Chief zal het Covid-19 vaccin niet nemen	The Chief of Pfizer refuses to get the Covid-19 vaccine	Washington Examiner

Pretest

The news messages and formatting of the critical thinking recommendation were pretested among 27 respondents (16 women, 11 men, *M*age = 32.18, *SD*age = 11.66, 48.1% academic

degree or higher). Two visual representations of the recommendation were tested: a textual banner in the blue colour of Twitter versus a white textual banner combined with a red exclamation mark. For the experiment, we selected the white banner with red exclamation mark, as it resulted in significantly more correctly recalled content ($t(26) = -3.22$, $p = .003$), was judged to be more informative ($t(26) = 2.56$, $p = .017$), motivating ($t(26) = 4.44$, $p < .001$) and visible (indicated by qualitative responses of the participants). Also, the pre-test examined whether the seven news messages were not obviously fake or true. Participants rated trust in each news message on a 4-point scale (1 = *Definitely not trust the news message*, 4 = *Definitely trust the news message*). Overall, true news messages were significantly more trusted ($M = 2.93$, $SD = 0.54$) than fake news messages ($M = 1.94$, $SD = 0.43$; $t(26) = -8.42$, $p < .001$). Nevertheless, there was still some desired variance in participants' evaluation of each news message: some people highly trusted fake news messages and distrusted true news messages. Therefore, all seven news messages were embedded in the news feed.

Measures

News Truth Discernment. The mediator news truth discernment was operationalized using a single question presenting the titles of the seven news messages. Participants were instructed to select the messages they perceived to be true. Participants' answers were converted into a news truth discernment score according to the formula: *proportion (hits) – proportion (false alarms)*. This formula follows Pennycook and colleagues' measures of news truth discernment (Pennycook & Rand, 2019; Pennycook et al., 2020). It compares to *d prime* (*d' = z(hits) – z(false alarms)*), which is a measure of discriminability or sensitivity (i.e., an individual's ability to detect a signal against noise while correcting for response biases) derived from signal detection theory (Batailler et al., 2021). To calculate our news truth discernment score, the proportion of false alarms (selected incorrect responses, i.e., fake news message perceived to be true, each assigned a weight of .25 as there were four fake news messages) was subtracted from the proportion of hits (selected correct responses, i.e., true news message perceived to be true, each assigned a weight of .33 as there were three true news message) per participant. As a result, scores could range from -1 to 1, with higher scores indicating that participants were better capable of discerning true from fake news, or in other words, that they were more sensitive to truth relative to falsity (cf. Pennycook et al., 2020), $M = 0.38$, $SD = 0.42$.

Trust in Fake News. The dependent variable trust in fake news was measured using the operationalization of Lutzke et al. (2019): "*To what extent do you think this news message is accurate?*" and "*To what extent do you think this news message is trustworthy?*". Participants are asked to rate each message on a 10-point scale (1 = *not at all accurate/trustworthy*, 10 = *very accurate/trustworthy*). The mean of the two items was calculated for each news message, after which the scores on the four fake news message were averaged into a single continuous variable called 'trust in fake news' (Cronbach's $\alpha = .88$, $M = 4.23$, $SD = 1.91$).

Information Literacy. The continuous moderator information literacy was measured using the Information Literacy Test (Boh Podgornik et al., 2016). Five multiple-choice items measured factual knowledge. For example, "*The most reliable, verified, concise, and comprehensive description of an unknown specialized concept can be found in:"* is accompanied with four responses: "*daily newspaper*", "*bilingual dictionary*", "*lexicon or*

encyclopaedia", or "*research article*". Each correct answer was worth 1 point, resulting in information literacy scores ranging from 0 to 5 (M = 3.00, SD = 1.21).

Impulsivity. Impulsivity was measured using the Dutch translation (Lijffijt & Barratt, 2005) of the Barratt Impulsiveness Scale - Brief (Steinberg et al., 2013). Eight items measured impulsivity on a four-point scale (1 = *rarely/never*, 4 = *always*). An example of an item is "*I do things without thinking*". Before creating the mean score, four items were reverse-coded. The variable was treated as a continuous moderator, with higher scores indicating higher trait impulsivity (Cronbach's α = .84, M = 2.57, SD = 0.51).

Sharing Intentions. Online and offline intentions to share the news messages are assessed as a means of assessing the convergent validity of the truth discernment and trust in fake news measures, as these intentions are known to be affected by prompts to judge news accuracy as well as perceptions of news accuracy and trust in news (Duffy et al., 2020; Lutzke et al., 2019; Pennycook et al., 2021). Sharing intentions are measured by asking participants: "*Would you share the news message online/offline?*" Participants had to indicate their intentions to share each message for both online and offline settings separately on a 7-point continuous scale (1 = *definitely not*, 7 = *definitely yes*). Overall, participants' intentions to share fake news online (M = 2.60, SD = 1.51) and offline (M = 3.81, SD = 1.62) were below the scale midpoint.

Participant Characteristics. Next to age, gender, and educational level the following control variables were assessed; "*How worried are you about Covid-19?*" (0 = *not worried at all*, 100 = *very worried*), "*How satisfied are you with the Dutch authorities Covid-19 approach?*" (0 = *not satisfied at all*, 100 = *very satisfied*), and "*How often do you check... news in general / Covid-19 specific news?*" (1 = *never*, 5 = *always*). In addition, participants were asked which channels they used to consume Covid-19 news. Also, general trust in Dutch news media was considered as a control variable because people who distrust media are more sceptical about the credibility of news information (Fletcher & Nielson, 2017). A scale consisting of five items measured the trust in news media (Meyer, 1988). Participants were asked whether Dutch news media is "*fair*", "*biased*" (reverse-coded), "*exact*", "*telling the whole story*", and "*trustworthy*" (1 = *completely disagree*, 5 = *completely agree*). Items were averaged into a reliable scale (Cronbach's α = .80). Furthermore, a question serving as manipulation check asked "*Did you notice a recommendation above each message?*" (*yes*/*no*/*not sure*).

