Supplementary material for "Public Health Communication Reduces COVID-19 Misinformation Sharing and Boosts Self-Efficacy"

January 22, 2024

Contents

1 Section A: Pre-registration

Prior to data collection, the analyses were preregistered on the Open Science Framework. Blinded versions of the pre-registrations for study 1, study 2 and all materials are available online. In this manuscript we focus on hypothesis 1, 2, and 3 for study 1.

Table 1. Study 1: Preregistered hypotheses

Hypothesis	**Result**
H1 (primary hypothesis): Compared to the control condition, all three interventions decrease the likelihood that citizens will be willing to share false headlines, but not real headlines about COVID-19 on social media.	The 3-minute interventions significantly decrease false headline sharing. The accuracy nudge and the 15-seconds intervention did not significantly affect false headline sharing (Figure 1 in the main text)
H2: Compared to the control condition, all three interventions increase the likelihood that citizens will be willing to share true headlines, but not false headlines about COVID-19 on social media.	Not supported. None of the interventions significantly increase real headline sharing (Figure 1 in the main text)
H3 (replication): A significant positive interaction between headline veracity (true or false headline) and treatment (accuracy induction) predicting likelihood to share, such that the treatment condition increases sharing discernment.	Support. In Section F of the appendix, Table 5 shows the coefficient for the interaction term between the interventions and the veracity of the headline
H4: The treatment effect of the more elaborate COVID-19 related accuracy-nudges on citizens' likelihood of fake news on social media is lower for respondents who has 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.	Not supported. Neither of the correlates significantly decrease the treatment effect. See Section F Table 8 and Table 9 of the appendix.
H5: There is 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 fake news, such that the effect is stronger for people who score high on attention to social comparison information	Not supported. We found no statistically significant interaction. See Section F Table 8 and Table 9 of the appendix.
H6: The effect of all three interventions decay gradually with number of rating tasks completed	Not supported. We do not find a statistically significant interaction. See Section F Table 11 as well as Figure 3 and Figure 4

Table 2. Study 2: Preregistered hypotheses

Hypothesis	**Result**
H1a: On average, participants in the long intervention condition display a higher threat perception, compared to participants in the control condition.	Not supported. See Figure 3 main text
H1b: On average, participants in the long intervention condition display higher self-efficacy, compared to participants in the control condition.	Support. See Figure 3 main text
H1c: On average, participants in the long intervention condition display higher response efficacy, compared to participants in the control condition.	Not supported. See Figure 3 main text

2 Section B: Stimuli

A screenshot of the stimuli from each condition is presented in figure 1 in the main text. The accuracy nudge resembles previous accuracy nudges (see Pennycook et al. (2021, 2020)). One of four non-covid-19 related headlines were presented to the participants which are available online. The 15-second and 3-minute video interventions from the Danish Health Authorities are freely available online.

2.1 Transcript 15-seconds intervention

[Text] Good advice for spotting misinformation

Ask yourself:

- Who is saying it?
- How many are saying it?
- Is the content too far out?

2.2 Transcript 3-minute intervention

[Text]: Can you trust what you read? [Speak]: We need knowledge to be able to combat COVID-19. And we are being told a lot all the time. But it is not all stories that are true and some may be misleading. The best defense against misleading information is common sense. And to strengthen it [common sense], here are some simple rules of thumb to help you decide whether you should believe what you hear – and not least whether to share it with others.

We require knowledge to combat COVID-19, and we are constantly bombarded with information. However, not all stories are true, and some may be misleading. The most effective defense against misleading information is common sense. To enhance your common sense, here are some guidelines to assist you in determining whether you should believe what you hear — and, not least, whether you should share it with others.

1. Who is saying it?

Is it someone you trust or someone having expertise within their field? For instance, is it an experienced researcher, an authority or an established media outlet? Do the the sender have a certain interest in saying it? If they have an interest in profiling themselves or their cause in a way, there is reason to think twice.

2. How many are saying it?

When multiple media outlets report the same story, it is more likely to have undergone fact-checking. If a story originates from only one media outlet, website, person, or group, it is likely that it hasn't been thoroughly fact-checked, and there's a reason to be cautious, even if the post is widely shared.

The same applies to experts publishing new research and new studies about COVID-19. It is important to pay attention to how other experienced researchers interpret the new studies and whether other studies validate the findings. The more studies that support the same conclusion, the greater the likelihood that it can be trusted.

3. Is the content too far out?

Good stories are not always true. If they sound too far-fetched or surprising, there's a reason to be skeptical and to wait and see before sharing.

Stories that evoke strong emotions like fear, anger, and great joy are more likely to captivate us and make us want to share them. But in those cases, it's important to be cautious. Information from authorities and credible experts is often more boring. However, in this situation, "boring" is actually a good thing.

Just like a virus, stories can spread rapidly. Sharing a crazy story may seem harmless, but false information can lead to confusion and insecurity, and it can directly impact people's well-being. That's why it's crucial for all of us to consider the stories we share with others.

We often share and comment on stories because they elicit happiness, anger, or worry. It's natural, but remember to maintain a civil tone when engaging in discussions with others, especially on social media.

You can help spread only accurate information by using common sense and thinking before

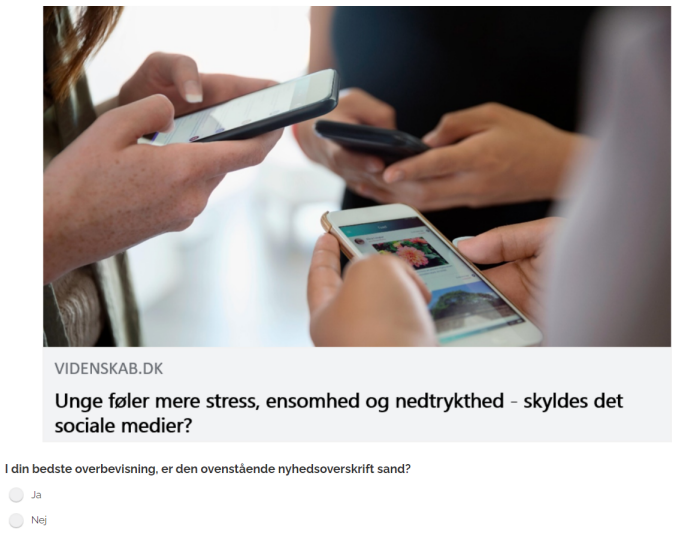
sharing stories with others. If you're uncertain, it's best not to share, and you can utilize the option to report false news on the platforms you use, such as Facebook. By doing so, you can help making a difference. And, if your common sense agrees, you can also contribute by sharing this video.

2.3 Screenshot of experimental stimuli

3 Section C: Pretest of headlines

All headlines headlines used in this study are available online and were pre-tested prior to the experiments. From July 2nd to July 5th 2021, the survey agency YouGov conducted 205 interviews in Denmark. Each respondent rated 44 headlines on the question: "In your best belief, is the above news headline true?" The results of the pretest is presented in figure 2.

