

RESEARCH ARTICLE 🕕 😳

Public Health Communication Reduces COVID-19 Misinformation Sharing and Boosts Self-Efficacy

Jesper Rasmussen⁽¹⁾, Lasse Lindekilde⁽¹⁾ and Michael Bang Petersen⁽¹⁾

Department of Political Science, Aarhus University, Aarhus, Denmark Corresponding author: Jesper Rasmussen; Email: jr@ps.au.dk

Abstract

During health crises, misinformation may spread rapidly on social media, leading to hesitancy towards health authorities. The COVID-19 pandemic prompted significant research on how communication from health authorities can effectively facilitate compliance with health-related behavioral advice such as distancing and vaccination. Far fewer studies have assessed whether and how public health communication can help citizens avoid the harmful consequences of exposure to COVID-19 misinformation, including passing it on to others. In two experiments in Denmark during the pandemic, the effectiveness of a 3-minute and a 15-second intervention from the Danish Health Authorities on social media was assessed, along with an accuracy nudge. The findings showed that the 3-minute intervention providing competences through concrete and actionable advice decreased sharing of COVID-19-related misinformation and boosted their sense of self-efficacy. These findings suggest that authorities can effectively invest in building citizens' competences in order to mitigate the spread of misinformation on social media.

Keywords: Misinformation; intervention; experiment; COVID-19; public health communication

Introduction

Misinformation about COVID-19 on social media has been a public concern during the COVID-19 pandemic. How can public health authorities communicate to mitigate the spread of misinformation? One line of research on countering COVID-19 misinformation suggests that subtly nudging people to think about accuracy reduces misinformation sharing (Pennycook et al. 2020). The strength of such interventions is their fast and frugal nature. At the same

This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

[©] The Author(s), 2024. Published by Cambridge University Press on behalf of American Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http:// creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

time, they are premised on the idea that people can discern between true and false, but "that when deciding what to share on social media, people are often distracted from considering the accuracy of the content" (Pennycook et al. 2021). As such, a key drawback of accuracy nudge interventions is that they simply remind people about accuracy, but leave individuals to rely on their own knowledge without providing tools or competences to deal with misinformation. The frugal nature of accuracy nudges also implies that their effects are small (Pennycook and Rand 2022), making recent studies question their effectiveness during the COVID-19 pandemic (Rathje et al. 2022; Roozenbeek et al. 2021; Gavin et al. 2022; Pretus et al. 2023).

Other interventions move beyond nudging accuracy motivations by seeking to equip people with better capabilities for identifying and avoiding sharing misinformation (Lee 2018; Guess et al. 2020; Hertwig and Grüne-Yanoff 2017; Sheeran and Orbell 2000; Sheeran et al. 2007; Van Der Linden 2022). This is consistent with research on risk communication that argues that feelings of competence are key in order to motivate people to respond effectively to risks by engaging in protective behaviors (Jørgensen et al. 2021b). People respond effectively when they are provided with trustworthy information about a threat, provided with actionable advice on how to respond to the threat and assured that this response will be efficient against the threat (Rogers 1975; Maddux and Rogers 1983; Rippetoe and Rogers 1987).

In this manuscript, we test the effectiveness of interventions that nudge accuracy to ones that additionally provide capabilities. Specifically, we test an accuracy nudge as well as two video-based, real-world interventions circulated by the Danish National Health Authority on social media in January 2021 during the COVID-19 pandemic. The accuracy nudge subtly primed people to think about their motivation to share accurate headlines, while the videos – a 15-second and a 3-minute intervention – provided capabilities through concrete instructions on how to avoid sharing COVID-19 related to misinformation.

For study 1, we predicted that all three interventions would decrease false headline sharing, increase real headline sharing, and increase sharing discernment (i.e., the relative sharing of real compared to false headlines). The accuracy nudge, the 15-second intervention and the 3-minute intervention all significantly increased sharing discernment. Only the 3-minute intervention, however, directly and significantly decreased false headline sharing, but did not alter real headline sharing. Neither the 15-second intervention nor the accuracy nudge had a statistically significant effect on either false or real headline sharing. Consistent with a capability perspective, study 2 showed that the 3-minute intervention increased participants' sense of self-efficacy in dealing with online misinformation. The intervention did not influence other aspects often highlighted in research on risk communication, specifically, participants' sense of the threat from misinformation and the effectiveness of remedies against misinformation. Overall, these results suggest that when health authorities communicate elaborate and actionable advice on how to avoid sharing COVID-19 misinformation, such communication can reduce the spread of false headlines and enhance people's sense of personal competence.

Two approaches to misinformation interventions

We examine two types of interventions to reduce misinformation sharing. One type of intervention is "accuracy nudges" which has received significant research interest (Pennycook et al. 2021,2020; Roozenbeek et al. 2021; Rathje et al. 2022; Gavin et al. 2022; Pretus et al. 2023). The psychological assumption behind accuracy nudges is that people are motivated to share accurate content on social media and are capable of distinguishing between true and false content. Yet, accuracy concerns do often not drive online sharing behavior because people are distracted from accuracy toward a desire to share emotionally engaging content and receive positive social feedback from friends on their sharing. Thus, reminding people to pay attention to accuracy through nudging should decrease the sharing of misinformation. Some research shows that subtle accuracy nudges where people are asked to rate the accuracy of a few news headlines decrease subsequent sharing of false headlines on social media (Pennycook et al. 2021, 2020).

Another type of intervention is capability interventions that seeks to mitigate misinformation sharing by building competences or resistance against the rhetorical techniques and strategies that are used to mislead people through misinformation (Roozenbeek and Van Der Linden 2022; Guess et al. 2020; Badrinathan 2021; Van Der Linden 2022; Lee 2018; Mo Jones-Jang et al. 2021). While these interventions are often rooted in distinct theoretical frameworks – such as inoculation theory or digital media literacy - the common denominator is that they go beyond merely priming accuracy motivations by providing concrete tools, actionable advice, or psychological competences to mitigate misinformation sharing. In other words, while the assumption behind accuracy nudges is that people already have the competence to avoid sharing misinformation and simply are in need of motivationoriented reminders, capability interventions go beyond the motivational component of accuracy nudges: They build capabilities by providing education (i.e., increasing knowledge) and training (i.e., imparting skills and tools) to avoid misinformation sharing. Where accuracy nudges rely on prompting a pre-existing motivation for accuracy, capability interventions aim to build reflective motivation where citizens contemplate and plan how to implement advice behaviorally.