Procedure

The study was conducted using the Qualtrics research platform. Participants received an invitation via the university's lab website or directly from the researchers. In the first part of the questionnaire, participants were informed about the general topic of the study - Covid-19-related news. At the time of data collection (end of April and early May 2021) very strict Covid-19 measures to contain the coronavirus (3rd wave) such as a night-time curfew were slowly released by the Dutch government (Rijksoverheid, n.d.). After participants provided informed consent, news consumption habits, level of Covid-19-related worries, and evaluation of the Dutch approach against Covid-19 were assessed. Subsequently, participants were randomly assigned to either the experimental or the control condition. They were instructed to look at the Twitter newsfeed as they normally would. After exposure, the titles of the seven news messages were presented in the same order as they appeared in the news feed. For each message, participants were asked about their sharing intentions, trust, and whether they thought the

message was true (i.e., truth discernment). Next, they answered a question serving as manipulation check, and questions assessing impulsivity, information literacy, and personal characteristics (i.e., age, gender, education, and trust in media). Finally, participants were thanked and debriefed. Here, people were informed about the aim of the study, manipulation, and which messages were fake.

Results

An overview of the participants' characteristics is presented in Table 2. As Table 2 shows, participants generally reported fairly high trust in news media, and they checked the news and Covid-19 news regularly. They mainly consumed Covid-19 news using social media, news apps and websites, and television. Participants were moderately worried about Covid-19 and somewhat dissatisfied with the Dutch Covid-19 approach.

Table 2. Participant Characteristics

Variable	Items/Scale	*N*	%	*Min*	*Max*	*M*	*SD*
Age				18.00	62.00	22.64	6.86
Gender	Female	165	79.7				
	Male	41	19.8				
	Other	1	.5				
Education	Less than academic degree	93	44.9				
	Academic degree or higher	114	55.1				
Covid-19 news sources	Social media	180	87.0				
	News apps and websites	156	75.4				
	Television	129	62.3				
	Radio	52	25.1				
	Newspaper	34	16.4				
Trust in news media	1 (*completely disagree*) to 5 (*completely agree*)			1.40	4.60	3.20	.60
News consumption frequency	1 (*never*) to 5 (*often*)			1.00	5.00	3.80	.93
Covid-19 news consumption frequency	1 (*never*) to 5 (*often*)			1.00	5.00	3.53	.91
Worries for Covid-19	0 (*not worried at all*) to 100 (*very worried*)			.00	96.00	55.21	20.44
Evaluation of Dutch approach against Covid-19	0 (*very dissatisfied*) to 100 (*very satisfied*)			1.00	87.00	43.57	18.35

Statistical Analysis

Data analysis was executed in IBM SPSS Statistics 27 (IBM Corp., 2020). The first hypothesis (mediation) was tested with PROCESS model 4 (Hayes, 2017; 5,000 bootstrapped samples). Hypotheses 2 and 3 (moderated mediation) were analysed using PROCESS model 7 (5,000 bootstrapped samples). We decided to test H2 and H3 separately because the incorporation of two moderators in one model complicates the interpretation of the results (Hayes, 2018). Before the analyses, assumptions were checked for all variables. All variables, except age, met the assumption of normality.[2] Furthermore, all variables met the assumptions of linearity and homogeneity of variances.

Randomisation Check

To check whether background variables and potentially confounding variables were equally distributed across the two conditions, a randomization check was conducted. Participants in the two conditions did not differ in terms of gender ($\chi2(2) = 2.30$, $p = .316$), education level ($\chi2(1) = 0.05$, $p = .817$), age ($t(205) = 1.04$, $p = .297$), news consumption frequency ($t(205) = -0.853$, $p = .395$), Covid-19 news consumption frequency ($t(205) = -0.66$, $p = .512$), evaluation of Dutch approach regarding Covid-19 ($t(205) = 0.22$, $p = .823$), concerns about Covid-19 ($t(205) = -0.26$, $p = .793$), and trust in media ($t(205) = 0.44$, $p = .663$). Therefore, none of the variables were included as covariates in the statistical analyses.

Manipulation Check

A manipulation check assessed whether participants were aware of the recommendation. Though a considerable number of participants were unsure about whether they had seen the recommendation (30.5% in the experimental condition and 44.1% in the control condition), significantly more participants reported to have noticed the recommendation in the experimental condition (59.0%) than in the control condition (6.9%; $\chi^2(1) = 70.94$, $p < .001$). Therefore, the manipulation is considered successful.

Validity Check

To check whether our measures of *news truth discernment* and *trust in fake news* adequately capture the phenomenon we aim to measure, we tested their correlation with associated concepts. Results show that news truth discernment was negatively associated with the intentions to share fake news messages both online ($r = -.49$, $p < .001$) and offline ($r = -.41$, $p < .001$). Trust in fake news messages was positively associated with intentions to share fake news online ($r = .61$, $p < .001$) and offline ($r = .58$, $p < .001$). These results show that less accurate truth discernment and more trust in fake news are associated with increased intentions to share fake news, as was expected based on the literature.

Hypotheses Tests

The Effect of the Critical Thinking Recommendations. Hypothesis 1 stated that participants exposed to a newsfeed with critical thinking recommendations are less likely to trust fake news than participants in the control condition and that this effect is mediated by people's ability to accurately discern true from fake news. The model including recommendation (independent variable), trust in fake news (dependent variable), and news truth discernment (mediator) explained 24.8% of the variance in trust in fake news ($F(2, 204) = 30,28$, $p < .001$), see Figure

1. There was no direct effect of recommendations on trust in fake news (b = -.31, t(205) = -1.16, p = .247). However, the effect of recommendations on news truth discernment was significant (b = .17, t(205) = 2.92, p = .004) as well as the effect of news truth discernment on trust in fake news (b = -2.30, t(204) = -8.10, p < .001). The indirect effect of recommendations on discernment and, consequently, trust in fake news was significant (b = -.38, 95% CI [-0.69;-0.12]), supporting H1. Thus, participants who were exposed to the recommendations were better capable to discern true news from fake news, and, in turn, had lower trust in fake news than those not exposed to the recommendations.

