



UvA-DARE (Digital Academic Repository)

Rhyme or reason

Medium-term effects of heuristic and traditional media literacy interventions

Dobber, Tom; Strikovic, Edina; Tulin, Marina

DOI

[10.1080/1369118X.2025.2533315](https://doi.org/10.1080/1369118X.2025.2533315)

Publication date

2026

Document Version

Final published version

Published in

Information Communication and Society

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Dobber, T., Strikovic, E., & Tulin, M. (2026). Rhyme or reason: Medium-term effects of heuristic and traditional media literacy interventions. *Information Communication and Society*, 29(2), 657-674 . <https://doi.org/10.1080/1369118X.2025.2533315>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Rhyme or reason: medium-term effects of heuristic and traditional media literacy interventions

Tom Dobber , Edina Strikovic  and Marina Tulin

Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, Netherlands

ABSTRACT

Media literacy interventions are at the forefront of anti-misinformation policy. Current interventions are often complex, time-consuming, and lack a clear theoretical basis, which might negatively affect their (longer-term) effectiveness. This study draws on Protection Motivation Theory to design and test a novel ‘heuristic intervention’ that is meant to be easily remembered and carried out. Since there is only limited information about the longer-term effectiveness of traditional interventions, this study uses a two-wave experimental design to test the immediate and medium-term effectiveness of heuristic interventions against the traditional approach regarding accuracy in discernment. Results show that the heuristic intervention shows promise by outperforming the control group and the traditional intervention, but only in the immediate term. Findings advance our understanding of how to design effective, theory-based, and potentially lasting interventions, offering insights into mitigating strategies addressing misinformation.

ARTICLE HISTORY



Received 29 November 2024
Accepted 8 July 2025


KEYWORDS

Media literacy;
misinformation; heuristic;
Protection Motivation
Theory; resilience

Media literacy interventions are premised upon the idea that knowledge of telltale signs of misinformation helps people not only recognize such information, but also act in ways to avoid misinformation. There are roughly two main categories of media literacy interventions: interactive interventions, and non-interactive interventions (see Droog et al., 2024). Interactive interventions include games (Basol et al., 2020) and workshops (Moore & Hancock, 2022). Non-interactive interventions include informational texts, graphics and videos. This current study focuses on the non-interactive media literacy interventions, which we will dub ‘traditional interventions’.

Traditional interventions are popular amongst policymakers and practitioners in part because of their scalability: it is easier and cheaper to unilaterally send information than to interact. In many cases traditional interventions should also require less effort from the

CONTACT Tom Dobber  t.dobber@uva.nl  University of Amsterdam, Nieuwe Achtergracht 166, Amsterdam 1018 WV, Netherlands

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/1369118X.2025.2533315>.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

receiver. Correspondingly, a large part of the academic literature on media literacy focuses on these traditional interventions (Droog et al., 2024).

Research has shown that traditional interventions *prima facie* show promise in making people more critical toward misinformation. However, emerging research points to potential negative spillover effects. Traditional interventions seem to make people more skeptical toward false *and* true information (Hameleers, 2023; Hoes et al., 2024; van der Meer et al., 2023). Some scholars even suggest that traditional media literacy interventions do more harm than good (Hoes et al., 2024). In this current study, we attempt to explore ways to improve upon traditional interventions by developing and testing a ‘heuristic media literacy intervention’. The heuristic intervention is premised upon the observation that traditional interventions are rational and complex, which might attenuate their (longer-term) effectiveness as they are challenging to remember and carry out. For example, traditional interventions seem complex in the sense that they ask participants to do multiple things, such as ‘check the source’, ‘search for the facts’, and evaluate to what extent ‘consequences [are] logically connected to causes’ (Hameleers, 2023, A2). Importantly, it seems that most traditional interventions lack a clear theoretical basis that explains why specific elements need to be communicated to target audiences. For instance, Guess et al. (2020) use Facebook’s ‘tips to spot false news’. This is an ecologically valid intervention, but the theoretical grounding of these tips is unclear if not absent. This current study draws from Protection Motivation Theory for a new type of intervention that encourages heuristic processing when consuming media (Rogers, 1975, 1983).

A central question Protection Motivation Theory, or PMT, addresses is: why do people fail to engage in behaviors meant to protect them from unfavorable outcomes? PMT has been applied successfully in interventions regarding for example health-related protective behaviors (e.g., Floyd et al., 2000), environmental-friendly behavioral intentions (Rainear & Christensen, 2017), and protective behavior against personalized advertisements (Strycharz et al., 2019). Yet, while most traditional interventions attempt to motivate people to engage in protective behavior (e.g., check multiple sources), PMT has been neglected by the literature on traditional ML interventions regarding problems. As a result, traditional interventions might not be as effective as they can be.

An effective media literacy intervention is not only effective immediately after exposure, but also on the longer-term. There are only few studies that test the longer-term effectiveness of their traditional media literacy interventions (Lu et al., 2024). In this two-wave experimental study, we will compare the immediate and medium-term effectiveness of a novel theory-based heuristic intervention to a traditional intervention and a control group by answering the following overarching key question:

To what extent can a theory-based heuristic media literacy intervention (vs. traditional vs. control) affect accuracy in discernment, confidence in judgment, and efficacy perceptions, and do effects persist in the medium term?

Theoretical framework

The literature on traditional interventions is expanding rapidly but needs a theoretical basis that can help explain how interventions can motivate people to engage in the

suggested behavior. According to PMT, a successful intervention conveys four elements: the magnitude of noxiousness (i.e., severity) of an event, the probability of occurrence, the efficacy of the response, and the extent to which the individual is efficacious in successfully carrying out that response (Rogers, 1983).

The media literacy literature does sparingly contain some (but not all four) elements of the model put forward in PMT (Rogers, 1983). For example, van der Meer and Hamelers (2021; A4d) refer in their intervention to the noxiousness of a one-sided media diet: ‘When people see only one side of an issue, it creates a highly polarized political and social environment, fueling extremism and hatred of ‘The Other Side.’ The second element, probability of occurrence, can also at times be found in media literacy interventions: ‘We encounter a lot of manipulated or inaccurate news – which is even more the case for sensitive topics such as migration.’ (Hameleers, 2023, A-A). However, the third and fourth element (i.e., response-efficacy, and self-efficacy) are rarely present, nor are all four elements together ever present in traditional interventions. As a result, traditional interventions might not optimally convince the citizen to engage in the suggested protective behavior. Building on PMT, we argue that the effectiveness of ML interventions can be increased by conveying the four elements of a protective behavior appeal: how harmful misinformation can be (noxiousness), how likely people are to encounter it (probability of occurrence), what citizens stand to gain by engaging in the behaviors suggested in the media literacy intervention (response efficacy), and how competent they themselves are in carrying out the suggested behavior (self-efficacy).