Overview of studies

In two studies, we conduct a preregistered test of a 15-second and a 3-minute capability-oriented intervention from the Danish Health Authorities as well as an accuracy nudge (Pennycook et al. 2020) and compare them to a control group. Study 1 tested the effect of the interventions on sharing of false and real headlines while Study 2 assessed the effect of the interventions on self-efficacy, response efficacy, and threat appraisal.¹ Whenever we report additional analyses that were not preregistered, we label them as exploratory. Table 1 provides an overview of the data collection.

¹All materials are available on OSF.

Downloaded from https://www.cambridge.org/core. Berklee College Of Music, on 06 Feb 2025 at 20:30:07, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/XPS.2024.2

Tuble 11 Overview of data concerto	Table 1.	Overview	of data	collection
------------------------------------	----------	----------	---------	------------

Study	Sampling	Sample	Purpose of study
Pretest†	Politically diverse sample in Denmark through YouGov	205	Validate headlines for study 1 and study 2
1	National sample in Denmark through YouGovs panel. Quota sampling to match population characteristics on age, gender, region, and education. July 2021	Wave 1: 2541	Measure individual differences (trust, need for cognition, cognitive reflection task, attention to social comparison information) of participants prior to the experiment
1	National sample in Denmark through YouGovs panel. Quota sampling to match population characteristics on age, gender, region, and education. August 2021	Wave 2: 2232 (88% retention)	Do the interventions reduce false headline sharing? Do the interventions increase real headline sharing? Do the interventions increase sharing discernment?
2	National sample in Denmark through YouGovs panel. Quota sampling to match population characteristics on age, gender, region, and education. December 2021	2012	Do the interventions boost threat appraisal, self-efficacy, and response efficacy?

Note: [†]We conducted a pretest to validate the headlines for study 1 and study 2. Details can be found in Section C of the Supplementary Material.

In both studies, each participant was assigned to one of four conditions. In the control condition, participants were not exposed to any treatment prior to the respective dependent measures of Study 1 and Study 2. In the accuracy nudge condition, participants rated the accuracy of a single headline (unrelated to COVID-19) framed as a pretest mimicking prior studies of accuracy nudges (Pennycook et al. 2020; Roozenbeek et al. 2021). In the 15-second condition and the 3-minute condition, participants were shown videos titled "Can you trust what you read?" containing guidance on how to recognize and avoid sharing COVID-19 misinformation.² Specifically, the actionable advice in the video is summed up in three questions that one should ask oneself when facing novel information on social media: (1) Who is saying it? (2) How many are saying it? (3) Is the content too far out? Besides the 3-minute video being longer than the 15-second video, there are two major differences between the interventions. First, while the 15-second video only contains text, the 3-minute video includes both text and audio, which makes the content more immersive. Second, the 3-minute video provides more elaborate advice. In other words, the advice provided by the 3-minute intervention is more concrete and actionable in terms of providing a plan for implementation, which is conducive to behavior change (Pearce et al. 2019; Sommestad et al. 2015). Through collaboration with the Danish Health

 $^{^{2}}$ Participants were not able to skip the videos and 88% passed the attention check regarding the stimuli. We provide additional information and robustness tests regarding attention checks in section F of the supplementary material.

Downloaded from https://www.cambridge.org/core. Berklee College Of Music, on 06 Feb 2025 at 20:30:07, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/XPS.2024.2

Authority, we have been given permission to use the actual videos which were shared on Facebook in 2020 and 2021.³

The studies were conducted in Denmark during in the Summer and Winter of 2021 during the COVID-19 pandemic. Denmark is characterized by high levels of interpersonal and institutional trust and low levels of political polarization. These factors materialized during the onset of the pandemic, as compliance with and support for governmental responses were high, while polarization was low, in contrast to other countries where governmental responses were more disputed (Lindholt et al. 2021; Van Bavel et al. 2022; Jørgensenet al. 2021b,a). Furthermore, it was widely accepted, even among political elites, that the COVID-19 virus constituted a significant public health threat that required public collaboration and responsiveness to contain. As policies and messages aimed at countering the COVID-19 pandemic are more effective when backed by cross-partisan coalitions of political elites (Flores et al. 2022), public health communication is more likely to be effective in Denmark, compared to countries where the nature of the COVID-19 pandemic as a health crisis was disputed. We further elaborate on the implications for the generalizability of the findings in the discussion section.

Study 1

We preregistered the following hypotheses for study 1. We expected that the interventions would reduce false headline sharing, and thus, we predicted that compared to the control condition, all three interventions decrease sharing of false headlines, but not real headlines about COVID-19 on social media (H1). Conversely, the intervention could work by increasing real headline sharing, and thus, we predicted that compared to the control condition, all three interventions increase sharing of real headlines, but not false headlines about COVID-19 on social media (H2). Furthermore, the interventions might prompt a relative increase in real versus false headlines, and thus, we predicted that compared to the control condition, the interventions increase sharing discernment (H3).⁴ Furthermore, we preregistered a range of robustness analyses of the treatment effects across covariates. We predicted that the treatment effect of 15-second and 3-minute interventions on citizens' likelihood of sharing false headlines is lower for respondents who have low trust in public institutions and government handling of the pandemic and low scores on cognitive reflection and attention to social comparison information, compared to respondents with high scores on these variables (H4). Next, we predicted a significant interaction between attention to social comparison information and all of the three interventions compared to the control condition on the willingness to share both real and false headlines, such that the effect is stronger for people who score high on attention to social comparison information (H5), and, finally, that the effect of all

 $^{^{3}}$ The 15-second and 3-minute video interventions from the Danish Health Authorities are freely available online while full transcripts of the interventions are included in section B of the appendix.