Figure 1. Experimental stimuli



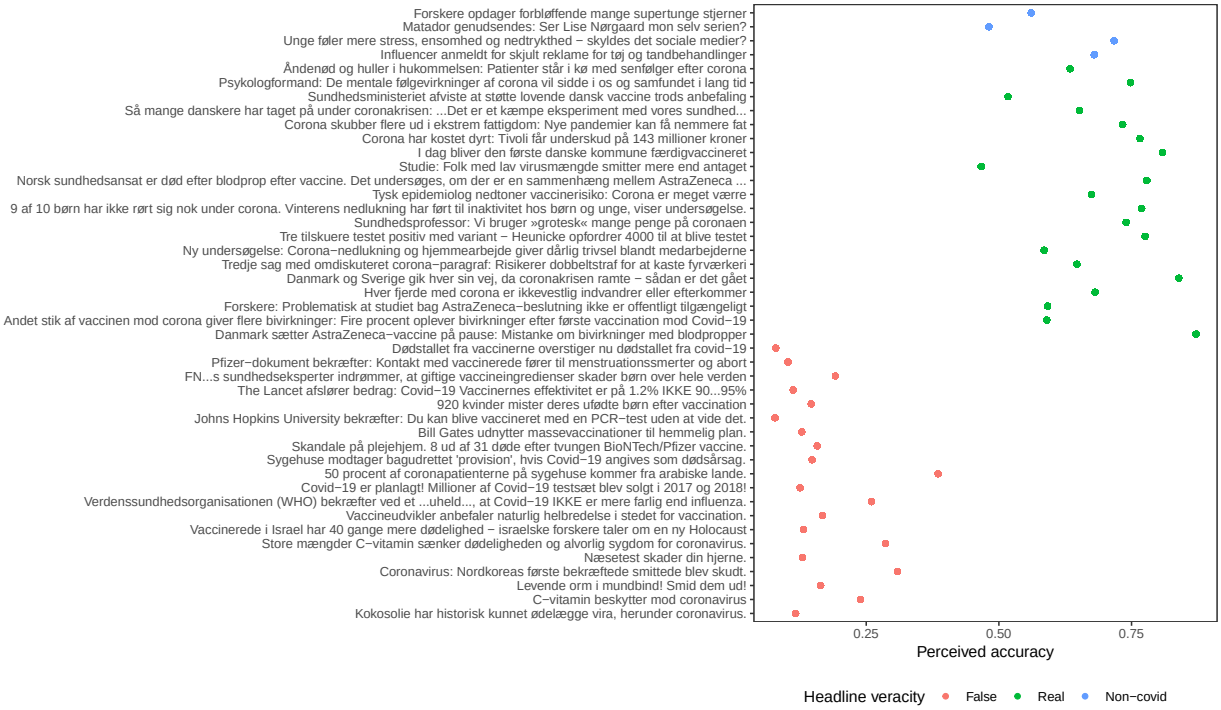
(a) Nudge

(b) 15-second video intervention

(c) 3-minute video intervention

Note: The images are screenshots of the interventions. The nudge (a) is a task where participants are asked to rate whether they believe one non-COVID-related headline is true. The (b) 15-second intervention contains three advice to spot misinformation. The (c) 3-minute intervention provides the same three advice and elaborate on how concretely to act on and implement the advice on how to spot misinformation.

Figure 2. Pretest of headlines



4 Section D: Descriptive statistics

Study 1: Sample (data provided by YouGov)

		Unweighted (n)		Weighted (n)	
		2541	100%	2541	100%
Gender	Female	1276	50,2	1286	50,6
	Male	1265	49,8	1255	49,4
Region	Capital Region	741	29,2	804	31,6
	Zealand	384	15,1	367	14,5
	Southern Denmark	576	22,7	536	21,1
	Central Denmark	578	22,7	574	22,6
	North Denmark	262	10,3	260	10,3
Age	18-34	628	24,7	697	27,4
	35-54	692	27,2	830	32,6
	55-69	702	27,6	569	22,4
	70+	519	20,4	446	17,5
Education	Upper secondary education	1598	62,9	1733	68,2
	Short/medium-cycle higher education	602	23,7	569	22,4
	Long cycle higher education	341	13,4	239	9,4

Study 2: Sample (data provided by YouGov)

		Unweighted (n)		Weighted (n)	
		2012	100%	2012	100%
Gender	Female	1080	53,7	1018	50,6
	Male	932	46,3	994	49,4
Age	18-34	454	22,6	552	27,4
	35-54	699	34,7	657	32,6
	55-69	491	24,4	450	22,4
	70+	368	18,3	353	17,5
Region	Capital Region	611	30,4	637	31,6
	Zealand	304	15,1	291	14,5
	Southern Denmark	423	21	424	21,1
	Central Denmark	460	22,9	454	22,6
	North Denmark	214	10,6	206	10,2
Education	Upper secondary education	1306	64,9	1373	68,2
	Short/medium-cycle higher education	498	24,8	450	22,4
	Long cycle higher education	208	10,3	189	9,4

Table 3. Summary statistics for dependent variables in Study 1

Treatment		Mean	SD	N
Control	False Sharing	0.16	0.22	566
	Real Sharing	0.24	0.23	566
	Sharing Discernment	0.09	0.18	566
Nudge	False Sharing	0.14	0.21	569
	Real Sharing	0.27	0.26	569
	Sharing Discernment	0.13	0.20	569
Short	False Sharing	0.15	0.22	552
	Real Sharing	0.27	0.25	552
	Sharing Discernment	0.12	0.20	552
Long	False Sharing	0.10	0.17	545
	Real Sharing	0.27	0.25	545
	Sharing Discernment	0.17	0.22	545

Sample size: 2232

Table 4. Summary statistics for dependent variables in Study 2

Treatment		Mean	SD	N
Control	Threat Appraisal	0.55	0.20	509
	Self Efficacy	0.78	0.22	509
	Response Efficacy	0.73	0.24	509
Nudge	Threat Appraisal	0.55	0.21	511
	Self Efficacy	0.79	0.23	511
	Response Efficacy	0.71	0.24	511
Short	Threat Appraisal	0.53	0.20	513
	Self Efficacy	0.79	0.23	513
	Response Efficacy	0.74	0.23	513
Long	Threat Appraisal	0.55	0.21	479
	Self Efficacy	0.84	0.21	479
	Response Efficacy	0.75	0.25	479

Sample size: 2012

5 Section E: Ethics

The research complies with Aarhus University's Code of Conduct and the ethical standards set by the Danish Code of Conduct for Research Integrity. Section 14(2) of the Act on Re-search Ethics Review of Health Research Projects, "notification of questionnaire surveys ... to the system of research ethics committee system is only required if the project involves human biological material". All participants provided informed consent prior to participating in the study and received a thorough debriefing after completing the study. The study did not involve deception. We elaborate on the ethical considerations in relation to the protocol, informed consent and debriefing below.

We initiated this study to assess whether communication from the public health authorities can mitigate the circulation of misinformation during pandemics. Throughout the pandemic, The Danish Health Authorities (Sundhedsstyrelsen) has invested heavily in communicating to the public. In collaboration with the Danish Health Authorities, we agreed to assess the effectiveness of two of their interventions, which were circulated on Facebook in December 2020 and January 2021.

The protocol of this experiment followed state-of-the-art paradigm of evaluating interventions against misinformation employed by a some of the worlds leading research teams (Pennycook et al., 2021; Roozenbeek et al., 2021). Within this paradigm, participants are exposed to true and false headlines, because it is the only way to measure whether people share less misinformation, when they are exposed to the intervention. The study did not involve deception. Participants were shown headlines they could encounter in their everyday life on social media in similar way as an extensive literature within this paradigm studies misinformation. On that basis, we consider the risks associated with this study to be low. The following steps were taken to ensure that participants were briefed thoroughly about the minimal risks and inconveniences associated with this study. All questionaires are avaiable on OSF.

Informed consent: We recruited participants through YouGov's panels. YouGov panelists are compensated for their time with points which can be redeemed for rewards. While YouGov

collects informed consent for their panelists, we collected additional informed consent from participants. First, the participants were briefed that the purpose the study which were about "news on social media about COVID-19". In addition the participants were informed that they would be "exposed for some headlines about COVID-19 which had circulated on social media" and that "It is important that you are aware that you in the study can read information which is not in line with the advice of the Health authorities". The briefing made it clear that clear that participation in the study was voluntary, anonymous, that the participant is free to withdraw their consent and participation in the study at any point even after completing the survey. Finally, the name including a direct email address to the responsible researcher providing an opportunity to address any concerns or questions.

Debriefing: After completing the study, the participants received a thorough debriefing. The participants were informed that the purpose of the study was to understand why false headlines are shared and what we can do to mitigate it. The participants were informed that some participants have seen the videos from the Danish Health Authorities to which a link was provided. It was stressed that some of the headlines were false, but has nevertheless circulated on Facebook in 2020 and 2021. The one's 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 true or false. Finally, we reiterated the direct email address to the responsible researcher providing yet another opportunity to address any concerns or questions.

6 Section F: Regression output and additional analyses

In this section, we a range of additional analyses in the following order:

- Regresssion output of the figures in the manuscript.
- Additional pre-registered analyses
- Additional supplementary analyses

6.1 Regression output from manuscript

Table 5. Regression output of figure 1 in main text: Willingness to share headlines.