The Effect of Information Literacy. Hypothesis 2 stated that the mediation effect of H1 is moderated by information literacy. The model including recommendation (independent variable), news truth discernment (mediator), trust in fake news (dependent variable), and information literacy (moderator) was significant (R^2 = .11; F(3, 202) = 6.96, p < .001). The entire model is presented in Figure 2. Information literacy directly influenced news truth discernment (b = .20, t(202) = 2.71, 95% CI [0.05;0.35], p = .008). However, the interaction between recommendation and information literacy on truth discernment was non-significant (b = -.08, t(202) = -1.88, 95% CI [-0.16;0.00], p = .062), rejecting H2. Given that the moderation was non-significant but the p-value approached significance at p =.062, we further explored this finding. The exploratory follow-up analysis showed that the positive effect of the recommendation on truth discernment was significant among participants with low (1 SD below average; b = -.53, 95% CI [-0.96;-0.16]) and moderate levels of information literacy (average; b = -.32, 95% CI [-0.61;-0.06]), whereas no effect of the recommendation was observed among participants with high levels of information literacy (1 SD above average; b = -.10, 95% CI [-0.42;0.21]), see Figure 3. Cautious interpretation of these exploratory findings suggests that participants with low and moderate levels of information literacy might benefit from a stop and think recommendation, whereas people with high information literacy levels might not.

The Effect of Impulsivity. Hypothesis 3 stated that the mediation effect of H1 is moderated by impulsivity. The model including recommendation (independent variable), news truth discernment (mediator), trust in fake news (dependent variable), and impulsivity (moderator) was significant (R^2 = .05; F(3, 203) = 3.21, p = .024). See Figure 4 for the entire model. The interaction effect of impulsivity and recommendation on truth discernment was not significant (b = -.09, t(203) = .31, 95% CI [-0.47;0.64], p = .759). Thus, impulsivity does not affect the relationship between recommendation and people's ability to discern true news from fake news, rejecting H3.

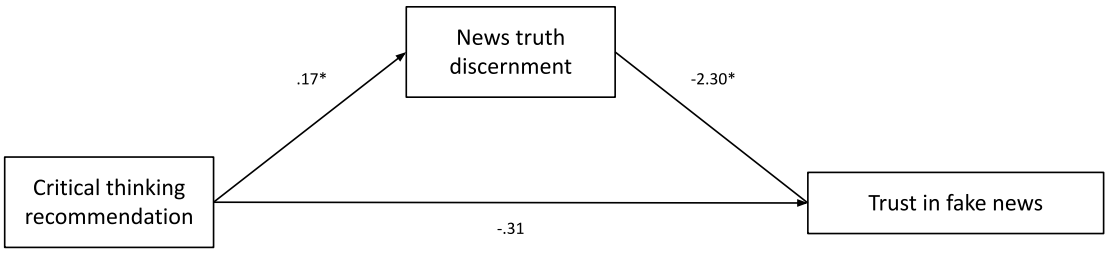
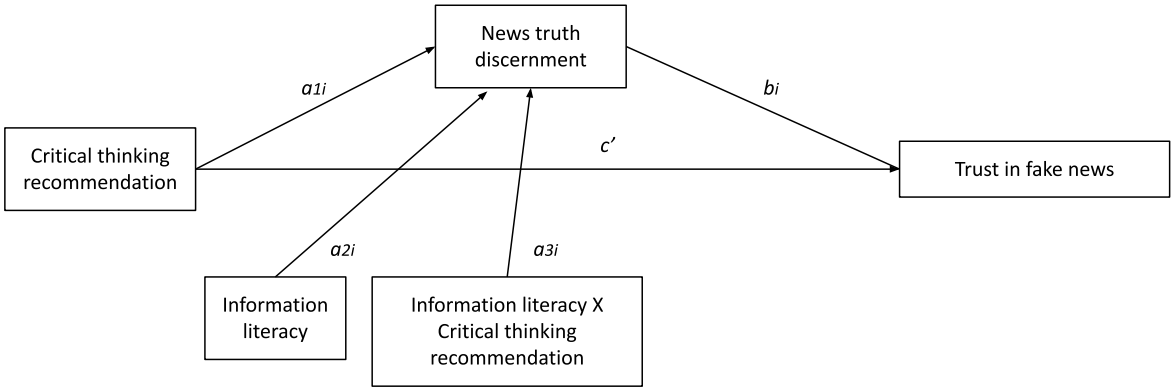


Figure 1. Testing the Mediating Role of Truth Discernment in the Effect of Recommendations on Trust in Fake News



a1i: b = .14, t(202) = 2.45, p = .015
a2i: b = .20, t(202) = 2.71 p = .008
a3i: b = -.08, t(202) = -1.88, p = .062
bi: b = -2.31, t(203) = -7.60, p < .001
c': b = .08, t(203) = .31, p = .760

