Appendix

Contents

1 Stimuli examples **33**

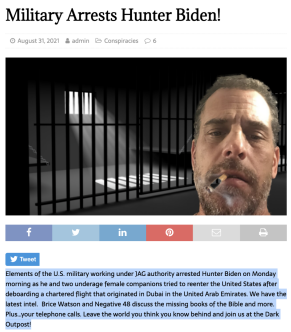
2 Tables **36**

3 Appendix study stimuli **43**

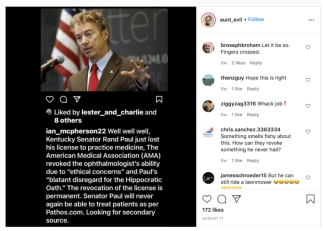
4 Study 2 and 3 stimuli **45**

5 Pre-registration documents **52**

1 Stimuli examples



(a) Fake news (Hunter Biden)



(b) Fake news (Rand Paul)

(c) (Fake news Japan

Figure A1: Fake content used in study 1

Figure A2: Examples of non-fact-checked content in Studies 1 and 2

2 Tables

Table A1: Summary of tested stimuli

Studies in which stimuli appeared	*Type*	*Summary*
1	Fact-checked false	Rand Paul loses medical license
1	Fact-checked false	Graves in Japan have QR codes
1	Fact-checked false	Military arrests Hunter Biden
2, 3	Fact-checked false	Ivanka Trump arrested
2, 3	Fact-checked false	Biden caused gas prices to rise
2, 3	Fact-checked false	Death of Ukrainian fighter pilot
2, 3	Fact-checked true	Southwest Airlines will not enforce mask mandate
2, 3	Not fact-checked	Cute baby photo
2, 3	Not fact-checked	Rihanna photo
2, 3	Not fact-checked	Crocodile photo
2, 3	Not fact-checked	Bruce Lee and Kareem Abdul-Jabbar photos
2, 3	Not fact-checked	Bruce Lee photos
2, 3	Not fact-checked	Saying about rainbows
2, 3	Not fact-checked	Joke about dogs
2, 3	Not fact-checked	Meme about reading
2, 3	Not fact-checked	Inspirational saying
2, 3	Not fact-checked	Chucky character
2, 3	Not fact-checked	Peanuts cartoon
2, 3	Not fact-checked	Bunnies cartoon

Table A2: Study 1 Balance

	Control		Treatment			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Biden Approval	1.9	1.0	2.0	1.0	0.0	0.0
Education	8.0	1.7	8.0	1.7	0.0	0.1
Income	12.4	6.2	12.2	6.3	-0.2	0.2
Age	38.7	11.3	39.3	12.1	0.6	0.4
Race	1.5	2.1	1.5	1.9	-0.1	0.1
Party ID	3.1	2.4	3.1	2.4	0.0	0.1

Table A3: Study 1: Effects on Belief Accuracy

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.14	0.03	-4.41	0.00	-0.21	-0.08	2862.00
2	(Intercept)	2.53	0.02	109.41	0.00	2.48	2.58	2862.00

Table A4: Study 1: Effects on Sharing Intention

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.25	0.06	-4.02	0.00	-0.37	-0.13	2862.00
2	(Intercept)	3.33	0.04	76.07	0.00	3.25	3.42	2862.00

Table A5: Study 1: Belief Accuracy and Political Partisanship

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	0.04	0.17	0.25	0.80	-0.30	0.38	738.00
2	Republican	-0.81	0.09	-9.52	0.00	-0.98	-0.65	738.00
3	Treatment:Republican	-0.06	0.12	-0.52	0.60	-0.30	0.17	738.00
4	(Intercept)	3.64	0.12	29.38	0.00	3.39	3.88	738.00

Table A6: Study 1: Sharing Intention and Political Partisanship

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.13	0.32	-0.42	0.67	-0.76	0.49	738.00
2	Republican	-1.69	0.16	-10.55	0.00	-2.00	-1.37	738.00
3	Treatment: Republican	-0.04	0.22	-0.19	0.85	-0.48	0.40	738.00
4	(Intercept)	5.77	0.22	26.05	0.00	5.33	6.20	738.00