	False headline sharing	Real headline sharing	All headlines (Intervention X veracity)
Accuracy nudge (vs. control)	−0.017 (0.013)	0.026 (0.014)	−0.017 (0.013)
15 sec video (vs. control)	−0.002 (0.013)	0.025 (0.014)	−0.002 (0.013)
3 min video (vs. control)	−0.055 (0.012)***	0.028 (0.015)	−0.055 (0.012)***
Headline veracity (0 = false, 1 = true)			0.088 (0.012)***
Accuracy nudge X veracity			0.043 (0.011)***
15 sec video X veracity			0.027 (0.010)*
3 min video X veracity			0.082 (0.013)***
Constant	0.156 (0.010)***	0.244 (0.013)***	0.156 (0.010)***
N	33 480	33 480	66 960
R2	0.01	0.00	0.05
R2 Adj.	0.01	0.00	0.05

Note:
'False headline sharing' shows the regression output of the left panel in figure 1 in the main text. 'Real headline sharing' shows the regression output of the middle panel. 'All headlines (Treatment X veracity)' shows a regression where all headlines are included with the interaction of the treatment interventions and the veracity of the headline on sharing intention. Coefficients are OLS estimates. Standard errors (in parentheses) are clustered at the respondent and headline level. All regressions are based on observations from 2232 respondents. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6. Regression output of figure 1 in main text: Willingness to share headlines.

	Sharing discernment
Accuracy nudge (vs. control)	0.043 (0.011)***
15 sec video (vs. control)	0.027 (0.011)*
3 min video (vs. control)	0.082 (0.012)***
Constant	0.088 (0.007)***
N	2232
R2	0.02
R2 Adj.	0.02

Note:
'Sharing discernment' shows the regression output of the right panel in figure 1 in the main text. The discernment score is calculated as the respondent level difference between mean real and false headline sharing, that is, sharing real - sharing false. Coefficients are OLS estimates. Standard errors (in parentheses) are clustered at the respondent level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7. Regression output of figure 2 in main text: Effect of interventions on protection motivation theory.

	Response efficacy	Self efficacy	Threat appraisal
Accuracy nudge (vs. control)	−0.019 (0.015)	0.008 (0.014)	0.008 (0.013)
15 sec video (vs. control)	0.016 (0.015)	0.009 (0.014)	−0.016 (0.013)
3 min video (vs. control)	0.025 (0.015)	0.060 (0.014)***	0.006 (0.013)
Constant	0.729 (0.011)***	0.778 (0.010)***	0.547 (0.009)***
N	2012	2012	2012
R2	0.00	0.01	0.00
R2 Adj.	0.00	0.01	0.00

Note:
The three columns shows the regression output of figure 2 in the main text for 1) Threat appraisal, 2) Self efficacy, and 3) Response efficacy respectively. Coefficients are OLS estimates. Standard errors (in parentheses) are clustered at the respondent level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.2 Pre-registered analyses

In this section, we present additional pre-registered analyses that assess the robustness of the findings.

- **Hypothesis 4** ask whether the treatment effects moderated by cognitive reflection, need for cognition, attention to social comparison, trust, trust in government and trust in health authorities? Table 8 reports regression results of the treatment effects interacted with the covariates: cognitive reflection, need for cognition, attention to social comparison, trust, trust in government and trust in health authorities.
- **Hypothesis 5** concerns whether the treatment effects on sharing fake and real news headlines moderated by attention to social comparison? As reported in Table 8 (Model 3) and Table 9 (Model 3), none of the treatments are significantly moderated by attention to social comparison on either sharing real or fake news headlines.
- **Hypothesis 6** asks whether the treatment effect decay gradually with the number of rating tasks completed?. Table 11 reports the treatments effect interacted with the number of rating tasks completed. The insignificant interaction effects indicate that the treatment effects does not change with the number of rating task increasing. This is visualised in Figure 3 and Figure 4 which shows the relatively stable means in all treatment groups across the number of rating tasks completed.

Table 8. Effect of interventions on sharing fake news headlines interacted with covariates. \

	(1)	(2)	(3)	(4)	(5)	(6)
Accuracy nudge (vs. control)	−0.028 (0.017)	0.040 (0.052)	−0.035 (0.035)	−0.057 (0.053)	−0.048 (0.030)	−0.013 (0.046)
15 sec video (vs. control)	−0.004 (0.018)	0.032 (0.055)	0.025 (0.035)	0.071 (0.057)	0.031 (0.032)	0.078 (0.048)
3 min video (vs. control)	−0.056*** (0.016)	−0.034 (0.046)	−0.081** (0.030)	−0.083 (0.053)	−0.068* (0.030)	−0.040 (0.045)
Cognitive reflection	−0.014 (0.025)					
Accuracy nudge X cognitive reflection	0.033 (0.036)					
15 sec video X cognitive reflection	0.005 (0.037)					
3 min video X cognitive reflection	0.003 (0.033)					
Need for cognition		0.061 (0.072)				
Accuracy nudge X need for cognition		−0.106 (0.093)				
15 sec video X need for cognition		−0.065 (0.104)				
3 min video X need for cognition		−0.039 (0.085)				
Attention to social comparison			−0.006 (0.051)			
Accuracy nudge X attention to social comparison			0.041 (0.074)			
15 sec video X attention to social comparison			−0.062 (0.074)			
3 min video X attention to social comparison			0.059 (0.064)			
Trust				−0.008 (0.063)		
Accuracy nudge X trust				0.061 (0.078)		
15 sec video X trust				−0.103 (0.085)		
3 min video X trust				0.045 (0.077)		
Trust government					−0.014 (0.038)	
Accuracy nudge X trust government					0.055 (0.048)	
15 sec video X trust government					−0.051 (0.050)	
3 min video X trust government					0.026 (0.047)	
Trust in Health Auth.						0.022 (0.046)
Accuracy nudge X trust in Health Auth.						−0.004 (0.060)
15 sec video X trust in Health Auth.						−0.101 (0.063)
3 min video X trust in Health Auth.						−0.014 (0.060)
Constant	0.161*** (0.012)	0.124** (0.039)	0.159*** (0.024)	0.160*** (0.043)	0.162*** (0.024)	0.137*** (0.035)
N	33 480	33 480	33 480	31 485	32 520	32 370
R2	0.01	0.01	0.01	0.01	0.01	0.01
R2 Adj.	0.01	0.01	0.01	0.01	0.01	0.01

Note:
Each model shows the interaction effects between treatments and one of the covariates: Cognitive reflection, need for cognition, attention to social comparison, trust, trust in government, and trust in health authorities. Coefficients are OLS estimates. All regressions are based on observations from 2232 respondents. Standard errors (in parentheses) are clustered at the respondent level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9. Effect of interventions on sharing real news headlines interacted with covariates. \

	(1)	(2)	(3)	(4)	(5)	(6)
Accuracy nudge (vs. control)	0.011 (0.020)	0.088 (0.054)	0.004 (0.039)	0.052 (0.062)	0.036 (0.036)	0.063 (0.053)
15 sec video (vs. control)	0.010 (0.019)	0.103 (0.058)	0.076* (0.037)	0.151* (0.060)	0.084* (0.034)	0.122* (0.051)
3 min video (vs. control)	0.010 (0.019)	−0.014 (0.056)	0.027 (0.038)	−0.031 (0.063)	0.014 (0.037)	0.013 (0.055)
Cognitive reflection	−0.039 (0.026)					
Accuracy nudge X cognitive reflection	0.048 (0.040)					
15 sec video X cognitive reflection	0.047 (0.039)					
3 min video X cognitive reflection	0.054 (0.041)					
Need for cognition		0.072 (0.070)				
Accuracy nudge X need for cognition		−0.116 (0.099)				
15 sec video X need for cognition		−0.150 (0.109)				
3 min video X need for cognition		0.079 (0.104)				
Attention to social comparison			0.048 (0.055)			
Accuracy nudge X attention to social comparison			0.050 (0.082)			
15 sec video X attention to social comparison			−0.117 (0.080)			
3 min video X attention to social comparison			0.000 (0.080)			
Trust				0.041 (0.065)		
Accuracy nudge X trust				−0.041 (0.091)		
15 sec video X trust				−0.187* (0.089)		
3 min video X trust				0.087 (0.094)		
Trust government					0.021 (0.040)	
Accuracy nudge X trust government					−0.016 (0.057)	
15 sec video X trust government					−0.093 (0.055)	
3 min video X trust government					0.026 (0.058)	
Trust in Health Auth.						0.046 (0.050)
Accuracy nudge X trust in Health Auth.						−0.049 (0.071)
15 sec video X trust in Health Auth.						−0.127 (0.068)
3 min video X trust in Health Auth.						0.025 (0.073)
Constant	0.257*** (0.013)	0.206*** (0.038)	0.223*** (0.025)	0.217*** (0.043)	0.230*** (0.024)	0.209*** (0.038)
N	33 480	33 480	33 480	31 485	32 520	32 370
R2	0.00	0.00	0.00	0.00	0.00	0.00
R2 Adj.	0.00	0.00	0.00	0.00	0.00	0.00