Four elements of an effective ML intervention

Magnitude of noxiousness

ML interventions often imply the noxiousness (i.e., severity) of misinformation, as their existence suggests its undesirability. According to PMT, emphasizing an event’s noxiousness motivates protective behaviors. This is not to say that the magnitude of the noxiousness of an event needs to be exaggerated. Rather, research has identified a spillover effect where media literacy interventions increase skepticism across the board, including toward correct information (Hameleers, 2023; Hoes et al., 2024; van der Meer et al., 2023).

Potentially, too strongly emphasizing the magnitude of noxiousness contributes to this spillover effect. For example, van der Meer et al. (2023, A-H) found a spillover effect resulting from a study that used an intervention warning that the impact of false information can be ‘extremely destructive’. However, spillover effects were also found in ML interventions that did not refer to the noxiousness of misinformation (Hameleers, 2023; Hoes et al., 2024), which suggests that a strong formulation of the magnitude of noxiousness is not the only driver of spillover effects. Regardless, according to PMT, it is crucial to highlight the magnitude of noxiousness of an event.

Probability of occurrence

According to PMT (Rogers, 1975, 1983), highlighting the probability of occurrence increases people’s expectance of exposure, which in turn contributes to the motivation to engage in the recommended behavior. Research suggests that the average citizen is unlikely to be exposed to misinformation (Acerbi et al., 2022). At the same time,

misinformation peaks around key events such as elections, and divisive issues (Zannettou et al., 2018). If our aim is to reduce spillover effects yet increase the likelihood of people engaging in protective behaviors, a crucial step is to convince people that there is a possibility, however small, that they will encounter misinformation. It is tempting to argue that exaggerating the probability of encountering or falling for misinformation contributes to spillover effects. After all, if you tell people that misinformation is omnipresent, it is not surprising that they become more skeptical of *all* information. However, even studies that downplay, or explicitly refer to the *unlikelihood* of misinformation exposure find evidence of spillover effects (Hoes et al., 2024). These paradoxical findings suggest that any warning that mentions the probability of misinformation, or the probability of encountering/falling for misinformation, however delicate the formulation, increases skepticism toward all information. However, as we will argue below, the spillover effects might also be the result of traditional interventions' complexity, and their demanding nature that undermines the efficacy perceptions.

Response efficacy

When people see an intervention as effective in addressing a problem, they are more likely to believe in its efficacy. For instance, highlighting misinformation's noxiousness and encouraging source verification can convince individuals of its value as a coping strategy. This, in theory, would also mean that their motivation to engage in the proposed behavior increases. However, similar to the two elements mentioned above, current media literacy interventions rarely attempt to convince people of the effectiveness of the proposed response (exceptions are Hameleers, 2023; Lewandowsky & Yesilada, 2021).

Moreover, ML interventions do not often propose one specific intervention behavior, but multiple. For example, Hameleers (2023) suggests three behaviors, Lutzke et al. (2019), and Soetekouw and Angelopoulos (2024) suggest four behaviors, Lewandowsky and Yesilada (2021) suggest six behaviors, Hoes et al. (2024) suggest eight behaviors, and Guess et al. (2020) suggest ten behaviors. It is challenging for the target audience to remember all these suggestions, which makes it unlikely that people will process these suggestions sufficiently to render them effective. In sum, rarely do media literacy interventions attempt to emphasize why a proposed behavior is effective, while often suggesting multiple behaviors. As a result, people are unlikely to perceive the proposed behaviors as efficacious in tackling the problem, which does not stimulate, and might even decrease, people's motivation to engage in the suggested behavior.

Self-efficacy

According to PMT (Rogers, 1983), people need to have confidence in their own capability to carry out the suggested response. However, current ML interventions are unlikely to convince people of their self-efficacy. The main reasons are that, first, suggested behaviors are overly rationalistic, and that, second, the suggestions are often too energy – and time-consuming to carry out.

First, media literacy interventions appeal to the rationality of their target audience. In doing so, media literacy interventions often contain a wealth of information. However, offering a lot of information does not equate to conveying this information effectively since people are limited information processors (Lang, 2000). Moreover, ML interventions, by virtue of their aim to increase literacy, often contain behavioral suggestions

that require a hefty dose of cognitive effort. For instance, people have been advised to reflect on whether ‘consequences are logically connected to causes’ (Hameleers, 2023, A-A), or to assess whether differences between the news consumer and others exist and whether the extent of differences is accurate (Lewandowsky & Yesilada, 2021), or to reflect on one’s own opinions and ‘check your biases’ (Hoes et al., 2024, A2), or to assess whether statements contain checkable facts (Merper et al., 2018). These suggestions are so cognitively taxing that one can expect that most people, most of the time, do not feel capable of carrying out the behaviors. As a result, people are arguably less inclined to carry out the behavior.

In addition to the cognitive load, a second problem pertaining to self-efficacy perceptions is the effort and time needed to carry out proposed behavior. For instance, studies ask participants to ‘look at other reports’ to corroborate information (Guess et al., 2020; A2), or to ‘Click away from the story to investigate the site, its trustworthiness and contact info’, and to ‘consult fact checkers’ (Hoes et al., 2024, A2). For most people, most of the time, these effortful behaviors would involve too much extra time and effort. In an experimental setting, people might be willing to devote their time and energy to these tasks. But when scrolling through social media, it is unlikely that people would do the same. Consequently, people are unlikely to have confidence in their capacity to carry out the suggested responses, which according to PMT is needed to motivate people to engage in the suggested behavior (Rogers, 1983).

Taken together, when viewing traditional interventions through the lens of PMT (Rogers, 1975; 1983) we see that existing interventions do not effectively implement the four elements needed to effectively motivate people to engage in a certain protective behavior. While noxiousness and probability of occurrence are sometimes present, traditional interventions lack the elements relating to response-efficacy and self-efficacy. Sometimes traditional interventions might even counteract potential efficacy perceptions by suggesting multiple cognitively demanding, and laborious behaviors.

The heuristic approach as an effective alternative

Response and self-efficacy are crucial elements of PMT (Rogers, 1983): people need to believe that the proposed response is effective and feel that they are able to carry out the suggested behavior, before they can be expected to carry out that behavior. To increase the likelihood of fulfilling these two PMT-prerequisites, the suggested behavior (i.e., the intervention) needs to be as easy to grasp as possible. The interventions that have been used (i.e., traditional interventions) do not meet those PMT-prerequisites because they are so complex, and are therefore unlikely to result in protection motivation behavior, while an intuitive and easy to process intervention has a higher chance to succeed. Thus, this study takes a heuristic approach to media literacy, which contains all four elements identified in PMT (Rogers, 1983).