 $^{^{4}}$ In line with previous research, we report on this hypothesis using a composite score of sharing discernment in the main text (Pennycook et al. 2020; Roozenbeek et al. 2021). Sharing discernment is equivalent to the interaction between veracity (i.e., whether the headline is false or real) and the interventions specified in the pre-registration as H3. We present the regression output of both sharing discernment and the interaction term in Section F of the appendix.

three interventions decays gradually with number of rating tasks completed (H6). We report briefly on all preregistered hypotheses in the main text and provide more details in Section A of the appendix.

Data were collected in collaboration with the market research institute YouGov through a two-wave panel among a national sample of the Danish population. YouGov sampled from their internet panels and employed quota sampling to match population characteristics on age, gender, region, and education. The power analysis suggested that 940 participants were required to have 90% power to replicate effect sizes from previous research on accuracy nudges (Pennycook et al. 2021) (see preregistration for formal power calculation). We recruited 2,541 participants between July 2 and 13, 2021 for the first wave. 2,232 participants (88% of the original sample) completed the second wave between August 2 and 23. In the first wave, we collected psychological correlates of participants, while in the second wave, participants were exposed to a survey experiment. We did not record any post-treatment attrition in wave 2.

To assess the effectiveness of the interventions in study 1, participants completed a news-sharing task consisting of 15 real and 15 false headlines. Participants were informed that they would be exposed to a range of articles from the past year concerning COVID-19. We opted for headlines instead of full articles because people often share articles on social media without reading the full article (Gabielkov et al. 2016). The headlines were presented one at a time in random order, and respondents were asked whether they were willing to share them.

The main outcome is headline sharing measured through a standard item: "If you were to see the above article on social media, how likely would you be to share it?" (1: Extremely unlikely, 2: Moderately unlikely, 3: Slightly unlikely, 4: Slightly likely, 5: Moderately likely, 6: Extremely likely; re-scaled to 0-1) (Pennycook et al. 2020). Studies validating the measure suggest that people report higher sharing intentions for headlines in surveys that do indeed receive more shares on Twitter (Mosleh et al. 2021). We use the outcome to measure both real (M = 0.26, SD = 0.32) and false (M = 0.14, SD = 0.26) headline sharing respectively. Furthermore, we use it to assess sharing discernment (M = 0.13, SD = 0.20) which is defined as the difference in sharing intentions between real and false headlines where a higher discernment score indicates that people share more real relative to false news.⁵

Results

We conduct the analyses using OLS regressions with standard errors clustered on subject and headline. Fig. 1 shows the effect of each intervention in three separate models with false and real headline sharing as well as sharing discernment, respectively, as the dependent variable (scaled $0-1)^6$.

⁵Each individual indicates sharing intentions for 15 real and 15 false headlines. The discernment score is calculated as the respondent-level difference between real and false headline sharing, that is, sharing_{real} – sharing_{false}. If a participant shared 12 out of 15 real headlines (0.8) and 6 out of 15 false headlines (0.4), their discernment score would be 0.8 - 0.4 = 0.4.

⁶Section A of the appendix provides an overview of all pre-registered hypotheses and results.



Figure 1. Willingness to share real and false headlines.

Note: Points are OLS estimates with 95% confidence interval bars based on clustered standard errors at the respondent level and headline level. The panels display estimates based on regressions of the interventions on false (n = 32,480) and real (n = 32,480) headline sharing as well as sharing discernment (n = 2,232) all re-scaled to 0-1.

Did the interventions decrease false headline sharing (H1)? The 3-minute video intervention significantly decreased willingness to share false headlines (b = -0.055, 95% CI = [-0.078, -0.031], p < 0.001, d = -0.222), while the 15-second video intervention (b = -0.002 95% CI = [-0.027, 0.023], p < 0.88, d = -0.008) and the accuracy nudge (b = -0.017, 95% CI = [-0.041, 0.008], p < 0.18, d = -0.064) did not, as shown in Fig. 1. In other words, only the 3-minute video intervention decreased false headline sharing.

Did the interventions increase real headline sharing (H2)? The 3-minute video intervention (b = 0.028, 95% CI = [-0.002, 0.052], p < 0.07, d = 0.089), the 15-second video intervention (b = 0.025, 95% CI = [-0.002, 0.052], p < 0.08, d = 0.081) and the accuracy nudge (b = 0.026, 95% CI = [-0.001, 0.054], p < 0.06, d = 0.084) did not significantly affect sharing of real headlines compared to the control condition.

Did the interventions increase sharing discernment (H3)? All three interventions significantly increased sharing discernment. In other words, both the accuracy nudge (b = 0.043, 95% CI = [0.021, 0.065], p < 0.001, d = 0.230), the 15-second video intervention (b = 0.027, 95% CI = [0.005, 0.049], p < 0.02, d = 0.145), and the 3-minute video intervention (b = 0.082, 95% CI = [0.059, 0.106], p < 0.001, d = 0.416) increased the relative sharing of real compared to false headlines.

We preregistered a range of robustness analyses of the treatment effects across covariates. We did not find that the treatment effect on sharing was significantly moderated by cognitive reflection, need for cognition, trust in government or health authorities (H4), and attention to social comparison (H5). Neither did the treatment effect significantly decay over time (H6) (See Section F in the appendix for details).