Table A7: Study 1: Belief Accuracy and Cognitive Reflection

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.20	0.08	-2.63	0.01	-0.35	-0.05	2858.00
2	CRT High	-0.49	0.06	-7.70	0.00	-0.62	-0.37	2858.00
3	CRT Low	0.19	0.06	3.01	0.00	0.07	0.32	2858.00
4	Treatment:CRT High	0.10	0.09	1.15	0.25	-0.07	0.27	2858.00 1
5	Treatment:CRT Low	0.06	0.09	0.70	0.48	-0.11	0.24	2858.00
6	(Intercept)	2.67	0.05	48.63	0.00	2.56	2.78	2858.00

Table A8: Study 1: Sharing Intention and Cognitive Reflection

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.27	0.14	-1.96	0.05	-0.53	-0.00	2858.00
2	CRT High	-0.96	0.11	-8.36	0.00	-1.18	-0.73	2858.00
3	CRT Low	0.42	0.11	3.70	0.00	0.20	0.65	2858.00
4	Treatment:CRT High	0.09	0.16	0.55	0.58	-0.23	0.41	2858.00
5	Treatment:CRT Low	0.01	0.16	0.08	0.94	-0.31	0.34	2858.00
6	(Intercept)	3.58	0.10	37.41	0.00	3.39	3.77	2858.00

Table A9: Study 2 Balance

	Control		Placebo		Treatment	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Biden Approval	1.8	0.9	1.8	0.9	1.8	0.9
Education	8.1	1.7	8.2	1.6	8.2	1.5
Income	12.1	6.2	12.4	6.1	12.5	6.3
Age	38.1	11.5	37.2	10.6	36.8	11.0
Race	1.5	2.1	1.6	2.3	1.6	2.2
Party ID	3.2	2.6	3.2	2.6	3.1	2.5

Table A10: Study 2: Effects on Sharing False Content

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.16	0.07	-2.20	0.03	-0.31	-0.02	2423.00
2	Placebo	0.04	0.07	0.48	0.63	-0.11	0.18	2423.00
2	(Intercept)	4.06	0.05	78.21	0.00	3.96	4.16	2423.00

Table A11: Study 2: Effects on Sharing Non-Fact-Checked Content

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.07	0.06	-1.18	0.24	-0.19	0.05	2419.00
2	Placebo	0.04	0.06	0.66	0.51	-0.08	0.17	2419.00
3	(Intercept)	4.29	0.04	96.20	0.00	4.20	4.38	2419.00

Table A12: Study 2: Effects on Sharing Fact-Checked True Content

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.20	0.08	-2.47	0.01	-0.35	-0.04	2429.00
2	Placebo	0.02	0.08	0.31	0.76	-0.13	0.18	2429.00
3	(Intercept)	4.06	0.06	71.94	0.00	3.95	4.17	2429.00

Table A13: Study 2: Sharing False Content Intention and Partisanship

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	0.54	0.31	1.73	0.08	-0.07	1.16	706.00
2	Republican	-1.74	0.17	-10.48	0.00	-2.06	-1.41	706.00
3	Treatment:Republican	-0.56	0.26	-2.14	0.03	-1.08	-0.05	706.00
4	(Intercept)	6.39	0.19	33.62	0.00	6.01	6.76	706.00

Table A14: Study 2: Sharing False Content Intention and CRT

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.19	0.13	-1.45	0.15	-0.45	0.07	2420.00
2	CRT High	-0.77	0.10	-7.58	0.00	-0.96	-0.57	2420.00
3	CRT Low	0.20	0.09	2.18	0.03	0.02	0.38	2420.00
4	Treatment:CRT High	-0.10	0.17	-0.59	0.56	-0.44	0.24	2420.00
5	Treatment:CRT Low	0.09	0.15	0.59	0.56	-0.21	0.40	2420.00
6	(Intercept)	4.31	0.08	55.33	0.00	4.16	4.46	2420.00

Table A15: Study 3 Balance

	Control		Placebo		Accuracy		Cialdini	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Biden Approval	1.8	0.9	1.8	1.0	1.8	1.0	1.8	0.9
Education	8.2	1.7	8.2	1.7	8.2	1.6	8.2	1.6
Income	11.6	5.9	11.8	5.9	12.0	6.1	11.8	5.7
Age	37.7	11.9	38.1	11.5	37.2	11.6	37.6	11.8
Race	1.4	1.6	1.4	1.9	1.5	2.0	1.4	1.9
Party ID	3.4	2.6	3.3	2.6	3.3	2.5	3.3	2.6