Heuristics are mental shortcuts that help people make sense of the world in a ‘fast and frugal’ way (Gigerenzer & Todd, 1999, p. 29). As limited information processors (Lang, 2000), people rely on heuristics to allow us to make judgments in an economic and relatively accurate manner (Fiske & Taylor, 2008). Heuristics can be understood as rules-of-thumb (‘e.g., if a message is long, then it is strong’; Bellur & Sundar, 2014, p. 118), and typically they are easy to remember, retrieve and apply (Chen & Chaiken, 1999). Research

shows that people often rely on rules-of-thumb precisely because of their efficiency (Gigerenzer & Todd, 1999).

This study introduces a rule-of-thumb, embedded in a broader PMT-intervention, that the target audience can use as a heuristic to ascertain whether information is correct or incorrect. Contrary to traditional interventions (e.g., Guess et al., 2020; Hameleers, 2023; Hoes et al., 2024; Lewandowsky & Yesilada, 2021; Lutzke et al., 2019; Merpert et al., 2018; Soetekouw & Angelopoulos, 2024), this heuristic should cost little effort or time to apply, and should also demand little cognitive effort. Indeed, by virtue of being a heuristic, it should be processed largely automatically.

People have an expectation of reality, and find information less credible when this expectation is violated. For instance, research suggests that people find deepfakes less credible when the person depicted in the deepfake makes statements they would normally not make (Hameleers et al., 2022, 2024a, 2024b). This study argues that media literacy interventions should appeal to people's intuitive 'nonsense detector'. As Hameleers et al. (2022, 2024a, 2024b) show, people have an expectation of reality, which is likely in part shaped by the three guiding principles of knowledge activation: availability, accessibility, and applicability (Chen & Chaiken, 1999). In other words, people are more likely to accept incoming information when this information is in line with expectations. But when this information deviates from expected reality, a 'nonsense detector' can be triggered (see also Truth-Default-Theory; Levine, 2014). When this deviation from expected reality occurs, this study proposes to offer people a rule-of-thumb that they can use in a 'fast and frugal' way to make judgments about the veracity of information (Gigerenzer & Todd, 1999, p. 29).

As such, this study, which has taken place in the Netherlands, proposes and tests the following rule-of-thumb: 'Voelt het raar? Dan is het niet waar!', which is a rhyming heuristic that translates to 'Does it feel odd? That means it is not true!'. Crucially, the rule-of-thumb is embedded in broader PMT-based intervention that instructs participants to relate incoming information to their existing knowledge, while emphasizing the four elements of PMT as follows: (1) probability of occurrence and (2) magnitude of noxiousness (i.e., Everyone sometimes falls for fake news. This is unfortunate, because it means you base your perceptions on wrong information). (3) response-efficacy (i.e., with the following rule-of-thumb, people are capable of recognizing fake news). (4) self-efficacy (several appeals, but for example: we ourselves often have a sense of what is true and what is odd; and: trust yourself and apply the rule-of-thumb when you read or hear something odd). Appendix A contains the original and translated heuristic media literacy intervention.

Considering that traditional interventions seem to be less effective than more interactive interventions in increasing accuracy in misinformation discernment (Lu et al., 2024), and since a PMT-based intervention would also be expected to impact confidence in assessment as well as perceptions of response and self-efficacy, we have formulated the following hypotheses.

H1. The heuristic intervention is most effective in increasing accuracy in misinformation discernment (compared to the traditional intervention and the control group).

H2. The heuristic intervention is most effective in increasing the confidence in accuracy judgments (compared to the traditional intervention and the control group).

H3. The heuristic intervention is most effective in increasing people's (a) response-efficacy and (b) self-efficacy perceptions (compared to the traditional intervention and the control group).

The over-time effects of ML interventions

A meta-analysis by Lu et al. (2024) shows that all experimental studies testing traditional interventions measure the effects directly after exposure. These experiments have been valuable, yet there is a need to study the extent to which effects persist over time because ML interventions aim to foster longer-term resilience to misinformation beyond the confines of the experiment. Thus, measuring only the immediate effects gives us limited insights on the extent to which resilience to misinformation persists in the real world. This study compares medium-term effects of the heuristic intervention and the traditional intervention within and between groups. Because the current study's heuristic approach fulfills all four PMT-elements aimed at increasing the likelihood of protective behavior, with a special emphasis on the two efficacy criteria, we expect that the effects of the heuristic intervention persist in the medium term (after 2 days). Because traditional interventions are limiting in terms of complexity, cognitive load and required time-effort, we expect that these effects do not persist in the medium term.

H4. The heuristic interventions effects on (a) accuracy in discernment, (b) accuracy confidence, (c) response-efficacy and (d) self-efficacy perceptions persist after two days, while the effects of the traditional media literacy interventions disappear after two days.

Method

We conducted a pre-registered two-wave online experiment.¹ The study design was approved by the ethical committee of [blinded for peer-review]. At T1, survey company Panelclix recruited 626 participants. Of these participants, 10 did not consent to data collection and were screened out prior to data collection. 85 participants failed the attention check. After the attention check, we asked two analytical thinking questions with four answer options. We kept the participants who failed the attention check but correctly answered both analytical thinking questions ($N = 9$), as this indicated that the participants did pay attention (i.e., we removed 76 participants who failed the attention check). We also removed 36 speeders, who completed the experiment within 200 s. This left us with a sample at T1 of $N = 504$. For T2, we re-invited only the participants who were part of the experimental conditions at T1, but not the participants from the control condition (because of resource constraints). At T2, Panelclix re-invited 357 people, of which 331 participants responded. Of those 331, 4 participants did not answer the questions measuring the dependent variables and were removed. Finally, 5 participants completed the survey at T2 in less than 60 s, and these speeders were also removed. This leaves us with a retention rate of 90%.

T1 was collected on 17 June, 2024. Participants of the experimental groups were re-invited for T2 48 hours later, on 19 June. Most participants ($N = 251$; 78%) completed the T2-survey on 19 June. Some participants ($N = 21$; 7%) completed the T2-survey on 20 June, or on 21 June ($N = 13$; 4%). The remainder ($N = 5$; 2%) completed the T2-survey no later than 24 June 2024.

The mean age was 48.90 ($SD = 15.90$). Regarding gender, the 50% of the sample identified as man ($N = 221$), 49% as woman ($N = 218$) and <1% as non-binary ($N = 2$). In terms of education, 19% of the sample comes from a 'low' or practically-oriented educational background ($N = 78$), 50% from a 'middle-level' educational background ($N = 219$), and 33% come from a 'high' or theory-oriented background ($N = 144$).