Note:
Each model shows the interaction effects between treatments and one of the covariates: Cognitive reflection, need for cognition, attention to social comparison, trust, trust in government, and trust in health authorities. Coefficients are OLS estimates. All regressions are based on observations from 2232 respondents. Standard errors (in parentheses) are clustered at the respondent level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10. Effect of interventions on sharing discernment interacted with covariates. \

	(1)	(2)	(3)	(4)	(5)	(6)
Accuracy nudge (vs. control)	0.038* (0.015)	0.049 (0.041)	0.039 (0.030)	0.108* (0.046)	0.084** (0.028)	0.075 (0.041)
15 sec video (vs. control)	0.014 (0.015)	0.071 (0.047)	0.051 (0.029)	0.080 (0.044)	0.053* (0.027)	0.044 (0.039)
3 min video (vs. control)	0.066*** (0.016)	0.020 (0.047)	0.108*** (0.032)	0.053 (0.050)	0.083** (0.030)	0.053 (0.043)
Cognitive reflection	−0.025 (0.019)					
Accuracy nudge X cognitive reflection	0.015 (0.030)					
15 sec video X cognitive reflection	0.042 (0.030)					
3 min video X cognitive reflection	0.051 (0.033)					
Need for cognition		0.011 (0.053)				
Accuracy nudge X need for cognition		−0.010 (0.074)				
15 sec video X need for cognition		−0.085 (0.087)				
3 min video X need for cognition		0.117 (0.087)				
Attention to social comparison			0.055 (0.044)			
Accuracy nudge X attention to social comparison			0.009 (0.067)			
15 sec video X attention to social comparison			−0.054 (0.063)			
3 min video X attention to social comparison			−0.059 (0.068)			
Trust				0.049 (0.045)		
Accuracy nudge X trust				−0.102 (0.068)		
15 sec video X trust				−0.084 (0.065)		
3 min video X trust				0.042 (0.074)		
Trust government					0.035 (0.031)	
Accuracy nudge X trust government					−0.071 (0.044)	
15 sec video X trust government					−0.042 (0.043)	
3 min video X trust government					0.000 (0.048)	
Trust in Health Auth.						0.024 (0.035)
Accuracy nudge X trust in Health Auth.						−0.045 (0.054)
15 sec video X trust in Health Auth.						−0.026 (0.051)
3 min video X trust in Health Auth.						0.039 (0.057)
Constant	0.096*** (0.010)	0.082** (0.029)	0.064*** (0.020)	0.057 (0.030)	0.068*** (0.019)	0.072** (0.026)
N	2232	2232	2232	2099	2168	2158
R2	0.02	0.02	0.02	0.02	0.02	0.02
R2 Adj.	0.02	0.02	0.02	0.02	0.02	0.02

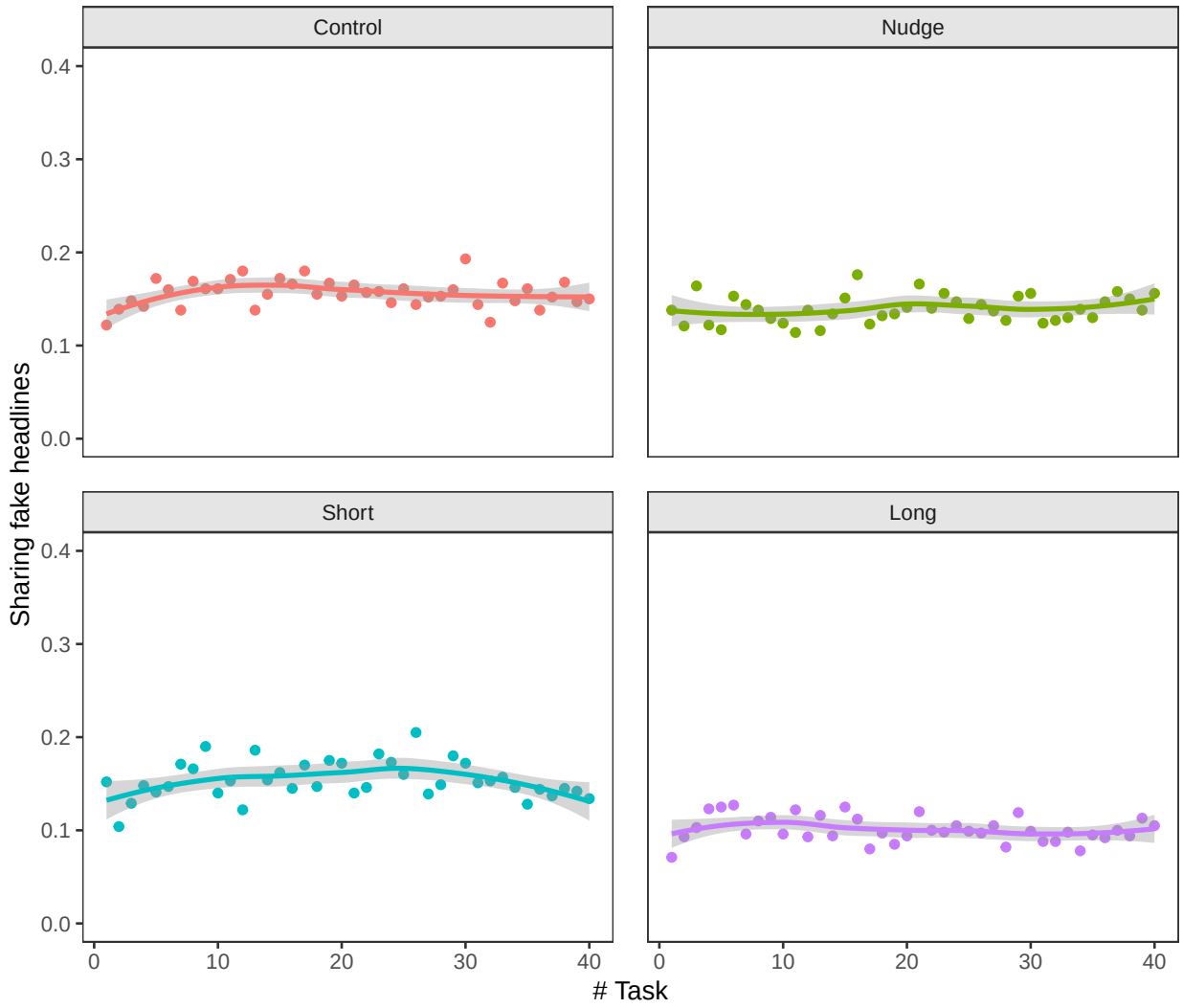
Note:
Each model shows the interaction effects between treatments and one of the covariates: Cognitive reflection, need for cognition, attention to social comparison, trust, trust in government, and trust in health authorities. Coefficients are OLS estimates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11. Effect of interventions on sharing fake and real news headlines interacted with number of tasks completed. \