Figure 2. Testing Information Literacy as a Moderator

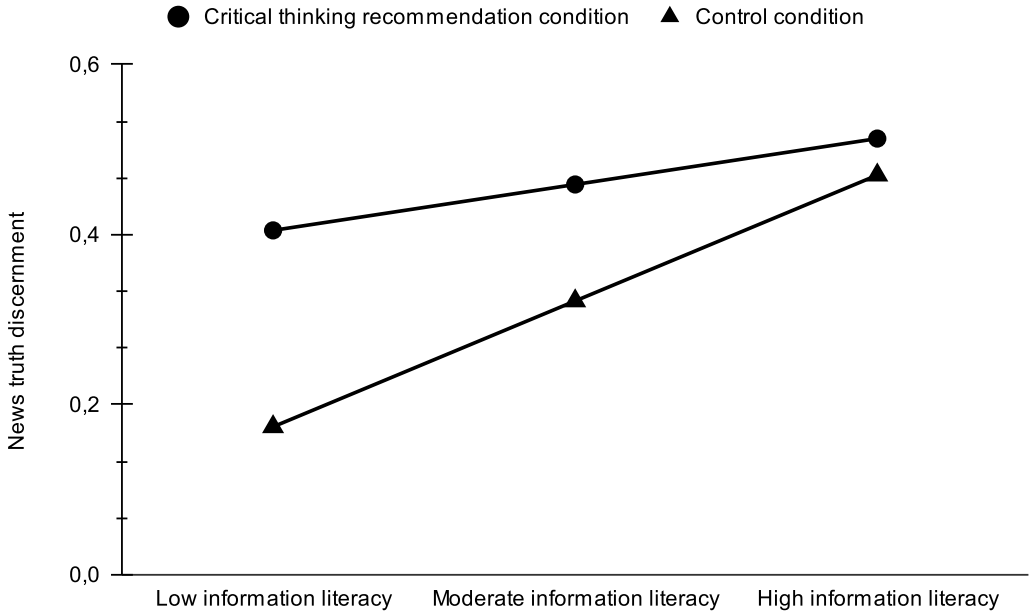
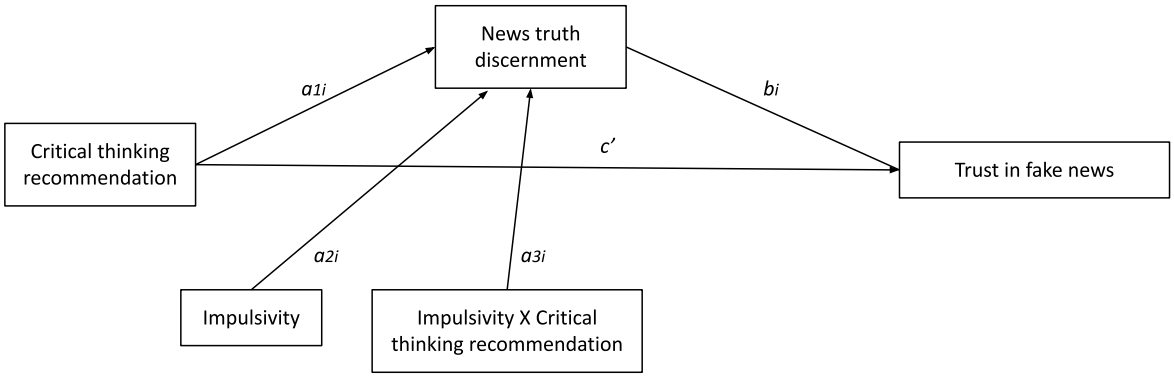


Figure 3. Exploring the Moderating Role of Information Literacy in the Effect of Recommendations on News Truth Discernment

Note. The difference between conditions was non-significant for high levels of information literacy but significant for moderate and low levels of information literacy.



a1i: b = -.13, t(203) = -.14, p = .890
a2i: b = -.25, t(203) = -.48, p = .634
a3i: b = -.09, t(203) = .31, p = .759
bi: b = -2.30, t(204) = -7.60, p < .001
c': b = .08, t(204) = .32, p = .752

Figure 4. Testing Impulsivity as a Moderator

Discussion

This study investigated whether and how recommendations attached to news messages can prevent people from trusting fake news. It thereby responds to a call for context specific, evidence-based interventions aimed at correcting fake news (Ha et al., 2021). Our research shows that a critical thinking recommendation can effectively help people to accurately discern true news from fake news, which in turn increases distrust in fake news. This is in line with prior studies showing that a lack of careful reasoning when scrolling through social media newsfeeds makes people susceptible to perceiving fake news as being true (Fazio, 2020; Salovich & Rapp, 2020; Pennycook & Rand, 2021; Pennycook et al., 2021) and evidences a mediating role for truth discernment in the previously demonstrated relation between 'stop and think' recommendations and trust in fake news (Lutzke et al., 2019; Fazio, 2020). Our finding that recommendations improve news truth discernment can be explained in light of dual process models, which propose that people can switch between states in which information is processed in a shallow, less critical versus a deep, more critical manner (cf. Kahneman, 2011). Whether people's increased discernment was indeed caused by such a shift in processing mode should be addressed in future research.

Moreover, this study builds on prior studies in which individuals were instructed to use guidelines (i.e., multiple questions) to evaluate the credibility of news messages on social media (Fazio, 2020; Lutzke et al., 2019), which successfully reduces the impact of fake news. However, these interventions are not optimal since people have to explicitly reflect on several questions to assess the news, which is unlikely to occur during real-life news consumption. As a next step, this study confirms that a single-sentence recommendation added to social media messages is sufficient to effectively stimulate people to thoughtfully evaluate news and, accordingly, reduce the trust in fake news. These findings suggest that recommendations can be an effective, scalable, and low-cost intervention against fake news that could be implemented on social media platforms. The idea to pause and think is also seen in practice (see the United Nations initiative #TakeCareBeforeYouShare; UNESCO, 2020). Platforms can

think about such recommendation-like interventions as a complementary solution besides technology-based solutions that aim to filter out fake news in general.