Study 2

In Study 1, we established that the 3-minute video intervention decreased sharing of false headlines. In Study 2, we used predictions derived from protection motivation theory to probe why the intervention worked. Protection motivation theory proposes that people protect themselves against risks – in our case believing in and sharing false headlines on social media – based on appraisals of (1) the threat from the risk and (2) their ability to cope with the risk (Rogers 1975; Maddux and Rogers

1983; Floyd et al. 2000; Sommestad et al. 2015; Pearce et al. 2019). Threat appraisals reflect the severity of the situation, the likelihood of a threat materializing, and individual vulnerability to the threat. Coping appraisals consist of two factors. The first factor is perceived "self-efficacy," understood as one's ability to carry out the recommended action and follow the advice successfully. The second factor is the perceived "response efficacy," understood as an individual's expectation that carrying out the recommended action and following the given advice will keep one safe from the threat. Research on protection motivation theory suggests that feelings of self-efficacy are, in general, the most important factor behind motivations to engage in protective behavior (Norman et al. 2015). Consistent with this, prior work on protection motivations in the context of the COVID-19 pandemic in Denmark has shown that compliance with public health authority advice is most strongly affected by self-efficacy (Jørgensen et al. 2021b). For study 2, we preregistered that compared to the control condition, the 3-minute intervention would increase threat appraisals (H7), feelings of self-efficacy (H8), and feelings of response efficacy (H9).7 As exploratory analyses, we also report the effects of the 15-second intervention and the accuracy nudge on these outcomes.⁸

To determine sample size, we conducted a two-sided t-test power calculation as noted in the preregistration. Given 500 participants in each group, significance level at 0.05 an effect size of d = 0.19 can be estimated with power = 0.9, and d = 0.16 with power = 0.8. A national sample of 2,012 participants quota sampled to match population characteristics on age, gender, region, and education was collected in Denmark between December 17, 2021, and December 23, 2021, by YouGov.

Instead of the news-sharing task used in Study 1, the outcome measure in Study 2 was a battery of six protection motivation items that were combined to three measures (threat appraisal, self-efficacy, and response efficacy). Besides the change of outcome, the experimental protocol was the same.

Threat appraisal, self-efficacy, and response efficacy were measured in a battery of six items where each factor was measured as the mean of its two corresponding items. The following items were included. Threat appraisal (M = 0.55, SD = 0.21): (1) "I am exposed in terms of false information regarding COVID-19," (2) "False information regarding COVID-19 is a threat to the Danish society." Self-efficacy (M = 0.80, SD = 0.22): (3) "It is easy for me to avoid spreading false information about COVID-19," (4) "I am confident that I can avoid spreading false information about COVID-19 if I want to." Response efficacy (M = 0.73, SD = 0.24): (5) "If I avoid falling for false information, I will be in greater security during the Corona epidemic," (6) "If I avoid spreading false information, I take part in protecting others against COVID-19." For each item, participants were asked to what extent they agreed or disagreed with the statements on a scale from 1, "Strongly disagree," to 7, "Strongly agree," and re-scaled to vary between 0 and 1.

⁷In the preregistration for study 2, these hypotheses are named H1A, H1B, and H1C, respectively.

⁸We did not predict specific hypotheses, but preregistered them as secondary analyses to test whether the accuracy nudge and the 15-second intervention affected threat appraisal, self-efficacy, and response efficacy.



Figure 2. Effect of interventions on threat appraisal, self-efficacy, and response efficacy. *Note:* Points are OLS estimates with 95% confidence interval bars based on clustered standard errors at the respondent level from three regressions. Each panel represents a regression of the treatment conditions on the respective pmt measure as the dependent variable. All regressions are based on samples of 2,012 respondents.

Results

Fig. 2 shows the effect of the interventions on threat appraisal, self-efficacy, and response efficacy.

Did the interventions affect threat appraisal (H7)? No. Neither the accuracy nudge (b = 0.01, 95% CI = [-0.02;0.03], p < 0.531, d = 0.039), the 15-second intervention (b = -0.02, 95% CI = [-0.04,0.01], p < 0.210, d = 0.078) or the 3-minute intervention (b = 0.01, 95% CI = [-0.02;0.03], p < 0.626, d = 0.031) significantly affected threat appraisal.⁹

Did the interventions affect self-efficacy (H8)? The 3-minute video intervention significantly increased self-efficacy (b = 0.06, 95% CI = [0.03;0.09], p < 0.001, d = 0.274), while there was no significant effect of the 15-second video intervention (b = 0.01, 95% CI = [-0.02,0.04], p < 0.529, d = 0.039) or the accuracy nudge (b = 0.01, 95% CI = [-0.02;0.04], p < 0.556, d = 0.037). This suggests that the 3-minute intervention boosts citizens' feelings of competence when they face COVID-19 misinformation.

Did the interventions affect response efficacy (H9)? No, neither the accuracy nudge (b = -0.02, 95% CI = [-0.05;0.01], p < 0.194, d = 0.081), the 15-second video intervention (b = 0.02, 95% CI = [-0.01,0.04], p < 0.288, d = 0.066), or the 3-minute video intervention (b = 0.03, 95% CI = [-0.01;0.06], p < 0.104, d = 0.103) significantly affected response efficacy.

In sum, across the 3 interventions, we only found evidence of a statistically significant effect of the 3-minute intervention on self-efficacy. Across study 1 and study 2, this suggests that the 3-minute intervention both increase people's personal feelings of being competent in terms of avoiding sharing misinformation (in study 2) and reduce people's sharing of false headlines (study 1).

⁹In the early stages of the COVID-19 pandemic, there were extensive campaigns from the Danish Health Authorities about the threat of COVID-19, including misinformation. One potential explanation that we do not observe an effect on threat appraisal is pre-treatment effects from this communication from the public authorities.

Downloaded from https://www.cambridge.org/core. Berklee College Of Music, on 06 Feb 2025 at 20:30:07, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/XPS.2024.2

Discussion

Are misinformation interventions effective against misinformation? The analyses showed that while the accuracy nudge and a 15-second capability-oriented intervention significantly increased sharing discernment - that is, the relative sharing of real vs. false headlines - they did not have a significant effect on neither false or real headline sharing compared to the control condition. The 3-minute capability-oriented intervention significantly increased sharing discernment and self-efficacy and reduced false headline sharing. In sum, we found mixed support for effectiveness of short capability-oriented messages and accuracy nudges against misinformation. These results add to recent academic research on the effectiveness of short messages or nudges on misinformation sharing (Rathje et al. 2022; Roozenbeek et al. 2021; Gavin et al. 2022; Pretus et al. 2023; Pennycook and Rand 2022). Notably, the 3-minute intervention was effective both in terms of boosting people's feelings of competence as well as reducing false headline sharing. The results expand our knowledge of the effectiveness of misinformation interventions by deploying different interventions in the same experimental framework. To assess the effectiveness of interventions, future research could adopt a similar approach in which they assess multiple interventions within the same experimental framework. Furthermore, while several studies assess whether interventions affect the belief in and sharing of misinformation (Pennycook et al. 2020; Roozenbeek et al. 2021; Guess et al. 2020; Badrinathan 2021; Bode and Vraga 2018; Jensen et al. 2022), few studies explicitly address the potential underlying mechanisms for why people alter their behavior (Lin et al. 2022; Altay et al. 2020). The finding that the most effective intervention also influenced participants' feelings of self-efficacy is consistent with the general finding in the risk communication literature that such feelings are key for motivating protective behavior (Norman et al. 2015).