	False headline sharing	Real headline sharing
Accuracy nudge (vs. control)	−0.023 (0.014)	0.019 (0.017)
15 sec video (vs. control)	−0.003 (0.015)	0.012 (0.016)
3 min video (vs. control)	−0.049*** (0.014)	0.018 (0.016)
Task order	0.000 (0.000)	0.000 (0.000)
Accuracy nudge X task order	0.000 (0.000)	0.000 (0.000)
15 sec video X task order	0.000 (0.000)	0.001 (0.000)
3 min video X task order	0.000 (0.000)	0.000 (0.000)
Constant	0.156*** (0.010)	0.249*** (0.011)
N	33 480	33 480
R2	0.01	0.00
R2 Adj.	0.01	0.00

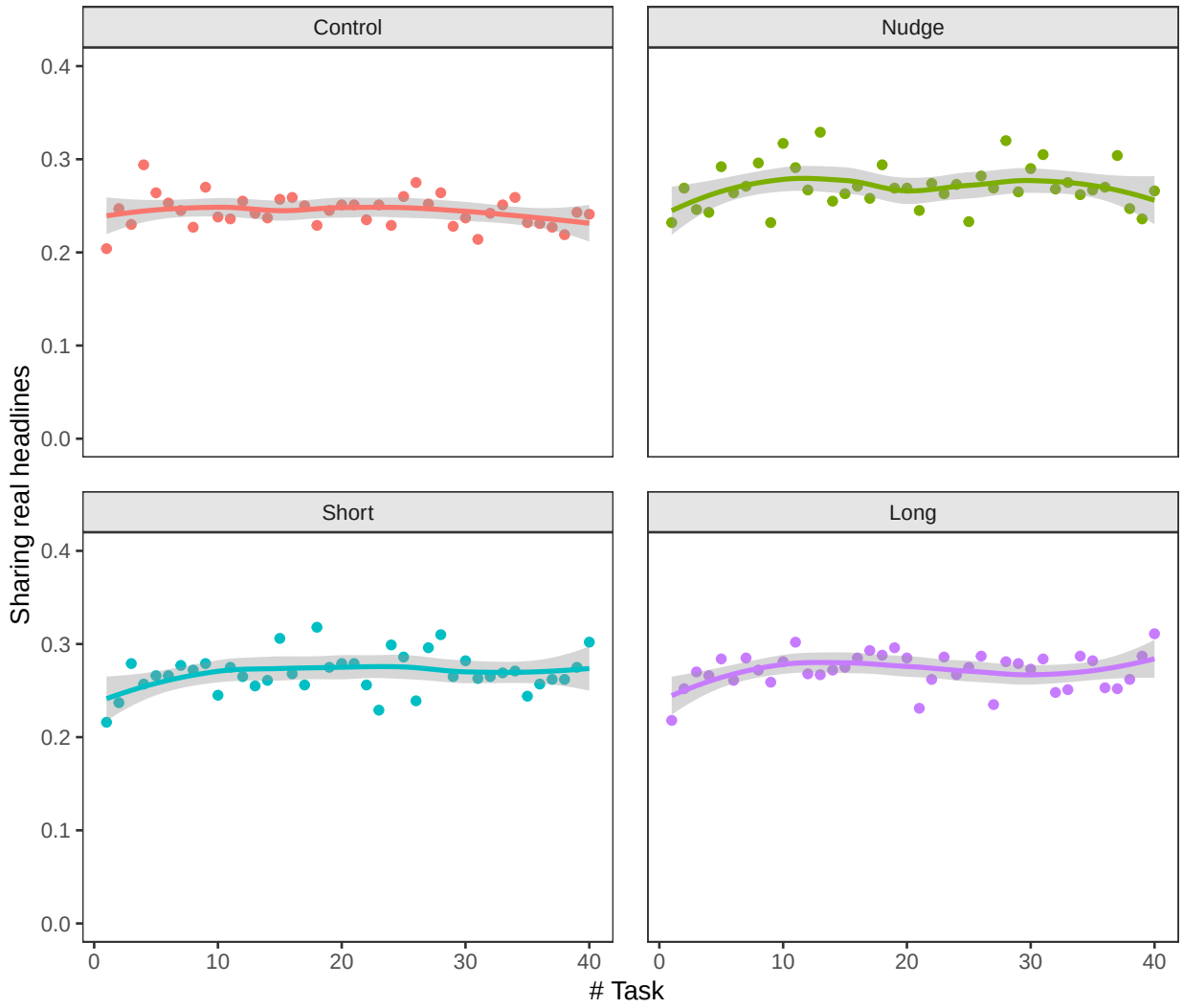
Note:
Coefficients are OLS estimates. All regressions are based on observations from 2232 respondents. Standard errors (in parentheses) are clustered at the respondent level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 3. Mean values of fake headlines sharing by treatment group across number of rating tasks completed.



Note: Means are calculated for each number of rating task.

Figure 4. Mean values of real headlines sharing by treatment group across number of rating tasks completed.



Note: Means are calculated for each number of rating task.

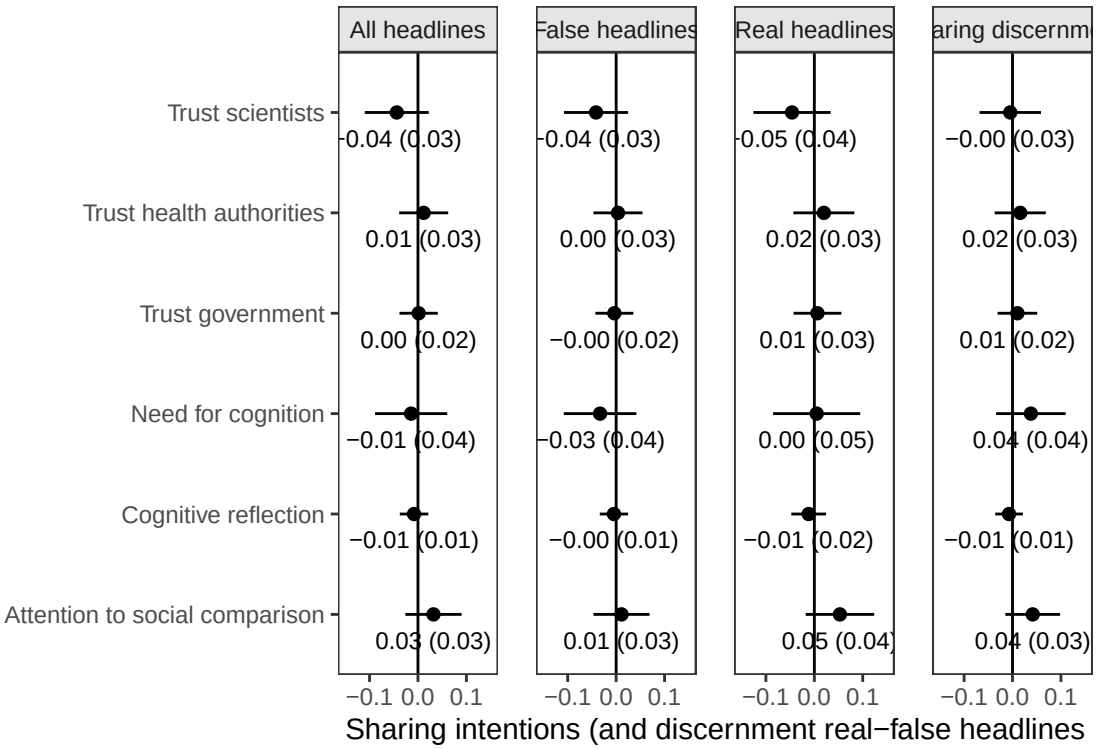
6.3 Additional supplementary analyses

6.3.1 Predictors of headline sharing - indices

We present associations between predictors of headline sharing collected in wave 1 for study 1 in figure 5. These are *trust in scientists*, *trust in health authorities*, *trust in government*, *need for cognition*, *cognitive reflection* and *attention to social comparison information*.