This study provides preliminary insights into whether personal characteristics might affect the effectiveness of fake news interventions. Overall, the findings show that people with higher levels of information literacy are better at identifying fake news than people with lower levels of information literacy, confirming Jones-Jang et al. (2021). Though the moderating effect of information literacy was non-significant (p = .062), the findings did suggest an interesting trend. Post-hoc exploration shows that people with low or moderate levels of information literacy might be better able to discern true from fake news and have lower trust in fake news when they are (versus are not) exposed to a newsfeed with critical thinking recommendations. No such trend is found for people with high levels of information literacy. In all, this exploratory finding suggests a potentially inverse relation: the lower people's literacy levels, the more they might benefit from the critical thinking recommendation. Rather than already speculate about the meaning of such an inverse relation, future confirmatory research should assess whether this trend holds. Nevertheless, the finding responds to the call of Scheufele and Krause (2019) to give special attention to the underserved lower literate people when developing an intervention that fights fake news. It is promising that those who have the most difficulty with correctly identifying fake news - from an information literacy perspective – seem to potentially benefit most from being exposed to a simple recommendation.

Lastly, impulsivity level did not moderate the effect of a critical thinking recommendation on truth discernment. It was expected that impulsive people are less likely to thoughtfully evaluate news after being prompted than less impulsive people, since they are less likely to shift to a reflective, critical thinking style (Strack & Deutsch, 2006). However, our null finding may be explained by studies showing that impulsive people tend to use social media more intensively than less impulsive people (Savci & Aysan, 2016). This might make them more aware of the mores on social media, advancing a critical processing mode in this group of people. Though highly speculative, this might explain why impulsivity did not moderate the recommendation effect, which is worth investigating further. Another potential explanation for the absence of a moderation is the self-reported measure of impulsivity. Potentially, participants may have dimmed their impulsivity score as this trait is seen as something that lies at the heart of dysfunctioning (DeYoung & Rueter, 2010).

Limitations and Future Research

Some limitations need to be addressed. Even though our study showed that a recommendation can be effective, this does not imply that prompting people to critically evaluate messages inherently reduces their susceptibility to misinformation. Literature from a motivated reasoning perspective (rather than our 'inattention' perspective) on the impact of misinformation suggests that people, even after deliberate evaluation, may conclude that an incorrect message is true if its content reinforces someone's political, religious, or social identity (Van der Linden, 2022). A limitation of our study is that we did not take people's *a priori* convictions into account. In future research, it is relevant to examine whether and how people's identity and prior convictions might impact our current findings. Evidence convincingly shows that existing beliefs are difficult to change (e.g., Kessler et al., 2019; Pluviano et al., 2017) and can affect the way in which people process and evaluate information (e.g., Meppelink et al., 2019). It is unlikely that our one-sentence recommendation is sufficient to impact those whose strong prior

convictions match the viewpoints described in the fake news item. This demonstrates that existing perceptions of identity, beliefs, or motivations are viable potential moderators of the presented recommendation effects, which should be addressed in future research. Additionally, various other individual characteristics might moderate the effect of recommendations on truth discernment and trust in fake news, such as conspiracy thinking, trust in authority, religiosity, political conviction, and preference for intuition or deliberation.

It should also be emphasized that, ideally, more participants should have been included in our study. This was not possible due to the limited resources in terms of time and money, as this study was performed as part of a Master's thesis. This has some implications, as the reported effect sizes in underpowered studies can differ from the population value and might therefore lead to biased conclusions (Crutzen & Peters, 2017). The findings of our study should therefore be interpreted with caution, and future research should test the same hypotheses with a larger sample size to confirm our results. Also, the convenience sample consisting of mainly students might not be representative of how the Dutch population deals with fake news on social media. Nevertheless, treatment effects are considerably similar when comparing convenience samples with nationally representative samples (Mullinix et al., 2015). On the other hand, older generations use social media differently than younger generations (Auxier & Anderson, 2021). Therefore, it would be valuable to investigate potential age or generational differences in how people deal with social media (mis)information.

Another limitation relates to the point that recommendations in and of themselves do not equip people with any new skills to identify and resist misinformation. Moreover, the effect of the recommendation might wear out on the long term and/or after several exposures. Since our study only consisted of one exposure, we are unsure whether and how people respond to the recommendation after multiple exposures. Future research should therefore explore how the recommendation is processed and evaluated during repeated exposures and whether effects hold over time, to get further insights into the recommendation's impact.

Furthermore, one-third of the participants in this study did not notice the recommendation. This fosters the question whether and how recommendations can be effective when they are not explicitly remembered and perhaps not even consciously processed. Studies on implicit memory or subliminal priming may offer interesting insights about the latent processing of recommendations and warnings. Yeu et al. (2013) found that people's implicit memory of a banner ad was better when they reported noticing but not remembering it, compared to when they reported not noticing it. This suggests that low levels of attentiveness can still affect memory in ways that people are unable to report about. Also, Verwijmeren et al. (2013) found that presenting a warning below consciousness-levels did not make people aware of the warning itself but did impact their behaviour. This signifies that although recommendations might not be consciously processed or explicitly remembered, they may affect the identification and evaluation of fake news.

A final limitation is the experimental setting. Although the stimulus Twitter Covid-19 newsfeed is an exact copy of the original newsfeed, the ecological validity could be improved as the participants were not able to react to posts, nor did the Twitter feed consist of a wide range of topics as is usually the case (e.g., entertainment, relatives, celebrities). This limitation is relevant, as accuracy and trustworthiness are not essential for all topics. Therefore, follow-up studies should explore whether the recommendations are still effective when they are unobtrusively embedded in a more natural setting. Also, as the messages circulated on social media, it could be possible that participants already saw the messages before participating in

the study. Some results suggest that previous exposure to fake news messages is negatively related to news truth discernment (Jones-Jang et al., 2021), whereas other results suggest that prior exposure to fake news increases discernment accuracy (Pennycook et al., 2018). These contradicting results do not offer clear insights on how prior exposure might have affected the truth discernment scores in the present study. Therefore, it is recommended to include an assessment of prior exposure in future studies.