In assessing these conclusions, it is worth noting important limitations of the study. While the study contributes by assessing interventions against misinformation beyond American or British samples (see also Badrinathan 2021; Guess et al. 2020; Gavin et al. 2022), it is important to note that the results are limited to a context where people are particularly responsive to government interventions. Danes held more trust in authorities, were more compliant, were less polarized, and were more concerned in the first stages of the COVID-19 pandemic than most other countries (Lindholt et al. 2021; Lieberoth et al. 2021). As a consequence, the Danes may be more responsive to messages from public health authorities. In other contexts with a higher degree of political polarization and skepticism regarding, for example, the threat of COVID-19, the effectiveness of interventions from public authorities may be more limited. To examine this proposition, future studies could use a comparative approach to assess the effectiveness of communication from public authorities across countries with varying degrees of trust, political polarization, and compliance with advice from authorities.

Practitioners should note that the interventions are mitigation strategies, rather than addressing the root cause of misinformation sharing. First, the effects of these types of interventions are small (Pennycook and Rand 2022). For instance, we included a shortened sharing task in study 2 and observed a smaller effect size than in study 1 for all interventions as the control condition was negatively affected by

the protection motivation theory items asking questions like "False information regarding COVID-19 is a threat to the Danish society."¹⁰ This in line with the literature suggesting that small nudges can boost sharing discernment (Pennycook et al. 2021). Nevertheless, the ephemeral effect sizes serve as an important reminder for researchers and practitioners alike that the effects of misinformation interventions often are small (Pennycook and Rand 2022), diminish over time (Carnahan et al. 2021; Carey et al. 2022), and thus require repeated intervention in practice (Ecker et al., 2022). Second, while a relatively small share of people share misinformation (Guess and Lyons 2020; Guess et al. 2019; Cinelli et al. 2021; Grinberg et al. 2019), recent studies suggest that social and political goals serve as important motivations (Uscinski et al. 2021; Osmundsen et al. 2021; Petersen et al. 2023; Rathje et al. 2023; Pickup et al. 2022) and some interventions may be less effective for people with certain political allegiances (Rathje et al. 2022). In Section F of the appendix, we conduct exploratory analysis of whether the interventions have heterogeneous treatment effects across trust in government, trust in health authorities, cognitive reflection, need for closure, attention to social comparison information, age, gender, income, education, and partisanship. In line with previous research, we find that partisanship predicts false headline sharing (Osmundsen et al. 2021), yet the effect of the 3-minute intervention on false headline sharing is consistent across partisanship. Furthermore, we do not find heterogeneous treatment effects for a range of other potentially relevant individual differences (specifically, cognitive reflection, need for closure, or attention to social comparison information, trust in government, or health authorities). In sum, while social, political, and accuracy motivations may shape people's overall propensity to share false headlines, the results suggest that the effect of the 3-minute video intervention is consistent. In other words, while interventions that provide competences do not address the root cause of misinformation sharing such as political motivations, they can be a reliable mitigation strategy with consistent effects across sub-populations.

The experimental design does not allow us to disentangle the effects of the length of the videos from the comprehensiveness of the advice. To be clear, our interpretation is that the 3-minute intervention is more effective in reducing false headline sharing and boosting self-efficacy, because it provides more elaborate guidance both in terms of providing information, tools, and devising actions to avoid sharing misinformation, compared to the 15-second intervention that only provides brief advice. Yet, further studies could benefit from testing interventions of similar length and assessing their efficacy.

The measures used in this study are based on sharing intentions, not actual online behavior, which clearly warrants concerns about whether the experimental results generalize beyond the experimental context. While sharing intentions is a standard way of measuring the experimental effects of misinformation interventions and it has

¹⁰We specified in the preregistration that the protection motivation theory battery would treat respondents in the control condition who would otherwise be untreated. Therefore the sharing task could not be "considered a direct replication [of study 1]." Consistent with this expectation, the reduced effect size was due to participants in the control condition being less willing to share false headlines (the mean sharing of false headlines decreases from 0.156 in study 1 to 0.131 in study 2 amounting to a difference of 0.0248 (p < 0.001). Thus, the reason the effect size is lower in the shortened sharing task is that the mean of the control condition decreases in study 2. See Table 12 and Figure 8 in the appendix for details).

been shown that sharing intentions correlates with actual sharing on social media (Mosleh et al. 2020), researchers are not able to address this caveat until social media platforms are willing to share data and conduct field experiments on their platforms. In this regard, we encourage other scholars to replicate these findings, including through field experiments, and urge social media platforms, relevant public authorities, and the like to redouble their efforts in collaborating with researchers.

In conclusion, this study has tangible policy implications for public health authorities. These findings suggest that elaborate public health communication on social media can be an effective tool for health authorities during a crisis to counter the circulation of misinformation. We demonstrated that communicating concrete advice on how to avoid sharing misinformation reduces false headline sharing and increases feelings of competence in the public.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10. 1017/XPS.2024.2

Data availability. The data, code, and any additional materials required to replicate all analyses in this article are available at the *Journal of Experimental Political Science Dataverse* within the Harvard Dataverse Network, at: doi: https://doi.org/10.7910/DVN/8UQV11 (Rasmussen et al. 2023)

Acknowledgements. The authors want to thank Jay van Bavel and members of his lab, Jon Roozenbeek, Clara Pretus, and members of the ROPH Project for helpful feedback at various stages of the project. We are grateful for feedback received during APSA 2022 and the Aarhus'22 Conference on Online Hostility and Bystanders. The authors thank Christian Jonasen Noer and Lea Pradella for excellent research assistance. Please direct correspondence pertaining to this article to Jesper Rasmussen: jr@ps.au.dk.