Design and stimuli

This pre-registered experiment consisted of two treatment conditions and one control condition. Participants were randomly assigned to one of those conditions. Participants in the treatment conditions saw one of two media literacy interventions at T1, a heuristic or a traditional intervention (see Appendix A). Participants in the heuristic treatment condition at T1 saw a media literacy intervention that contains a rule-of-thumb, serving as a heuristic, as well as all four elements specified in Protection Motivation Theory (Rogers, 1983). The rule-of-thumb is the following rhyming heuristic: 'Voelt het raar? Dan is het niet waar!' ('Does it feel odd? That means it is not true!'). Probability of occurrence is cued by the sentence 'Everyone sometimes falls for fake news'. Magnitude of noxiousness is cued by 'This is unfortunate, because it means you base your perceptions on wrong information.' Response efficacy is cued as follows 'with the following rule-of-thumb, people are capable of recognizing fake news'. Self-efficacy is cued by multiple sentences 'we often have a sense of what is true and what is odd'; and 'trust yourself and apply the rule-of-thumb when you read or hear something odd', and '(...) we already have a lot of knowledge and have a clear image of what we can expect'.

The other treatment group saw a traditional media literacy intervention at T1, that was based on Hameleers (2023). It is a rationalistic intervention that did not include appeals to intuition, or any element of PMT (Rogers, 1983). Similar to many traditional interventions, the intervention explains that people should check the source, search for the facts, and asks people to reflect on the extent to which consequences are logically connected to causes (see Appendix A). The control condition at T1 was exposed to a news article about Formula 1 (see Appendix A). After exposure, we measured the dependent variables.

At T2, the participants from both experimental conditions (heuristic, traditional) were invited to the second survey (48 h after the invite for the first survey). There was no stimulus, but the survey did measure accuracy and confidence in judgment once more (using different news articles than at T1), as well as the remaining key dependent variables (i.e., response efficacy, self-efficacy). The control group was only invited to T1. This means that the comparison of both experimental groups' T2 scores is benchmarked against the T1 control group scores.

Dependent variables

Accuracy in discernment was measured based on Chan et al. (2024), who made Facebook-style news posts that were either true or false (see Appendix B). They based this on information from factcheckers. Participants randomly saw real and false posts from a larger pool, and measured accuracy on a 5-point scale. In our current study we took a similar approach but did not randomize the news posts from a larger pool (i.e., all participants

were exposed to all news posts, but they were exposed in random order). We showed participants seven news posts and asked them to rate the veracity of each news post. Each participant rated the same news posts, to exclude the risk that some conditions rated (slightly) more realistic posts than other conditions. Accuracy in discernment was measured at T1 using seven news posts, and at T2 also using seven news posts, different from those at T1. At T1, four out of seven news posts were false. At T2, five out of seven were false. We changed the base rate from 4/7–5/7 because we informed our participants at the T1-debriefing that four out of seven posts were false. We feared that keeping the base rate at 4/7 could bias the results through a learning effect.

For each news posts, participants were asked *to what extent is the information in this news post correct, do you think?* In line with Chan et al. (2024), participants could answer *certainly not correct, probably not correct, I am not sure, probably correct, certainly correct*. For a true news post, there were two ‘correct’ answering options: ‘probably correct’ and ‘certainly correct’. For a false news post, the correct answering options were ‘probably not correct’ and ‘certainly not correct’. The variable was recoded so a correct answer was scored with a 1, and an incorrect answer with a 0. Since there were seven accuracy questions, 7 would be a perfect accuracy score, and 0 would be the worst accuracy score. At T1, the average accuracy score was $M = 4.53$ ($SD = 1.42$). At T2 the average accuracy score (which does not include the control group) is $M = 4.04$ ($SD = 1.26$).

Confidence in accuracy judgment was measured on the basis of the five-point scale that also measured accuracy in discernment. Participants could indicate how certain they were of their accuracy assessment (i.e., *certainly not correct, probably not correct, I am not sure, probably correct, certainly correct*). The ends of the scale signal the most confidence, even if the answer is wrong. The midpoint signals the least confidence. The options probably (not) correct signal some confidence. This measurement was transformed into a 3-point scale, that was the average score of all seven accuracy measurements at each point in time. This variable at each timepoint took per participant seven times their accuracy score, and each time gave 1 point for the answer ‘*I am not sure*’, 2 points for the answers ‘*probably (not) correct*’, and 3 point for the answers ‘*certainly (not) correct*’. These answers were averaged (divided by 7) to form a scale that ranged from 1 to 3. At T1, the average confidence score was $M = 2.05$ ($SD = 0.33$). At T2, the average confidence score was $M = 2.06$ ($SD = 0.34$).

Response efficacy was measured with the following two items measured on a 7-point scale. *The tips for recognizing fake news are easy to apply*, and *the tips for recognizing fake news put me in a better position to spot fake news*. Both items correlated strongly (Pearson’s $r = .63$ at T1, and Pearson’s $r = .60$ at T2). The average score at T1 was $M = 4.68$ ($SD = 1.07$). The average score at T2 was $M = 5.05$ ($SD = 1.17$). Response efficacy was not measured in the control group since the control group was not exposed to tips for recognizing fake news.

Self efficacy was measured on the basis of Rasmussen et al. (2024) with the following two items, on a 7-point scale. *It is easy for me to recognize fake news*, and *I am able to distinguish fake news from reliable news*. Both items correlated strongly (Pearson’s $r = .67$ at T1, and T2). The average score at T1 was $M = 4.20$ ($SD = 1.05$). The average score at T2 was $M = 4.15$ ($SD = 1.07$).

Procedure

Upon entering the survey, participants agreed to the informed consent (10 did not consent, and their participation was immediately terminated). Then, participants were asked questions about their background (age, gender, education) and saw an attention check, as well as two items measuring analytical thinking. Participants were then randomly assigned to either the control condition, or the heuristic condition, or the traditional condition (see Appendix A). After that, we measured participants' accuracy in discernment by showing them a total of seven news posts and asking participants to rate the veracity of each news post. Thereafter, we measured the remaining dependent variables, and debriefed the participants. Exactly 48 hours after their invitation for T1, participants were invited for T2. At T2, all participants first agreed to the informed consent and were then asked to rate the accuracy of seven news post (different posts than at T1). After that, we measured the remaining dependent variables and debriefed the participants.

Randomization checks

We conducted randomization checks on age, gender, and education, and found that randomization for these variables was successful. A one-way ANOVA revealed no significant differences regarding age between conditions: ($F(2, 502) = .28, p = .76$). A χ^2 test revealed no significant differences regarding education ($X^2(4) = 4.92, p = .30$). A Fisher-Freeman-Halton test revealed no significant differences regarding gender ($p = .49$).