We collected six additional covariates that may influence the sharing of true and false headlines. As specified in the pre-registration, the main purpose of the study is to evaluate the effectiveness of the interventions, but we report these correlations here. To avoid pre-treatment effects in the experiment, these covariates were measured separately in wave 1. First, *institutional trust* is related to concern (Lieberoth et al., 2021), vaccine uptake and conspiratorial thinking (Lindholt et al., 2021) during the COVID pandemic. We measured institutional trust by asking the following question: "To what extent do you trust the following institutions regarding 1) the national health authorities, 2) the police, 3) the media, 4) scientists and 5) the government?" Second, and for similar reasons, we asked participants about their vaccination status. Third, we included a cognitive reflection task (CRT) to assess people's propensity to rely on their intuitions (Frederick, 2005). Fourth, we included a measure of political knowledge (Hansen and Stubager, 2020). Fifth, the need for cognition is associated with conspiratorial thinking (Marchlewska et al., 2018) and was measured using an 18-item scale (Cacioppo et al., 1984). Sixth, to measure the degree to which participants use information about what others do and reactions to guide their behavior, we included the attention to social comparison information (ATSCI) scale (Lennox and Wolfe, 1984).

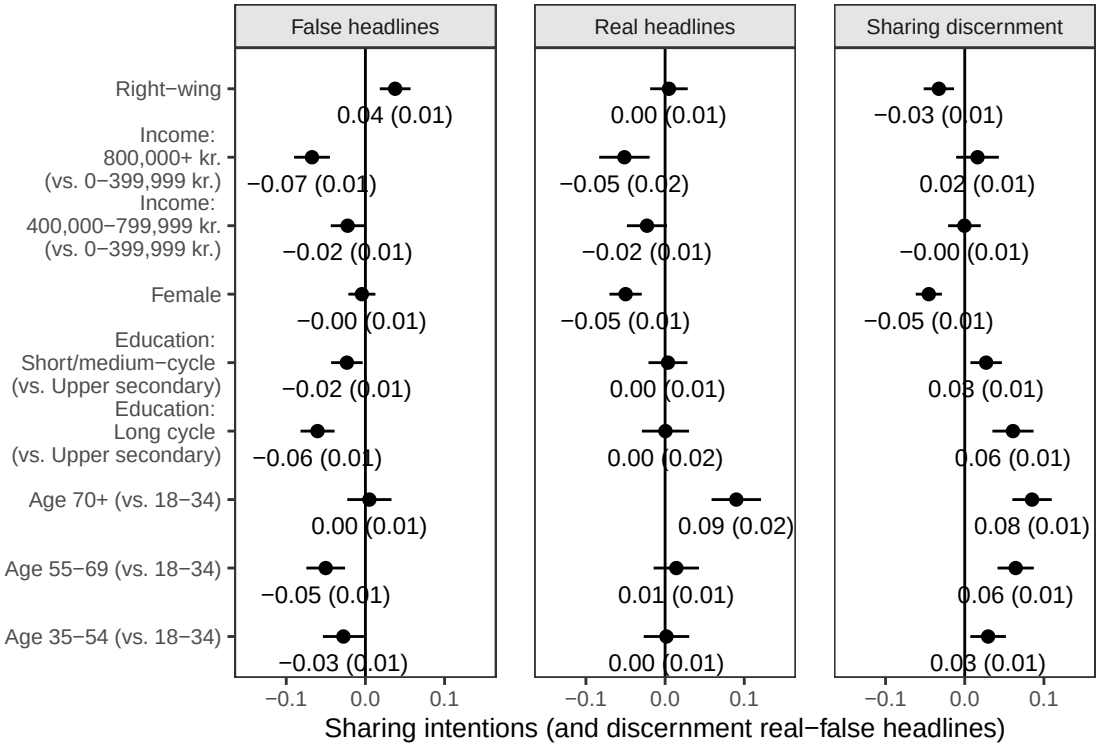
Figure 5. Associations between predictors and sharing headlines



Note: Points are OLS estimates with 95 % confidence interval bars. All models are estimated separately with each independent variable as the single predictor controlled for gender, age, income, education. The number of participants are 1801 (cognitive reflection), 1801 (need for cognition), 1801 (attention to social comparison information), 1747 (trust in government), 1736 (trust in health authorities), and 1722 (trust in scientists). The dependent variables for the four panels are sharing all headlines, sharing false headlines, sharing real headlines, and sharing discernment (scaled 0-1). Sharing discernment is calculated at the respondent level as the mean difference between real and false headline sharing between all headlines viewed: A higher sharing discernment indicates that the predictor is associated with a higher propensity to share real compared to false headlines.

6.3.2 Predictors of headline sharing - demographics

Figure 6. Associations between demographics and sharing headlines



Note: Points are OLS estimates with 95 % confidence interval bars. All models are estimated separately with each independent variable as the single predictor. The dependent variables for the panels are false and real headline sharing as well as sharing discernment (scaled 0-1). Sharing discernment is calculated at the respondent level as the mean difference between real and false headline sharing between all headlines viewed: A higher sharing discernment indicates that the predictor is associated with a higher propensity to share real compared to false headlines.

6.3.3 Partisanship interaction

Table 12 show coefficients of a regression with an interaction term between the interventions and partisanship

Table 12. Effect of interventions X partisanship on sharing. \

	False headline sharing	Real headline sharing
Accuracy nudge (vs. control)	−0.031	0.012
	(0.018)	(0.023)
15 sec video (vs. control)	0.010	0.049*
	(0.019)	(0.021)
3 min video (vs. control)	−0.035*	0.059**
	(0.017)	(0.023)
Partishanship (Right-wing)	0.050*	0.018
	(0.020)	(0.023)
Accuracy nudge X Partishanship (Right-wing)	0.013	0.021
	(0.027)	(0.032)
15 sec video X Partishanship (Right-wing)	−0.023	−0.028
	(0.029)	(0.032)
3 min video X Partishanship (Right-wing)	−0.041	−0.046
	(0.025)	(0.033)
Constant	0.134***	0.242***
	(0.014)	(0.018)
N	25 440	25 440
R2	0.01	0.00
R2 Adj.	0.01	0.00
SE clusters	by: resid & headline	by: resid & headline

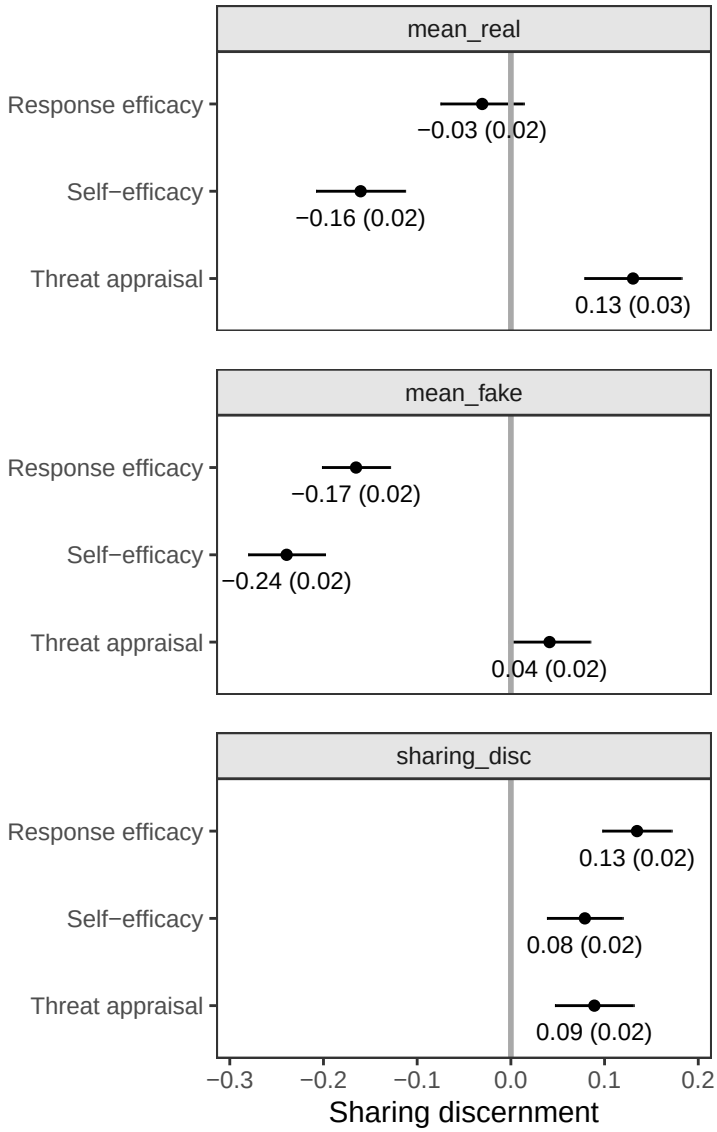
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.3.4 Are threat appraisal, self-efficacy, and response efficacy related to greater sharing discernment?