Funding statement. This study was funded by the Carlsberg Foundation with Grant CF20-0044 to M.B.P.

Competing interests. During the preparation of this study, Michael Bang Petersen received salary from the Danish Health Authority as a member of the National Vaccination Council, a council of independent experts advising the Danish Health Authority on matters of vaccination. The Danish Health Authority had no role in the execution of the present research. Jesper Rasmussen and Lasse Lindekilde declare no competing interests.

Ethics statement. The authors affirm that this article adheres to APSA's Principles and Guidance on Human Subject Research. The studies were exempt by Danish law from formal review by an Institutional Review Board. As per Section 14(2) of the act underlying the Danish National Research Ethics Committee, "notification of questionnaire surveys . . . to the system of research ethics committee system is only required if the project involves human biological material." Informed consent was obtained from all participants. Upon completion of the surveys, respondents received extensive debriefing and were informed that some of the headlines were false, but have nevertheless circulated on Facebook in 2020 and 2021. The ones we deemed false were fact-checked by Danish or International fact-checkers, and we provided a complete list of all the headlines used in the study and whether they were deemed for cash. See Section E of the appendix for an elaborate discussion of ethics.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Pre-registrations

Study 1: https://osf.io/akybg?view_only = 7ca987412b6d464c982f93a31d95df19 Study 2: https://osf.io/uqxv9?view_only = 7ca987412b6d464c982f93a31d95df19

References

- Altay, Sacha, Anne-Sophie Hacquin, and Hugo Mercier. 2020. "Why Do So Few People Share Fake News? It Hurts Their Reputation." *New Media & Society*: 1461444820969893. ISSN 1461-4448. doi: 10.1177/ 1461444820969893.
- Badrinathan, Sumitra. 2021. "Educative Interventions to Combat Misinformation: Evidence from a Field Experiment in India." American Political Science Review: 1–17. ISSN 0003-0554. doi: 10.1017/ s0003055421000459.
- Bode, Leticia, and Emily K. Vraga. 2018. "See Something, Say Something: Correction of Global Health Misinformation on Social Media." *Health Communication* 33 (9): 1131–40. ISSN 1041-0236. doi: 10. 1080/10410236.2017.1331312.
- Carey, John M., Andrew M. Guess, Peter J. Loewen, Eric Merkley, Brendan Nyhan, Joseph B. Phillips, and Jason Reifler. 2022. "The Ephemeral Effects of Fact-Checks on COVID-19 Misperceptions in the United States, Great Britain and Canada." *Nature Human Behaviour* 6 (2): 236–43. ISSN 2397-3374. doi: 10.1038/s41562-021-01278-3.
- Carnahan, Dustin, Daniel E. Bergan, and Sangwon Lee. 2021. "Do Corrective Effects Last? Results from a Longitudinal Experiment on Beliefs Toward Immigration in the U.S." *Political Behavior 43* (3): 1227–46. ISSN 0190-9320. doi: 10.1007/s11109-020-09591-9.
- Cinelli, Matteo, Andraz Pelicon, Igor Mozetic, Walter Quattrociocchi, Petra Kralj Novak, and Fabiana Zollo. 2021. "Dynamics of Online Hate and Misinformation." *Scientific Reports 11* (1). ISSN 2045-2322. doi: 10.1038/s41598-021-01487-w.
- Ecker, Ullrich K. H., Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K. Vraga, and Michelle A. Amazeen. 2022. "The Psychological Drivers of Misinformation Belief and Its Resistance to Correction." *Nature Reviews Psychology 1* (1): 13–29. ISSN 2731-0574. doi: 10.1038/s44159-021-00006-y.
- Flores, Alexandra, Jennifer C. Cole, Stephan Dickert, Kimin Eom, Gabriela M. Jiga-Boy, Tehila Kogut, Riley Loria, Marcus Mayorga, Eric J. Pedersen, Beatriz Pereira, Enrico Rubaltelli, David K. Sherman, Paul Slovic, Daniel Västfjäll, and Leaf Van Boven. 2022. "Politicians Polarize and Experts Depolarize Public Support for COVID-19 Management Policies Across Countries." Proceedings of the National Academy of Sciences 119 (3). doi: 10.1073/pnas.2117543119.
- Floyd, Donna L., Steven Prentice-Dunn, and Ronald W. Rogers. 2000. "A Meta-Analysis of Research on Protection Motivation Theory." *Journal of Applied Social Psychology* 30 (2): 407–29. ISSN 0021-9029. doi: 10.1111/j.1559-1816.2000.tb02323.x.
- Gabielkov, Maksym, Arthi Ramachandran, Augustin Chaintreau, and Arnaud Legout. 2016. Social Clicks: What and Who Gets Read on Twitter?. https://hal.inria.fr/hal-01281190.
- Gavin, Lyndsay, Jenna McChesney, Anson Tong, Joseph Sherlock, Lori Foster, and Sergiu Tomsa. 2022. "Fighting the Spread of COVID-19 Misinformation in Kyrgyzstan, India, and the United States: How Replicable are Accuracy Nudge Interventions?" *Technology, Mind, and Behavior 3* (3). ISSN 2689-0208. doi: 10.1037/tmb0000086.
- Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. "Fake News on Twitter during the 2016 U.S. Presidential Election." *Science* 363 (6425): 374–8. ISSN 1095-9203. doi: 10.1126/science.aau2706.
- Guess, Andrew, Jonathan Nagler, and Joshua Tucker. 2019. "Less than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook." *Science Advances*, 5 (1). doi: 10.1126/sciadv. aau4586.
- Guess, Andrew M., Michael Lerner, Benjamin Lyons, Jacob M. Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar. 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*: 201920498. doi: 10.1073/pnas.1920498117.
- Guess, Andrew M. and Benjamin A. Lyons. 2020. Misinformation, Disinformation, and Online Propaganda. In Social Media and Democracy, ed. Nathaniel Persily and Joshua A. Tucker. Cambridge University Press, 1 edition, 10–33. ISBN 978-1-108-89096-0 978-1-108-83555-8 978-1-108-81289-4. doi: 10.1017/9781108890960.003.