Results

We ran one-way ANOVAs with repeated measures with a within and between subjects focus, to test the hypotheses. H1 posited that the heuristic intervention is most effective in increasing accuracy in misinformation discernment (compared to the traditional intervention and the control group). As [Figure 1](#) displays, at T1, the heuristic condition scores significantly higher ($M = 4.77, SD = 1.23$) than the control condition ($M = 4.40, SD = 1.38$) and the traditional condition ($M = 4.43, SD = 1.61$) on accuracy in misinformation discernment: ($F(2, 438) = 3.17, p = .04$). The generalized η^2 is .014, which suggests a small effect. However, running a one-way ANOVA with Holm correction reveals a borderline significant difference between conditions ($p = .052$). Running a series of one-sided t-tests with Holm correction for pairwise comparison also reveals a significant difference between the heuristic condition and the control condition (adjusted $p = .034$), but borderline significant for the heuristic condition and the traditional condition ($p = .061$). Considering the above, we conclude that H1 is partially supported. The heuristic condition displayed significantly higher accuracy in misinformation discernment than the control, but not the traditional condition.

Hypothesis 4a held that the effect on accuracy in discernment persisted over time. This is not the case, as the right-hand panel of [Figure 1](#) illustrates. At T2, participants in the heuristic condition ($M = 4.07, SD = 1.19$) scored significantly worse than their own score at T1 ($M = 4.77, SD = 1.23; p < .001$), and also significantly lower than the control group at T1 ($M = 4.40, SD = 1.38; p = .03$). At T2, the heuristic condition had a similar score as

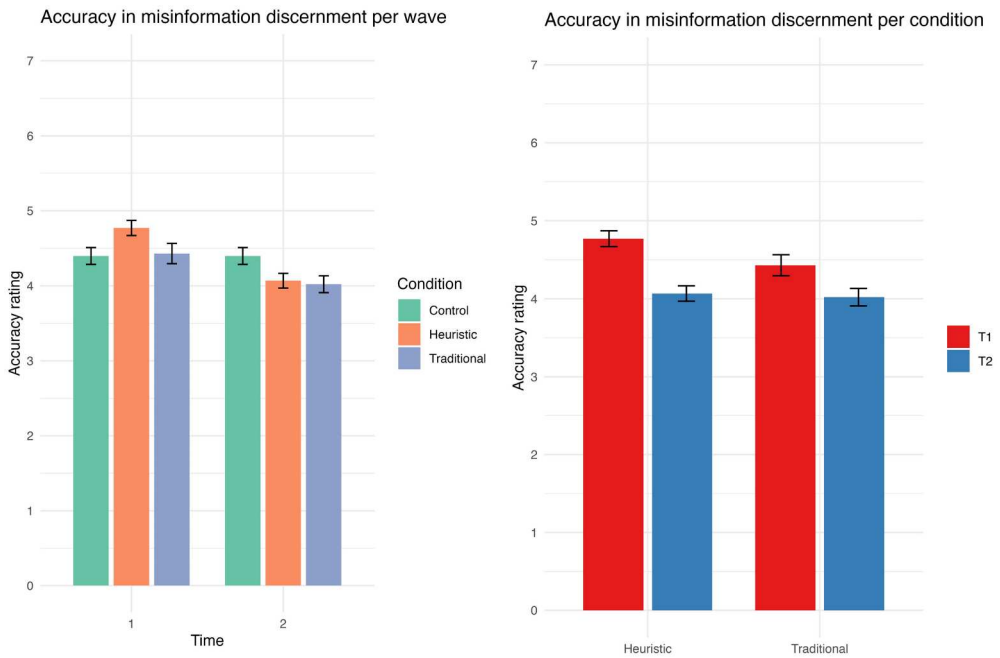


Figure 1. Accuracy in misinformation discernment.

Note: left-hand panel shows between-effect per wave. Right-hand panel shows within-effect. The control group was only measured at T1.

the traditional condition ($M = 4.02$, $SD = 1.33$, $p = .38$). This means that H4a is not supported.

Does the heuristic intervention increase confidence in accuracy judgment?

Hypothesis 2 expected that participants in the heuristic condition would be more confident in their accuracy judgments. Scores could range between 0 and 3, and were the average of seven judgments per wave, per participant. The heuristic condition scored on average $M = 2.10$ ($SD = .32$), slightly higher than the traditional condition $M = 2.03$ ($SD = .35$), and the control condition $M = 2.01$ ($SD = .32$). A repeated-measures ANOVA revealed no significant main effects of the condition ($F(2, 438) = 1.99$, $p = .14$), or time ($F(1, 438) = .14$, $p = .71$). There was also no significant interaction effect of condition and time ($F(2, 438) = 2.06$, $p = .13$). Indeed, the scores at T2 remained stable: heuristic condition ($M = 2.06$; $SD = .34$) scored similar to the traditional condition ($M = 2.05$; $SD = .35$), and slightly higher than the control condition ($M = 2.01$; $SD = .32$). In conclusion, and as shown by Figure 2, H2 is not supported, and H4b (which held that the effect would persist) is not supported.

Is response efficacy highest for the heuristic intervention?

H3a expected that people would perceive higher response efficacy for the heuristic intervention than the traditional intervention. For the analysis, we only compare the

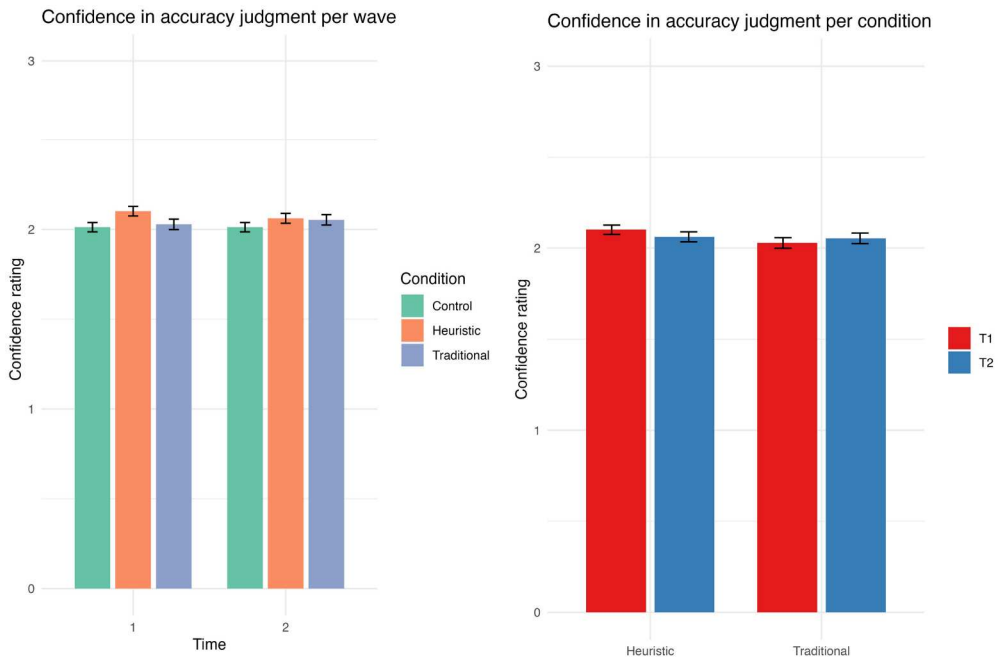


Figure 2. Confidence in accuracy judgment.