Table A16: Study 3: Effects on Sharing Fact-Checked True Content

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	
1	Treatment	-0.16	0.07	-2.36	0.02	-0.29	-0.03	4824.00
2	Placebo	-0.12	0.07	-1.77	0.08	-0.25	0.01	4824.00
3	Accuracy Nudge	-0.05	0.07	-0.71	0.48	-0.18	0.09	4824.00
4	(Intercept)	4.00	0.05	84.40	0.00	3.90	4.09	4824.00

Table A17: Study 3: Sharing False Content Intention and Partisanship

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	0.07	0.26	0.26	0.80	-0.44	0.57	1471.00
2	Republican	-1.93	0.11	-17.46	0.00	-2.15	-1.72	1471.00
3	Treatment:Republican	-0.18	0.22	-0.80	0.42	-0.61	0.26	1471.00
4	(Intercept)	6.51	0.13	50.25	0.00	6.26	6.76	1471.00

Table A18: Study 3: Sharing False Content Intention and CRT

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.05	0.10	-0.50	0.62	-0.26	0.15	4814.00
2	CRT High	-0.51	0.07	-7.46	0.00	-0.64	-0.37	4814.00
3	CRT Low	0.29	0.07	4.35	0.00	0.16	0.42	4814.00
4	Treatment:CRT High	-0.08	0.13	-0.57	0.57	-0.33	0.18	4814.00
5	Treatment:CRT Low	-0.14	0.13	-1.11	0.27	-0.39	0.11	4814.00
6	(Intercept)	4.10	0.05	75.02	0.00	3.99	4.21	4814.00

3 Appendix study stimuli

Literacy intervention

"Tips to Spot False News"

Be skeptical of headlines. False news stories often have catchy headlines in all caps with exclamation points. If shocking claims in the headline sound unbelievable, they probably are.

Look closely at the URL. A phony or look-alike URL may be a warning sign of false news. Many false news sites mimic authentic news sources by making small changes to the URL. You can go to the site to compare the URL to established sources.

Investigate the source. Ensure that the story is written by a source that you trust with a reputation for accuracy. If the story comes from an unfamiliar organization, check their "About" section to learn more.

Watch for unusual formatting. Many false news sites have misspellings or awkward layouts. Read carefully if you see these signs.

Consider the photos. False news stories often contain manipulated images or videos. Sometimes the photo may be authentic, but taken out of context. You can search for the photo or image to verify where it came from.

Inspect the dates. False news stories may contain timelines that make no sense, or event dates that have been altered.

Check the evidence. Check the author's sources to confirm that they are accurate. Lack of evidence or reliance on unnamed experts may indicate a false news story.

Look at other reports. If no other news source is reporting the same story, it may indicate that the story is false. If the story is reported by multiple sources you trust, it's more likely to be true.

Is the story a joke? Sometimes false news stories can be hard to distinguish from humor or satire. Check whether the source is known for parody, and whether the story's details and tone suggest it may be just for fun.

Some stories are intentionally false. Think critically about the stories you read, and only share news that you know to be credible.

Figure A3: Literacy intervention

Accuracy nudge

We would also like to pretest actual news headline from 2016 and 2017 for future studies. We are interested in whether people think it is accurate or not.

We only need you to give your opinion about the accuracy of a single headline.

Note: The image may take a moment to load.

Figure A4: Accuracy nudge screen 1

Because Of The Lack Of Men, Iceland Gives $5,000 Per Month To Immigrants Who Marry Icelandic Women!

Breaking news about Iceland country incredible but true if you are interested read the full story Iceland team was able to achieve an unprecedented achievement in the European…

HOWAFRICA.COM

+ Add page break

neutral_fake_impl

To the best of your knowledge, how accurate is the claim in the above headline?

Not at all accurate	Not very accurate	Somewhat accurate	Very accurate
○	○	○	○

Figure A5: Accuracy nudge screen 2

Q8

Woman who had ovary frozen in childhood gives birth

She is believed to be the first woman in the world to have a baby after having ovarian tissue frozen before the onset of puberty

CBSNEWS.COM

+ Add page break

neutral_real_plaus

To the best of your knowledge, how accurate is the claim in the above headline?