Conclusion

A simple, single-sentence recommendation to stop and think before trusting information can effectively help people to accurately discern true news from fake news and accordingly reduce their trust in fake news. It is promising that people with low and moderate levels of information literacy seem to benefit most from the recommendation, but this should be confirmed in further research. Impulsivity does not play a role in the effectiveness of the recommendation. The practical implication of this study is that a recommendation is a scalable, low-cost, and easily implementable intervention to fight misinformation on social media in terms of increasing people's identification of, and reducing their trust in, fake news. However, future research is urged to further explore whether this holds over time and potentially prevents the spread and consumption of fake news, as well as examine the boundary conditions of the recommendations' impact. Hopefully, recommendations are a meaningful addition to today's media landscape that has to deal with hazardous fake news.

Notes

1. Ideally, we would have aimed for a larger sample size that was based on an a priori power analysis. However, as this study is based on a Master's thesis we operated with limited monetary resources and under strict time constraints. A post hoc sensitivity power analysis using G*Power ($\alpha = .05$, $\beta = .80$) showed that with the current sample size we were able to detect effect sizes of $\geq .39$, meaning that our study was likely underpowered.
2. This can be explained by the fact that the minimum age in the sample was 18, the maximum age 62, and that most of the participants were students.

Acknowledgements

We thank the editors and reviewers for their valuable suggestions. This article is based on Jasmijn Kruijt's Master's thesis as part of the Research Master Communication Science, University of Amsterdam.

Ethical Approval

The research reported in this manuscript was approved by the Ethical Review Board of the Amsterdam School of Communication Research, University of Amsterdam (nr: 2021-PC-13364).

Funding

Participant compensation (gift card) was paid by the first author. Research time of the second and third author was funded by their respective institutions, i.e., ASCoR and BSI.

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.

References

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

Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, *56.* https://doi.org/10.1016/j.tele.2020.101475

Auxier, B., & Anderson, M. (2021). Social media use in 2021. *Pew Research Center*. https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf

Baron, J. (2019). Actively open-minded thinking in politics. *Cognition*, *188*, 8-18. https://doi.org/10.1016/j.cognition.2018.10.004

Batailler, C., Brannon, S. M., Teas, P. E., & Gawronski, B. (2021). A signal detection approach to understanding the identification of fake news. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620986135

Boczkowski, P. J., Mitchelstein, E., & Matassi, M. (2018). "News comes across when I'm in a moment of leisure": Understanding the practices of incidental news consumption on social media. *New Media & Society*, *20*(10), 3523-3539. https://doi.org/10.1177/1461444817750396

Boh Podgornik, B., Dolničar, D., Šorgo, A., & Bartol, T. (2016). Development, testing, and validation of an information literacy test (ILT) for higher education. *Journal of the association for Information Science and Technology*, *67*(10), 2420-2436. https://doi.org/10.1002/asi.23586

Brashier, N. M., Eliseev, E. D., & Marsh, E. J. (2020). An initial accuracy focus prevents illusory truth. *Cognition*, *194*, 104054. https://doi.org/10.1016/j.cognition.2019.104054

Bridgman, A., Merkley, E., Loewen, P. J., Owen, T., Ruths, D., Teichmann, L., & Zhilin, O. (2020). The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. *Harvard Kennedy School Misinformation Review*, *1*(3). https://doi.org/10.37016/mr-2020-028

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, *8*(1), 108-117. https://doi.org/10.1016/j.jarmac.2018.09.005

Cardoso Durier da Silva, F., Vieira, R., & Garcia, A. C. (2019). Can machines learn to detect fake news? a survey focused on social media. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2019.332

Casero-Ripollés, A. (2020). Impact of Covid-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak. *El profesional de la información*, *29*(2). https://doi.org/10.3145/epi.2020.mar.23

Chan, M. P. S., Jones, C. R., Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, *28*(11), 1531-1546. https://doi.org/10.1177/0956797617714579

Coutlee, C. G., Politzer, C. S., Hoyle, R. H., & Huettel, S. A. (2014). An abbreviated impulsiveness scale constructed through confirmatory factor analysis of the Barratt impulsiveness scale version 11. *Archives of Scientific Psychology*, *2*(1), 1-12. https://doi.org/10.1037/arc0000005

Crutzen, R., & Peters, G. J. Y. (2017). Targeting next generations to change the common practice of underpowered research. *Frontiers in Psychology*, *8*, 1184. https://doi.org/10.3389/fpsyg.2017.01184

Dentith, M. R. (2016). The problem of fake news. *Public Reason, 8*(1-2): 65-79. https://philpapers.org/rec/DENTPO-31

De Paor, S., & Heravi, B. (2020). Information literacy and fake news: How the field of librarianship can help combat the epidemic of fake news. *The Journal of Academic Librarianship*, *46*(5), 102218. https://doi.org/10.1016/j.acalib.2020.102218

DeYoung, C. G., & Rueter, A. R. (2010). Impulsivity as a personality trait. *Handbook of self-regulation: Research, theory, and applications*, *2*, 485-502. Guilford Press.

Diviani, N., & Meppelink, C. S. (2017). The impact of recommendations and warnings on the quality evaluation of health websites: An online experiment. *Computers in Human Behavior, 71*, 122-129. https://doi.org/10.1016/j.chb.2017.01.057

Duffy, A., Tandoc, E., & Ling, R. (2020). Too good to be true, too good not to share: the social utility of fake news. *Information, Communication & Society*, *23*(13), 1965-1979. https://doi.org/10.1080/1369118X.2019.1623904

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

Fazio, L. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School Misinformation Review*, *1*(2). https://doi.org/10.37016/mr-2020-009

Fletcher, R., & Nielsen, R. K. (2017). People don't trust news media–and this is key to the global misinformation debate. *AA. VV., Understanding and Addressing the Disinformation Ecosystem*, 13-17. https://firstdraftnews.org/wp-content/uploads/2018/03/The-Disinformation-Ecosystem-20180207-v4.pdf

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692. https://doi.org/10.1037/0033-2909.132.5.692

Ha, L., Andreu Perez, L., & Ray, R. (2021). Mapping recent development in scholarship on fake news and misinformation, 2008 to 2017: Disciplinary contribution, topics, and impact. *American Behavioral Scientist*, *65*(2), 290-315. https://doi.org/10.1177/0002764219869402

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, *85*(1), 4-40. https://doi.org/10.1080/03637751.2017.1352100

Iacobucci, G. (2019). Vaccination: "fake news" on social media may be harming UK uptake, report warns. *BMJ: British Medical Journal (Online)*, *364*. https://doi.org/10.1136/bmj.l365

IBM Corp. (2020). *IBM SPSS Statistics for Windows*, Version 27.0. Armonk, NY: IBM Corp.

Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S., & Lupu, Y. (2020). The online competition between pro- and anti-vaccination views. *Nature, 582*(7811), 230–233. https://doi.org/10.1038/s41586-020-2281-1

Jones-Jang, S. M., Mortensen, T., & Liu, J. (2021). Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist*, *65*(2), 371-388. https://doi.org/10.1177/0002764219869406

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

Kessler, E. D., Braasch, J. L. G., & Kardash, C. M. (2019). Individual differences in revising (and maintaining) accurate and inaccurate beliefs about childhood vaccinations. *Discourse Processes,* 56(5–6), 415–428. https://doi.org/10.1080/0163853X.2019.1596709

Khan, M. L., & Idris, I. K. (2019). Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behaviour & Information Technology*, *38*(12), 1194-1212. https://doi.org/10.1080/0144929X.2019.1578828

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., … Zittrain, J. L. (2018). The science of fake news. *Science, 9*, 1094–1096. https://doi.org/10.1126/science.aao2998

Lee, J. J., Kang, K., Wang, M. P., Zhao, S. Z., Wong, J. Y. H., O'Connor, S., . . . Shin, S. (2020). Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: Cross-sectional online study. *Journal of Medical Internet Research, 22*(11), e22205. https://doi.org/10.2196/22205

Lijffijt, M., & Barratt, E. S. (2005). *BIS-11 Dutch Version*. http://www.impulsivity.org/measurement/bis11_Dutch

Lutkenhaus, R. O., Jansz, J., & Bouman, M. P. A. (2019). Mapping the Dutch vaccination debate on Twitter: Identifying communities, narratives, and interactions. *Vaccine: X, 1*, 100019. https://doi.org/10.1016/j.jvacx.2019.100019

Lutzke, L., Drummond, C., Slovic, P., & Árvai, J. (2019). Priming critical thinking: Simple interventions limit the influence of fake news about climate change on Facebook. *Global Environmental Change*, *58*, 101964. https://doi.org/10.1016/j.gloenvcha.2019.101964

Mele, N., Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K., ... & Mattsson, C. (2017). Combating fake news: An agenda for research and action. *https://shorensteincenter.org/wp-content/uploads/2017/05/Combating-Fake-News-Agenda-for-Research-1.pdf.*

Meppelink, C.S., Smit, E. G., Fransen, M. L. & Diviani, N. (2019). "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of Health Communication, 24*(2), 129-140. https://doi.org/10.1080/10810730.2019.1583701

Messer, S. B. (1976). Reflection-impulsivity: A review. *Psychological Bulletin, 83*(6), 1026–1052. https://doi.org/10.1037/0033-2909.83.6.1026

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology, 58*(13), 2078-2091. https://doi.org/10.1002/asi.20672

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413-439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Meyer, P. (1988). Defining and measuring credibility of newspapers: Developing an index. *Journalism Quarterly*, *65*(3), 567-574. https://doi.org/10.1177/107769908806500301

Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, *2*(2), 109-138. https://doi.org/10.1017/XPS.2015.19

Naeem, S. B., Bhatti, R., & Khan, A. (2021). An exploration of how fake news is taking over social media and putting public health at risk. *Health Information & Libraries Journal*, *38*(2), 143-149. https://doi.org/10.1111/hir.12320

Pavleska, T., Školkay, A., Zankova, B., Ribeiro, N., & Bechmann, A. (2018). Performance analysis of fact-checking organizations and initiatives in Europe: a critical overview of online platforms fighting fake news. *Social Media and Convergence*, *29*. https://www.researchgate.net/profile/Nelson-Ribeiro/publication/327121444

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865-1880. https://doi.org/10.1037/xge0000465

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. In *Communication and Persuasion* (pp. 1-24). Springer , New York, NY.

Pluviano, S., Watt, C., & Sala, S. D. (2017). Misinformation lingers in memory: Failure of three pro-vaccination strategies. *PLOS ONE*, 12(7), e0181640. https://doi.org/10.1371/journal.pone.0181640

Rapp, D. N., & Salovich, N. A. (2018). Can't we just disregard fake news? The consequences of exposure to inaccurate information. *Policy Insights from the Behavioral and Brain Sciences*, *5*(2), 232-239. https://doi.org/10.1177/2372732218785193

Reimann, N. (2020). Some Americans Are Tragically Still Drinking Bleach As A Coronavirus 'Cure'. Forbes. https://www.forbes.com/sites/nicholasreimann/2020/08/24/some-americans-are-tragically-still-drinking-bleach-as-a-coronavirus-cure/?sh=4ce6610c6748

Rijksoverheid (n.d.). *April 2021: De maand van het volhouden en eerste stap openingsplan*. Coronavirus tijdlijn, Rijksoverheid. Retrieved March 30, 2022, from https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/april-2021-de-maand-van-het-volhouden-en-eerste-stap-openingsplan.