Note: left-hand panel shows between-effect per wave. Right-hand panel shows within-effect. The control group was only measured at T1.

participants exposed to an intervention, so we exclude the control condition. We also exclude those who indicated that they did not remember the response tips ($N = 9$). At T1, the heuristic intervention ($M = 4.68$; $SD = 1.17$) scores similar to the traditional intervention ($M = 4.69$; $SD = 1.00$). A repeated-measured ANOVA shows no significant main effect of condition ($F(1, 253) = .48$, $p = .49$), which means that H3a is not supported (see the left-hand panel of Figure 3).

Testing H4c, which expects that effects persist over time, we see that the heuristic condition at T2 ($M = 5.03$; $SD = 1.09$) scores higher than the traditional condition at T2 ($M = 4.88$; $SD = 1.00$). The repeated-measured ANOVA shows a significant direct effect of time on response efficacy ($F(1, 253) = 12.46$, $p < .001$). The generalized η^2 is .012, which suggests a small effect (see the right-hand panel of Figure 3). A pairwise comparison with Holm correction shows that the heuristic condition scored significantly higher on response efficacy at T2 than at T1 ($p = .01$). While the traditional condition also scored higher at T2 than at T1, this difference was not significant ($p = .12$). Comparing both conditions at T2, a pairwise comparison with Holm correction shows that both conditions do not differ significantly ($p = .27$). While the results suggest a lagged effect of the heuristic intervention on perceived response efficacy (see the right-hand panel of Figure 3), this was not part of H4c. Since H4c expected that effects would persist, rather than materialize at T2, we conclude that H4c is not supported.

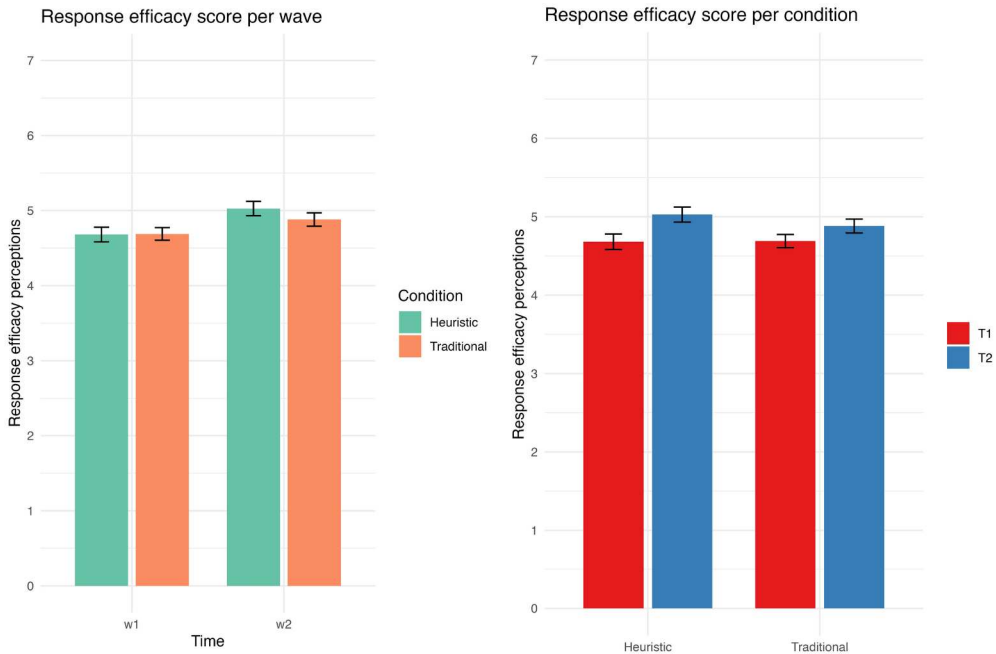


Figure 3. Perceived response efficacy.

Note: left-hand panel shows between-effect per wave. Right-hand panel shows within-effect.

Is self-efficacy highest for the heuristic intervention?

H3d stated that people in the heuristic condition would hold higher self-efficacy perceptions at T1, compared to the traditional and control conditions. As the left-hand panel of Figure 4 illustrates, there are only slight differences between conditions at T1. The heuristic condition ($M = 4.31$; $SD = 1.06$) scores higher than the traditional ($M = 4.20$; $SD = 1.01$) and control condition ($M = 4.10$; $SD = 1.08$). A repeated-measures ANOVA reveals no significant main effect of condition on self-efficacy perceptions: ($F(2, 438) = .92, p = .40$). When using an uncorrected one-sided t-test to zoom in on the expected difference between the heuristic condition and the control condition, we first find a marginally significant difference ($p = .046$), but after a Holm correction this difference becomes non-significant ($p = .14$).

Hypothesis 4d stated that the heuristic intervention effects on self-efficacy expected by H3d would persist. Turning to the within-subjects effects, the data show that compared to T1 the scores at T2 have slightly decreased for the heuristic condition ($M = 4.20$; $SD = 1.11$), and traditional condition ($M = 4.10$; $SD = 1.03$). The self-efficacy perceptions of the control condition were only measured at T1. The repeated-measures ANOVA shows a borderline significant main effect of time on self-efficacy perceptions ($F(1, 438) = 3.85, p = .05$). However, after a pairwise comparison with Holm correction the differences between T1 and T2 turn out to be non-significant for both the heuristic condition ($p = .21$) and the traditional condition ($p = .14$; see right-hand panel of Figure 4). As such, hypotheses H3d and H4d are not supported.