As a supplementary analysis, we pre-registered that we would "[...] test to what extent protection motivation theory (PMT) items (Risk perception, Self-efficacy, and Response efficacy) predict 1) false and 2) real headline sharing and 3) sharing discernment. We assess the association between PMT items and the sharing outcomes in models introducing one item at time and a full model with all PMT items controlling for other covariates (e.g. demographic variables). We predict that the PMT items decrease the willingness to share false headlines.

In Figure 7, we assess whether threat appraisal, self-efficacy, and response efficacy are related to sharing discernment. Both threat appraisal ($\beta = 0.09$, 95 % CI = [0.05;0.13]), self-efficacy ($\beta = 0.08$, 95 % CI = [0.04;0.12]) and response efficacy ($\beta = 0.13$, 95 % CI = [0.10;0.17]) are related to a higher sharing discernment, suggesting that all three factors predict a relative increase in real compared to false headline sharing. People share less false compared to real headlines when they hold a higher threat appraisal, feel more efficacious in terms of identifying COVID-19 misinformation and believe not sharing COVID-19 misinformation will protect themselves and others more.

Figure 7. Sharing discernment predicted by threat appraisal, self-efficacy and response efficacy



Note: Points are OLS estimates with 95 % confidence interval bars. The dependent variable (sharing discernment) is calculated at the respondent level as the mean difference between real and false headline sharing between all headlines viewed. A higher sharing discernment indicates that the predictor is associated with a higher propensity to share real compared to false headlines. Based on 2,012 respondents.

6.3.5 Replication of treatment effects in Study 2

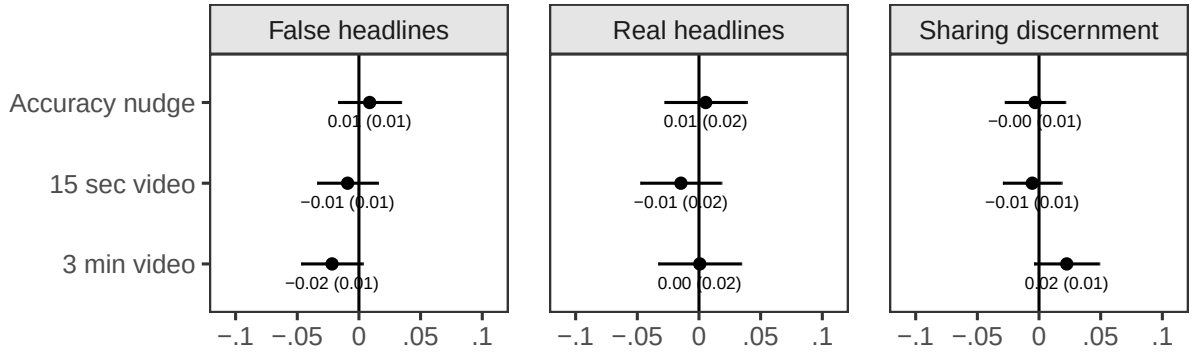
In the pre-registration for study 2, we registered the following as an additional analysis "[...] after answering the primary outcome measures, participants will complete 20 sharing tasks (10 real and 10 false headline) similar to the primary task in wave 2. We conduct the analyses using OLS with standard errors clustered on participant id, where we test the effect of each condition on the willingness to share a) real and b) false headlines respectively. **This is in principle a replication of wave 2** in our project (https://osf.io/akybg), Pennycook et al. (2020), and Roozenbeek et al. (2021), **but cannot be considered a direct replication as the 6-item battery with protection motivation theory items (Risk perception, Self-efficacy, and Response efficacy) are positioned between the intervention and the sharing task**" (our highlight). Thus, the protection motivation survey battery treats the control condition which would otherwise be untreated. We report these results in table 13, figure 8 and the coefficients here: *Accuracy nudge*: $b_{false} = .009$, p<.485, $d = .035$; $b_{real} = .005$, p<.735, $d = .017$; $b_{Sharing\ discernment}$ = -0.003, p<.797 $d = .02$. *15-seconds intervention*: $b_{false} = .009$, p<.445 $d = .038$; b_{real} = -.015, p<0.36, $d = .047$; $b_{Sharing\ discernment}$ = -0.005, p<.653 $d = .03$. *3-minute intervention*: b_{false} = -.022, p<.083 $d = .091$; $b_{real} = .001$,p<0.967, $d = .002$; $b_{Sharing\ discernment} = .022$, p<.095, $d = .11$

Table 13. Regression output of figure 8: Willingness to share headlines (Study 2).

	False headline sharing	Real headline sharing	Sharing discernment	All headlines (Intervention X veracity)
Accuracy nudge (vs. control)	0.009 (0.012)	0.005 (0.016)	−0.003 (0.012)	0.009 (0.012)
15 sec video (vs. control)	−0.009 (0.012)	−0.015 (0.016)	−0.005 (0.012)	−0.009 (0.012)
3 min video (vs. control)	−0.022 (0.012)	0.001 (0.016)	0.022 (0.013)	−0.022 (0.012)
Headline veracity (0 = false, 1 = true)				0.130 (0.013)***
Accuracy nudge X veracity				−0.003 (0.013)
15 sec video X veracity				−0.005 (0.012)
3 min video X veracity				0.022 (0.013)
Constant	0.131 (0.010)***	0.261 (0.014)***	0.130 (0.009)***	0.131 (0.010)***
N	20 120	20 120	2012	40 240
R2	0.00	0.00	0.00	0.05
R2 Adj.	0.00	0.00	0.00	0.05

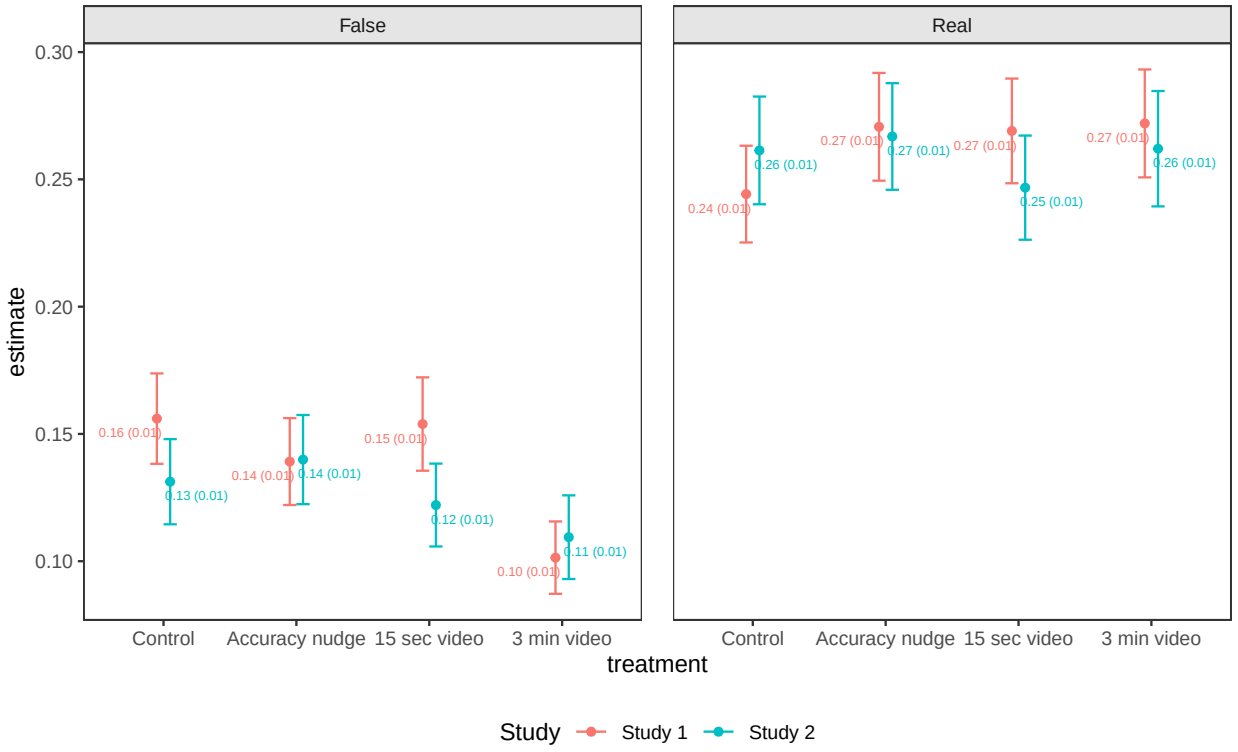
Note:
'False headline sharing' shows the regression output of the left panel in figure 9. 'Real headline sharing' shows the regression output of the middle panel. 'Sharing discernment' shows the regression output of the right panel. The discernment score is calculated as the respondent level difference between mean real and false headline sharing, that is, sharing real - sharing false. 'All headlines (Treatment X veracity)' shows a regression where all headlines are included with the interaction of the treatment interventions and the veracity of the headline on sharing intention. Coefficients are OLS estimates. Standard errors (in parentheses) are clustered at the respondent level. All regressions are based on observations from 20120 participants * p < 0.05, ** p < 0.01, *** p < 0.001