Rosenberg, H., Syed, S., & Rezaie, S. (2020). The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. *Canadian Journal of Emergency Medicine*, *22*(4), 418-421 https://doi.org/10.1017/cem.2020.361

Savci, M., & Aysan, F. (2016). Relationship between impulsivity, social media usage and loneliness. *Educational Process: International Journal*, *5*(2), 106-115. https://doi.org/10.12973/edupij.2016.52.2

Salovich, N. A., & Rapp, D. N. (2020). Misinformed and unaware? Metacognition and the influence of inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000977

Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, *116*(16), 7662-7669. https://doi.org/10.1073/pnas.1805871115

Shahi, G. K., Dirkson, A., & Majchrzak, T. A. (2021). An exploratory study of COVID-19 misinformation on twitter. *Online Social Networks and Media*, *22*, 100104. https://doi.org/10.1016/j.osnem.2020.100104

Spielvogel, I., Naderer, B., & Matthes, J. (2021). Disclosing product placement in audiovisual media services: a practical and scientific perspective on the implementation of disclosures across the European Union. *International Journal of Advertising*, *40*(1), 5-25. https://doi.org/10.1080/02650487.2020.1781478

Steinberg, L., Sharp, C., Stanford, M. S., & Tharp, A. T. (2013). New tricks for an old measure: The development of the Barratt Impulsiveness Scale–Brief (BIS-Brief). *Psychological Assessment*, *25*(1), 216-226. https://doi.org/10.1037/a0030550

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220-247. https://doi.org/10.1207/s15327957pspr0803_1

Strack, F., & Deutsch, R. (2006). Reflective and impulsive determinants of consumer behavior. *Journal of Consumer Psychology*, *16*(3), 205-216. https://doi.org/10.1207/s15327663jcp1603_2

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137-153. https://doi.org/10.1080/21670811.2017.1360143

UNESCO (2020, July 7). *Pause. Take care before you share*. https://en.unesco.org/news/pause-take-care-you-share

Van der Linden, S. (2022). Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine, 28,* 460-467. https://doi.org/10.1038/s41591-022-01713-6

Van Reijmersdal, E. A., Fransen, M. L., van Noort, G., Opree, S. J., Vandeberg, L., Reusch, S., van Lieshout, F., & Boerman, S. C. (2016). Effects of disclosing sponsored content in blogs: How the use of resistance strategies mediates effects on persuasion. *American Behavioral Scientist*, *60*(12), 1458-1474. https://doi.org/10.1177/0002764216660141

Van Stekelenburg, A., Schaap, G., Veling, H., & Buijzen, M. (2020). Correcting misperceptions: The causal role of motivation in corrective science communication about vaccine and food safety. *Science Communication*, 42(1), 31-60. https://doi.org/10.1177%2F1075547019898256

Verwijmeren, T., Karremans, J. C., Bernritter, S. F., Stroebe, W., & Wigboldus, D. H. (2013). Warning: You are being primed! The effect of a warning on the impact of subliminal ads. *Journal of Experimental Social Psychology*, *49*(6), 1124-1129. https://doi.org/10.1016/j.jesp.2013.06.010

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). Information disorder: The essential glossary. *Harvard, MA: Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School*. Harvard, MA, USA.

Watson, A. (2021, September). *Coronavirus fake news frequency in the UK in September 2020*. Statista. https://www.statista.com/statistics/1112492/coronavirus-fake-news-frequency-in-the-uk/

Williams, D. A., & Wavell, C. (2007). Secondary school teachers' conceptions of student information literacy. *Journal of Librarianship and Information Science*, *39*(4), 199-212. https://doi.org/10.1177/0961000607083211

Yeu, M., Yoon, H. S., Taylor, C. R., & Lee, D. H. (2013). Are banner advertisements in online games effective? *Journal of Advertising*, *42*(2-3), 241-250. https://doi.org/10.1080/00913367.2013.774604

Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). Fake news: Fundamental theories, detection strategies and challenges. In *Proceedings of the twelfth ACM international conference on web search and data mining* (pp. 836-837). https://doi.org/10.1145/3289600.3291382

Zhu, B., Chen, C., Loftus, E. F., He, Q., Chen, C., Lei, X., Lin, C., & Dong, Q. (2012). Brief exposure to misinformation can lead to long-term false memories. *Applied Cognitive Psychology*, *26*(2), 301-307. https://doi.org/10.1002/acp.1825

Zuckerman, A., & Chaiken, S. (1998). A heuristic-systematic processing analysis of the effectiveness of product warning labels. *Psychology & Marketing*, *15*(7), 621-642. https://doi.org/10.1002/(SICI)1520-6793(199810)15:7%3C621::AID-MAR2%3E3.0.CO;2-H

Author Contributions

Conceptualisation (main idea, theory): Jasmijn Kruijt, Corine S. Meppelink, Lisa Vandeberg
Funding acquisition: Jasmijn Kruijt
Project administration: Jasmijn Kruijt
Methodology (design, operationalisation): Jasmijn Kruijt
Data collection: Jasmijn Kruijt
Data analysis: Jasmijn Kruijt
Writing – original draft: Jasmijn Kruijt
Writing – review & editing: Jasmijn Kruijt, Corine S. Meppelink, Lisa Vandeberg

Author Biographies

Jasmijn Kruijt (MSc in Communication Science, University of Amsterdam) is a former Communication Science Research Master student at the University of Amsterdam. Currently, she is working as a Junior Lecturer at the University of Amsterdam.

Corine S. Meppelink (PhD in Communication Science, University of Amsterdam) is an Assistant Professor in Persuasive Communication at the Amsterdam School of Communication Research, University of Amsterdam. Her research focuses on digital literacy skills, particularly in the health domain, and how inequalities in this respect impact the processing of and consequences of online (mis)information.

Lisa Vandeberg (PhD in Cognitive Psychology, Erasmus University Rotterdam) is an Assistant Professor in Persuasive Communication at the Behavioral Science Institute, Radboud University, as well as a post-doctoral researcher at IQ Healthcare, Radboud University Medical Centre. Her research focuses on cognitive biases in prosocial communication.