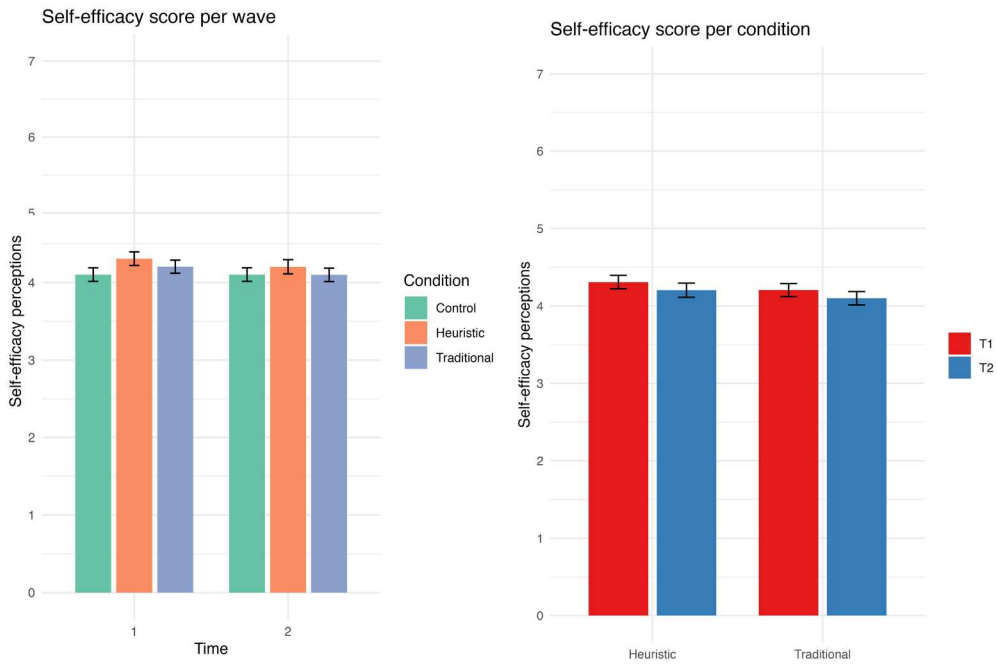


Figure 4. Self-efficacy perceptions.

Note: left-hand panel shows between-effect per wave. Right-hand panel shows within-effect. The control group was only measured at T1.

Discussion

This study set out to test the immediate and medium-term effectiveness of a theory-based heuristic ML intervention, and compare this to a traditional intervention. This study proposed that the media literacy literature could learn from Protection Motivation Theory (Rogers, 1975, 1983), by offering people an intervention that conveys four elements: noxiousness, probability, response efficacy, and self-efficacy. Moreover, in contrast to the often rationalistic and complex existing ML interventions, this study offered people a heuristic, a rule-of-thumb, as a key-takeaway message that they could use to efficiently assess incoming information.

The heuristic intervention shows promise in increasing accuracy in discernment at T1, by significantly outperforming the control condition and marginally significantly outperforming the traditional intervention. At the same time, the effect size is small, and the effect seems to disappear at T2. On the one hand this suggests that the heuristic approach is by no means a panacea. On the other hand, persisting accuracy in discernment would be a hard needle to move with a novel and single-shot intervention.

There is little knowledge about the short-to-medium term effectiveness of media literacy interventions (Lu et al., 2024). The results of this study paint a sobering picture: the positive effects decrease between T1 and T2 (> 48 hours). In fact, the two experimental groups score significantly worse at T2 than the control group did at T1. While this decrease is substantial, we cannot say for certain that the effects disappear completely because we have no measurement of the control group at T2. For instance, an important

driver of the decrease could be that the measurement was more challenging at T2 than at T1: discernment could have been easier for the first set of seven news posts, than for the second set. However, where the heuristic condition outperformed the traditional condition in accuracy at T1 by a marginally significant margin ($p = .061$), both groups scored similarly at T2. This at least suggests that the positive discernment effects of the heuristic intervention, compared to the traditional intervention, disappeared in the medium term, regardless of the complexity of the measurement of T2, suggesting that repeated exposure to the intervention is needed to reinforce learning effects.

This study expected that people in the heuristic condition would have stronger efficacy perceptions (i.e., response and self), but found no significant differences at T1. The heuristic condition scored significantly higher on response efficacy perceptions at T2 than at T1 which might suggest a lagged effect of the intervention. However, seeing as we also saw a (non-significant) increase in the traditional condition, it could also be that the remaining participants at T2 were simply more engaged with the issue of media literacy and this correlates with their efficacy-perceptions.

When it comes to self-efficacy perceptions, the results show that the heuristic condition trends upwards at T1 but decreases relatively strongly at T2. Relatedly, confidence in judgment was highest for the heuristic condition, but did not differ significantly between conditions and over time. While the intervention contained different self-efficacy cues, also meant to increase confidence, it seems that self-efficacy perceptions are too stable to be affected by this single-shot intervention.

Implications and limitations

The results raise questions about the longer-term effectiveness of (single shot) media literacy interventions: it could be that single shot media literacy interventions are not effective beyond the immediate moment. At the same time, this one study is limited in several ways. First, according to PMT (Rogers, 1975, 1983), people should be motivated to engage in a certain behavior. This study did not measure motivation, and so we cannot say for sure whether the intervention was effective in motivating people to engage in the behavior. In other words, while the heuristic intervention had strong self-efficacy cues, it might have needed stronger cues regarding the remaining PMT-elements (i.e., noxiousness, probability and response efficacy) to motivate people to adapt their behavior for a longer time-period. In a similar vein, most PMT-based interventions contain a relatively strong fear appeal (Rogers, 1975). This intervention likely did not arouse fear in the participants, which might have affected the effectiveness of the interventions. One might ask to what extent fear-based media literacy interventions are desirable, since fear-based appeals might convey an unrealistic (too negative) perspective about the threat of misinformation which might induce spillover effects (Acerbi et al., 2022).

Our measurement of accuracy in discerning misinformation is narrow, focusing solely on participants' ability to evaluate Facebook-style news posts, similar to Chan et al. (2024; Appendix B), limiting the findings to this specific context. Additionally, these posts lacked clear source cues, forcing participants to engage directly and solely with the content. While narrow, our study's findings suggest that it is worthwhile to test the effectiveness of heuristic media literacy interventions in a broader context.

It is important to note that while the intervention follows from theory (Lang, 2000; Rogers, 1975; Tversky & Kahneman, 1974) and builds upon prior ML literature (e.g., Guess et al., 2020; Hameleers, 2023; Hoes et al., 2024; Lewandowsky & Yesilada, 2021; Lutzke et al., 2019; Merpert et al., 2018; Soetekouw & Angelopoulos, 2024), it is by no means perfect. It is possible or even likely that, if we tell citizens to rely on their internal nonsense detector when in doubt, there will be systematic biases in what citizens recognize as ‘not true’. Such biases will partly be explained by motivated reasoning processes. Moreover, research on the ‘illusory truth effect’ (e.g., Pennycook et al., 2018) shows that familiarity can trump critical evaluation. Therefore, while our heuristic aims to cue intuitive doubt, it must be embedded in a broader (theory-based) context to minimize adverse effects. Future research should explore how to calibrate such heuristics to minimize confirmation bias while retaining their simplicity, for instance by adding the word ‘probably’ to the rule-of-thumb.

This study makes a first attempt to understand the effectiveness of heuristic ML interventions. While medium-term effects are inconclusive, heuristic cues outperform traditional ones in the short term. With PMT’s success in other fields, further research is encouraged to overcome the study’s limitations and enhance media literacy strategies.