Figure 8. Replication of treatment effects on sharing false and real headlines (Study 2)



Note: The full regression table is reported in Table 13.

Figure 9. Marginal means for experimental conditions in study 1 and study 2



Note: This figure compares the marginal means for the sharing task in study 1 and study 2. The control group has a statistically significant lower mean in study 2 (Mean = 0.131) than in study 1 (Mean = 0.156) amounting to a difference of 0.0248 (p<.001). In contrast, this difference is higher for the 3-minute intervention in study 2 (mean = 0.109) than study 1 (mean = 0.101) corresponding to a difference of 0.008 (p<.052). This suggests that — in line with the preregistration — that the introduction of the pmt measures affects the mean for the control group

6.3.6 Subset on attention check and social media users

As specified in the pre-registration, "We do not plan to exclude any participants prior to the analysis. Yet, we plan to do sensitivity tests for potential moderators including 1) attention screeners and 2) Social media use". In this section, we provide these analyses:

- **Attention check**: In study 1, we included an attention check to assess to what extent people could recall the advice. Among 1097 respondents who received either the 15 sec. or 3 min. video 916 (83,5 %) gave at least one right answer when asked about which headlines they saw in the video and 371 (34%) gave the right answer on all three headlines.
- **Attention check**: Among all 2232 respondents 1963 (88%) gave the right answer when asked whether they did not see a video (control and nudge condition), saw a video that lasted less than 1 minute (15 sec. condition) or saw a video that lasted more that 1 minute (3 min. condition). Table 14 reports results of regression analyses among attentive respondents based on the second measure.
- **Social media users**: Table 15 reports results of regression analyses among respondents who are Facebook users (85.1% of the sample).

Table 14. Willingness to share headlines among attentive respondents. \

	False headline sharing	Real headline sharing	Sharing discernment	All headlines (Intervention X veracity)
Accuracy nudge (vs. control)	−0.018 (0.013)	0.025 (0.015)	0.044 (0.012)***	−0.018 (0.013)
15 sec video (vs. control)	−0.015 (0.014)	0.019 (0.015)	0.034 (0.012)**	−0.015 (0.014)
3 min video (vs. control)	−0.059 (0.012)***	0.030 (0.015)*	0.089 (0.013)***	−0.059 (0.012)***
Headline veracity (0 = false, 1 = true)				0.090 (0.008)***
Accuracy nudge X veracity				0.044 (0.012)***
15 sec video X veracity				0.034 (0.012)**
3 min video X veracity				0.089 (0.013)***
Constant	0.149 (0.009)***	0.239 (0.010)***	0.090 (0.008)***	0.149 (0.009)***
N	29 445	29 445	1963	58 890
R2	0.01	0.00	0.03	0.05
R2 Adj.	0.01	0.00	0.02	0.05

Note:
Sharing discernment is calculated as the respondent level difference between mean real and false headline sharing, that is, sharing real - sharing false. 'All headlines (Treatment X veracity)' shows a regression where all headlines are included with the interaction of the treatment interventions and the veracity of the headline on sharing intention. Coefficients are OLS estimates. Standard errors (in parentheses) are clustered at the respondent level. All regressions are based on observations from 1963 respondents. * p < 0.05, ** p < 0.01, *** p < 0.001

Table 15. Willingness to share headlines among Facebook users. \

	False headline sharing	Real headline sharing	Sharing discernment	All headlines (Intervention X veracity)
Accuracy nudge (vs. control)	−0.015 (0.014)	0.023 (0.016)	0.039 (0.012)**	−0.015 (0.014)
15 sec video (vs. control)	−0.001 (0.014)	0.018 (0.016)	0.019 (0.012)	−0.001 (0.014)
3 min video (vs. control)	−0.054 (0.013)***	0.023 (0.016)	0.077 (0.013)***	−0.054 (0.013)***
Headline veracity (0 = false, 1 = true)				0.092 (0.008)***
Accuracy nudge X veracity				0.039 (0.012)**
15 sec video X veracity				0.019 (0.012)
3 min video X veracity				0.077 (0.013)***
Constant	0.155 (0.010)***	0.247 (0.011)***	0.092 (0.008)***	0.155 (0.010)***
N	28 485	28 485	1899	56 970
R2	0.01	0.00	0.02	0.05
R2 Adj.	0.01	0.00	0.02	0.05

Note:
Sharing discernment is calculated as the respondent level difference between mean real and false headline sharing, that is, sharing real - sharing false. 'All headlines (Treatment X veracity)' shows a regression where all headlines are included with the interaction of the treatment interventions and the veracity of the headline on sharing intention. Coefficients are OLS estimates. Standard errors (in parentheses) are clustered at the respondent level. All regressions are based on observations from 1899 respondents. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

John T. Cacioppo, Richard E. Petty, and Chuan Feng Kao. The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment*, 48(3):306–307, 1984. ISSN 0022-3891, 1532-7752. doi: 10.1207/s15327752jpa4803_13.

Shane Frederick. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4):25–42, 2005. ISSN 0895-3309. doi: 10.1257/089533005775196732.

Kasper M. Hansen and Rune Stubager. The Danish National Election Study 2019. *CVAP Working Paper Series*, 2020.

Richard D. Lennox and Raymond N. Wolfe. Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology*, 46(6):1349–1364, 1984. ISSN 1939-1315. doi: 10.1037/0022-3514.46.6.1349.

Andreas Lieberoth, 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. 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), 2021. doi: 10.1098/rsos.200589.

Marie Fly Lindholt, Frederik Jø rgensen, Alexander Bor, and Michael Bang Petersen. Public acceptance of COVID-19 vaccines: Cross-national evidence on levels and individual-level predictors using observational data. *BMJ Open*, 11(6), 2021. ISSN 2044-6055. doi: 10.1136/bmjopen-2020-048172.

M. Marchlewska, A. Cichocka, and M. Kossowska. Addicted to answers: Need for cognitive closure and the endorsement of conspiracy beliefs. *European Journal of Social Psychology*, 48(2):109–117, 2018. ISSN 0046-2772. doi: 10.1002/ejsp.2308.

Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science*, 31(7):770–780, 2020. ISSN 0956-7976. doi: 10.1177/0956797620939054.

Gordon Pennycook, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand. Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855):590–595, 2021. ISSN 0028-0836. doi: 10.1038/s41586-021-03344-2.

Jon Roozenbeek, Alexandra L. J. Freeman, and Sander Van Der Linden. How Accurate Are Accuracy-Nudge Interventions? A Preregistered Direct Replication of Pennycook et al. (2020). *Psychological Science*, page 09567976211024535, 2021. ISSN 0956-7976. doi: 10.1177/09567976211024535.