- Hertwig, Ralph and Till Grüne-Yanoff. 2017. "Nudging and Boosting: Steering or Empowering Good Decisions." *Perspectives on Psychological Science* 12 (6): 973–86. ISSN 1745-6916. doi: 10.1177/ 1745691617702496.
- Jensen, Ulrich T., Stephanie Ayers, and Alexis M. Koskan. 2022. "Video-based Messages to Reduce COVID-19 Vaccine Hesitancy and Nudge Vaccination Intentions." *PLOS ONE 17* (4). ISSN 1932-6203. doi: 10.1371/journal.pone.0265736.
- Jørgensen, Frederik, Alexander Bor, Marie Fly Lindholt, and Michael Bang Petersen. 2021a. "Public Support for Government Responses against COVID-19: Assessing Levels and Predictors in Eight Western Democracies During 2020." West European Politics, 44 (5-6): 1129–58. ISSN 0140-2382. doi: 10. 1080/01402382.2021.1925821.
- Jørgensen, Frederik, Alexander Bor, and Michael Bang Petersen. 2021b. "Compliance without Fear: Individual-Level Protective Behaviour During the First Wave of the COVID-19 Pandemic." *British Journal of Health Psychology*, 26 (2): 679–96. ISSN 1359-107X. doi: 10.1111/bjhp.12519.
- Lee, Nicole M. 2018. "Fake News, Phishing, and Fraud: A Call for Research on Digital Media Literacy Education Beyond the Classroom." *Communication Education* 67 (4): 460–66. ISSN 0363-4523. doi: 10. 1080/03634523.2018.1503313.
- Lieberoth, Andreas, Shiang-Yi Lin, Sabrina Stöckli, Hyemin Han, Marta Kowal, Rebekah Gelpi, Stavroula Chrona, Thao Phuong Tran, Alma Jeftić, Jesper Rasmussen, Huseyin Cakal, and Taciano L. Milfont. 2021. "Stress and Worry in the 2020 Coronavirus Pandemic: Relationships to Trust and Compliance with Preventive Measures Across 48 Countries in the COVIDiSTRESS Global Survey." Royal Society Open Science 8 (2). doi: 10.1098/rsos.200589.
- Lin, Hause, Gordon Pennycook, and David Rand. 2022. "Thinking more or Thinking Differently? Using Drift-Diffusion Modeling to Illuminate Why Accuracy Prompts Decrease Misinformation Sharing." *PsyArxiv.* doi: 10.31234/osf.io/kf8md.
- Lindholt, Marie Fly, Frederik Jørgensen, Alexander Bor, and Michael Bang Petersen. 2021. "Public Acceptance of COVID-19 Vaccines: Cross-National Evidence on Levels and Individual-Level Predictors Using Observational Data." BMJ Open 11 (6). ISSN 2044-6055. doi: 10.1136/bmjopen-2020-048172.
- Maddux, James E., and Ronald W. Rogers. 1983. "Protection Motivation and Self-efficacy: A Revised Theory of Fear Appeals and Attitude Change." *Journal of Experimental Social Psychology*, *19* (5): 469–79. ISSN 0022-1031. doi: 10.1016/0022-1031(83)90023-9.
- Mo Jones-Jang, S., Tara Mortensen, and Jingjing Liu. 2021. "Does Media Literacy Help Identification of Fake News? Information Literacy Helps, but Other Literacies Don't. *American Behavioral Scientist* 65 (2): 371–88. ISSN 0002-7642. doi: 10.1177/0002764219869406.
- Mosleh, Mohsen, Gordon Pennycook, and David G. Rand. 2020. "Self-Reported Willingness to Share Political News Articles in Online Surveys Correlates with Actual Sharing on Twitter." *PLOS ONE*, *15* (2). ISSN 1932-6203. doi: 10.1371/journal.pone.0228882.
- Mosleh, Mohsen, Gordon Pennycook, and David G. Rand. 2021. "Field Experiments on Social Media." *Current Directions in Psychological Science*. ISSN 0963-7214. doi: 10.1177/09637214211054761.
- Norman, Paul, Henk Boer, Erwin R. Seydel, and Barbara Mullan. 2015. "Protection Motivation Theory." In *Predicting and Changing Health Behaviour: Research and Practice with Social Cognition Models*, Vol. 3. Maidenhead: Open University Press, 70–106.
- Osmundsen, Mathias, Alexander Bor, Peter Bjerregaard Vahlstrup, Anja Bechmann, and Michael Bang Petersen. 2021. "Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter." *American Political Science Review*: 1–17. ISSN 0003-0554. doi: 10.1017/ s0003055421000290.
- Pearce, J. M., L. Lindekilde, D. Parker, and M. B. Rogers. 2019. "Communicating with the Public About Marauding Terrorist Firearms Attacks: Results from a Survey Experiment on Factors Influencing Intention to 'Run, Hide, Tell' in the United Kingdom and Denmark." *Risk Analysis*, 39 (8): 1675–94. ISSN 1539-6924 (Electronic) 0272-4332 (Linking). doi: 10.1111/risa.13301.
- Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand. 2021. "Shifting Attention to Accuracy Can Reduce Misinformation Online." *Nature*, 592 (7855): 590–95. ISSN 0028-0836. doi: 10.1038/s41586-021-03344-2.

- Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. 2020. "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention." *Psychological Science* 31 (7): 770–80. ISSN 0956-7976. doi: 10.1177/ 0956797620939054.
- Pennycook, Gordon and David Rand. 2022. "Accuracy Prompts are a Replicable and Generalizable Approach for Reducing the Spread of Misinformation." *Nature Communications 13* (1). ISSN 2041-1723. doi: 10.1038/s41467-022-30073-5.
- Petersen, Michael Bang, Mathias Osmundsen, and Kevin Arceneaux. 2023. "The 'Need for Chaos' and Motivations to Share Hostile Political Rumors." American Political Science Review: 1–20. ISSN 0003-0554. doi: 10.1017/s0003055422001447.
- Pickup, Mark, Dominik Stecuła, and Sander Van Der Linden. 2022. "Who Shares Conspiracy Theories and Other Misinformation about Covid-19 Online: Survey Evidence from Five Countries." *Journal of Quantitative Description: Digital Media 2.* ISSN 2673-8813. doi: 10.51685/jqd.2022.024.
- Pretus, Clara, Ali Javeed, Diána R. Hughes, Kobi Hackenburg, Manos Tsakiris, Oscar Vilarroya, and Jay J. Van Bavel. 2023. "The Misleading Count: An Identity-Based Intervention to Mitigate the Spread of Partisan Misinformation." *PsyArxiv.* doi: 10.31234/osf.io/7j26y.
- Rasmussen, Jesper, Lasse Lindekilde, and Michael Bang Petersen. 2023. "Replication Data for: Public Health Communication Reduces COVID-19 Misinformation Sharing and Boosts Self-Efficacy." *Harvard Dataverse*. doi: 10.7910/DVN/8UQV11.
- Rathje, Steve, Jon Roozenbeek, Cecilie Steenbuch Traberg, Jay J. Van Bavel, and Sander Van Der Linden. 2022. "Letter to the Editors of Psychological Science: Meta-Analysis Reveals that Accuracy Nudges Have Little to No Effect for U.S. Conservatives: Regarding Pennycook et al. (2020)." *PsyArxiv.* doi: 10.31234/osf.io/945na.
- Rathje, Steve, Jon Roozenbeek, Jay J. Van Bavel, and Sander Van Der Linden. 2023. "Accuracy and Social Motivations Shape Judgements of (Mis)information." *Nature Human Behaviour*. ISSN 2397-3374. doi: 10.1038/s41562-023-01540-w.
- **Rippetoe**, **Patricia A.**, and **Ronald W. Rogers**. 1987. "Effects of Components of Protection-Motivation Theory on Adaptive and Maladaptive Coping with a Health Threat." *Journal of Personality and Social Psychology*, 52 (3): 596. ISSN 1939-1315.
- Rogers, Ronald W. 1975. "A Protection Motivation Theory of Fear Appeals and Attitude Changel." The Journal of Psychology 91 (1): 93–114. ISSN 0022-3980. doi: 10.1080/00223980.1975.9915803.
- Roozenbeek, Jon, Alexandra L. J. Freeman, and Sander Van Der Linden. 2021. "How Accurate Are Accuracy-Nudge Interventions? A Preregistered Direct Replication of Pennycook et al. (2020)." *Psychological Science*, 09567976211024535. ISSN 0956-7976. doi: 10.1177/09567976211024535.
- Roozenbeek, Jon and Sander Van Der Linden. 2022. "How to Combat Health Misinformation: A Psychological Approach." American Journal of Health Promotion 36 (3): 569–75. ISSN 0890-1171. doi: 10.1177/08901171211070958.
- Sheeran, Paschal, Richard Aubrey, and Stephen Kellett. 2007. "Increasing Attendance for Psychotherapy: Implementation Intentions and the Self-regulation of Attendance-Related Negative Affect." *Journal of Consulting and Clinical Psychology* 75 (6): 853–63. ISSN 1939-2117, 0022-006X. doi: 10.1037/0022-006X. 75.6.853.
- Sheeran, Paschal, and Sheina Orbell. 2000. "Using Implementation Intentions to Increase Attendance for Cervical Cancer Screening." *Health Psychology* 19 (3): 283. ISSN 1930-7810. doi: 10.1037/0278-6133.19.3.283.
- Sommestad, Teodor, Henrik Karlzén, and Jonas Hallberg. 2015. "A Meta-Analysis of Studies on Protection Motivation Theory and Information Security Behaviour." *International Journal of Information Security and Privacy 9* (1): 26–46. ISSN 1930-1650, 1930-1669. doi: 10.4018/IJISP. 2015010102.
- Uscinski, Joseph E., Adam M. Enders, Michelle I. Seelig, Casey A. Klofstad, John R. Funchion, Caleb Everett, Stefan Wuchty, Kamal Premaratne, and Manohar N. Murthi. 2021. "American Politics in Two Dimensions: Partisan and Ideological Identities versus Anti-Establishment Orientations." *American Journal of Political Science* 65 (4): 877–95. ISSN 0092-5853, 1540-5907. doi: 10.1111/ajps. 12616.

- Van Bavel, Jay J., Aleksandra Cichocka, Valerio Capraro, Hallgeir Sjåstad, John B. Nezlek, Tomislav Pavlović, Mark Alfano, Michele J. Gelfand, Flavio Azevedo, Michèle D. Birtel, Aleksandra Cislak, Patricia L. Lockwood, Robert Malcolm Ross, Koen Abts, Elena Agadullina, John Jamir Benzon Aruta, Sahba Nomvula Besharati, Alexander Bor, Becky L. Choma, ... Paulo S. Boggio. 2022.
 "National Identity Predicts Public Health Support During a Global Pandemic." Nature Communications 13 (1). ISSN 2041-1723. doi: 10.1038/s41467-021-27668-9.
- Van Der Linden, Sander. 2022. "Misinformation: Susceptibility, Spread, and Interventions to Immunize the Public." *Nature Medicine*. ISSN 1078-8956. doi: 10.1038/s41591-022-01713-6.

³⁴² Jesper Rasmussen et al.

Cite this article: Rasmussen J, Lindekilde L, and Petersen MB (2024). Public Health Communication Reduces COVID-19 Misinformation Sharing and Boosts Self-Efficacy. *Journal of Experimental Political Science* 11, 327–342. https://doi.org/10.1017/XPS.2024.2

Downloaded from https://www.cambridge.org/core. Berklee College Of Music, on 06 Feb 2025 at 20:30:07, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/XPS.2024.2