Note

1. https://osf.io/hcety/?view_only=e7ad83b52379449a8435c08063988296

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Tom Dobber is assistant professor at the Amsterdam School of Communication Research (ASCoR), University of Amsterdam.

Dina Strikovic is assistant professor at the Amsterdam School of Communication Research (ASCoR), University of Amsterdam.

Marina Tulin is assistant professor at the Amsterdam School of Communication Research (ASCoR), University of Amsterdam.

ORCID

Tom Dobber  <http://orcid.org/0000-0002-6657-4037>

Edina Strikovic  <http://orcid.org/0000-0001-8377-5001>

References

- Acerbi, A., Altay, S., & Mercier, H. (2022). Fighting misinformation or fighting for information? *Harvard Kennedy School Misinformation Review*, <https://doi.org/10.37016/mr-2020-87>
- Basol, M., Roozenbeek, J., & Linden, S. v. d. (2020). Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of Cognition*, 3(1), 2. <https://doi.org/10.5334/joc.91>

- Bellur, S., & Sundar, S. S. (2014). How can we tell when a heuristic has been used? Design and analysis strategies for capturing the operation of heuristics. *Communication Methods and Measures*, 8(2), 116–137. <https://doi.org/10.1080/19312458.2014.903390>
- Chan, M., Vaccari, C., & Yamamoto, M. (2024). Examining the relationship between dispositional news literacy and discernment of real and misleading news. *International Journal of Public Opinion Research*, 36(2), <https://doi.org/10.1093/ijpor/edae020>
- Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York, NY: The Guilford Press.
- Droog, E., Vermeulen, I., Van Huijstee, D., Harutyunyan, D., Tejedor, S., & Pulido, C. (2024). Combatting the misinformation crisis: A systematic review of the literature on characteristics and effectiveness of media literacy interventions. Pre-Print. <https://osf.io/preprints/osf/ympmur>.
- Fiske, S., & Taylor, S. (2008). *Social cognition: From brains to culture*. Boston, MA: McGraw-Hill Higher Education.
- Floyd, D. L., Prentice-Dunn, S., & Rogers, R. W. (2000). A meta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology*, 30(2), 407–429. <https://doi.org/10.1111/j.1559-1816.2000.tb02323.x>
- Gigerenzer, G., Todd, P., & ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford University Press.
- Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27), 15536–15545. <https://doi.org/10.1073/pnas.1920498117>
- Hameleers, M. (2023). The (Un)Intended consequences of emphasizing the threats of mis- and disinformation. *Media and Communication*, 11(2), <https://doi.org/10.17645/mac.v11i2.6301>
- Hameleers, M., van der Meer, T., & Dobber, T. (2022). You won't believe what they just said! The effects of political deepfakes embedded as Vox Populi on social media. *Social Media + Society*, 8(3), 20563051221116346. <https://doi.org/10.1177/20563051221116346>
- Hameleers, M., van der Meer, T., & Dobber, T. (2024a). Distorting the truth versus blatant lies: The effects of different degrees of deception in domestic and foreign political deepfakes. *Computers in Human Behavior*, 152, 108096. <https://doi.org/10.1016/j.chb.2023.108096>
- Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2024b). They would never say anything like this! reasons to doubt political deepfakes. *European Journal of Communication*, 39(1), 56–70. <https://doi.org/10.1177/02673231231184703>
- Hoes, E., Aitken, B., Zhang, J., Gackowski, T., & Wojcieszak, M. (2024). Prominent misinformation interventions reduce misperceptions but increase scepticism. *Nature Human Behaviour*, 1–9. <https://doi.org/10.1038/s41562-024-01884-x>
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
- Levine, T. R. (2014). Truth-Default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33(4), 378–392. <https://doi.org/10.1177/0261927X14535916>
- Lewandowsky, S., & Yesilada, M. (2021). Inoculating against the spread of islamophobic and radical-islamist disinformation. *Cognitive Research: Principles and Implications*, 6(1), 57. <https://doi.org/10.1186/s41235-021-00323-z>
- Lu, C., Hu, B., Bao, M.-M., Wang, C., Bi, C., & Ju, X.-D. (2024). Can media literacy intervention improve fake news credibility assessment? A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 27(4), 240–252. <https://doi.org/10.1089/cyber.2023.0324>
- 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>
- Merpert, A., Furman, M., Anauati, M. V., Zommer, L., & Taylor, I. (2018). Is that even checkable? An experimental study in identifying checkable statements in political discourse.

- Communication Research Reports*, 35(1), 48–57. <https://doi.org/10.1080/08824096.2017.1366303>
- 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>
- 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>
- Rainear, A. M., & Christensen, J. L. (2017). Protection motivation theory as an explanatory framework for proenvironmental behavioral intentions. *Communication Research Reports*, 34(3), 239–248. <https://doi.org/10.1080/08824096.2017.1286472>
- Rasmussen, J., Lindekilde, L., & Petersen, M. B. (2024). Public health communication reduces COVID-19 misinformation sharing and boosts self-efficacy. *Journal of Experimental Political Science*, 1–16. <https://doi.org/10.1017/XPS.2024.2>
- Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude Change1. *The Journal of Psychology*, 91(1), 93–114. <https://doi.org/10.1080/00223980.1975.9915803>
- Rogers, R. W. (1983). Cognitive and physiological processes in fear-based attitude change: A revised theory of protection motivation. In J. Cacioppo & R. Petty (Eds.), *Social psychophysiology. A sourcebook* (pp. 153–176). New York, NY: The Guilford Press.
- Soetekouw, L., & Angelopoulos, S. (2024). Digital resilience through training protocols: Learning To identify fake news on social media. *Information Systems Frontiers*, 26(2), 459–475. <https://doi.org/10.1007/s10796-021-10240-7>
- Strycharz, J., van Noort, G., Smit, E., & Helberger, N. (2019). Protective behavior against personalized ads: Motivation to turn personalization off. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 13(2), 2. <https://doi.org/10.5817/CP2019-2-1>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- van der Meer, T., & Hamelers, M. (2021). Fighting biased news diets: Using news media literacy interventions to stimulate online cross-cutting media exposure patterns. *New Media & Society*, 23(11), 3156–3178. <https://doi.org/10.1177/1461444820946455>
- van der Meer, T., Hamelers, M., & Ohme, J. (2023). Can fighting misinformation have a negative spillover effect? *Journalism Studies*, 0(0), 1–21. <https://doi.org/10.1080/1461670X.2023.2187652>
- Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G., & Blackburn, J. (2018). Disinformation warfare: Understanding state-sponsored trolls on twitter and their influence on the web. *arXiv*, 218–226.