Not at all accurate	Not very accurate	Somewhat accurate	Very accurate
○	○	○	○

Figure A6: Accuracy nudge screen 3

4 Study 2 and 3 stimuli

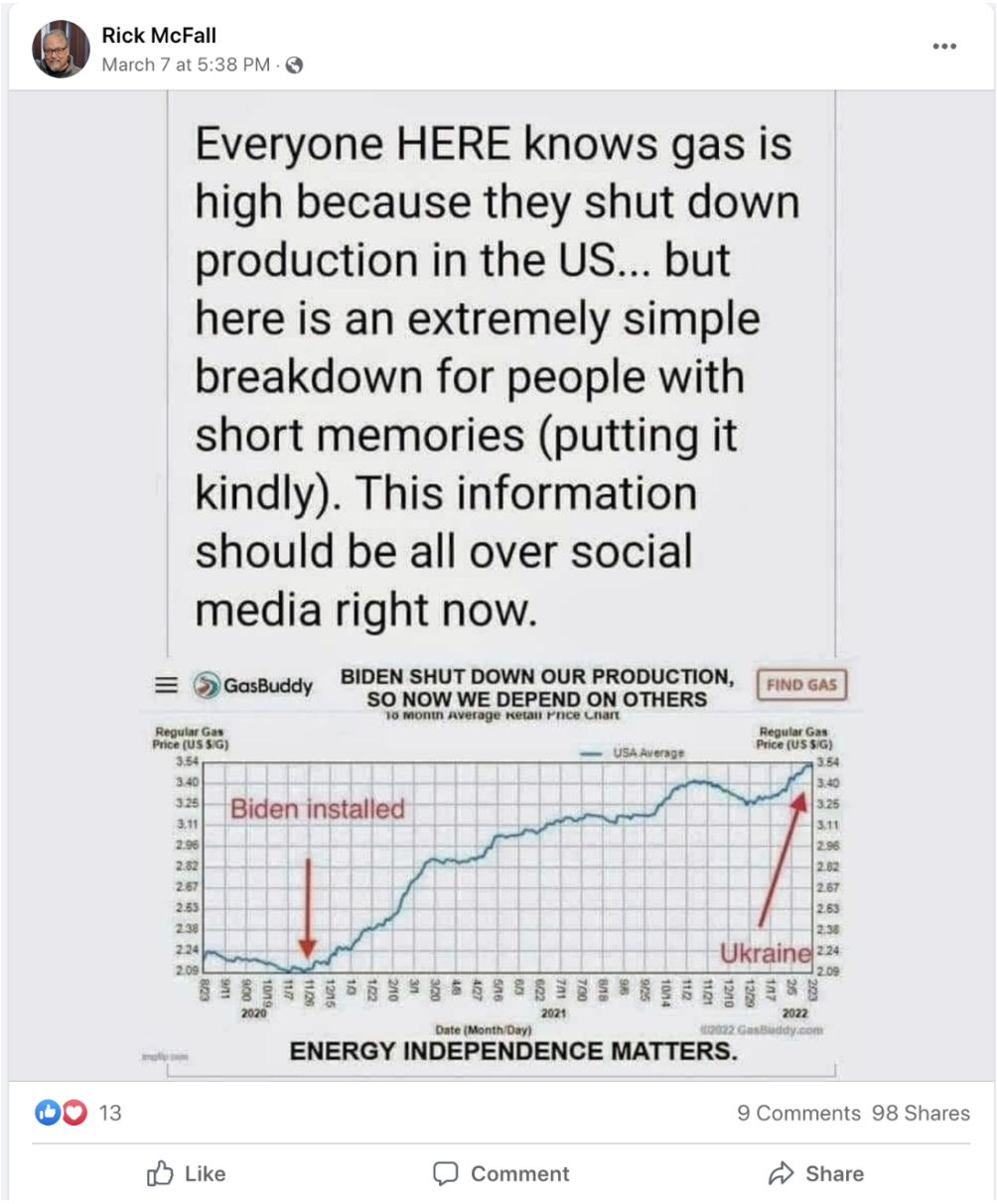


Figure A7: Biden post fact-checked as false

Figure A8: Trump post fact-checked as false

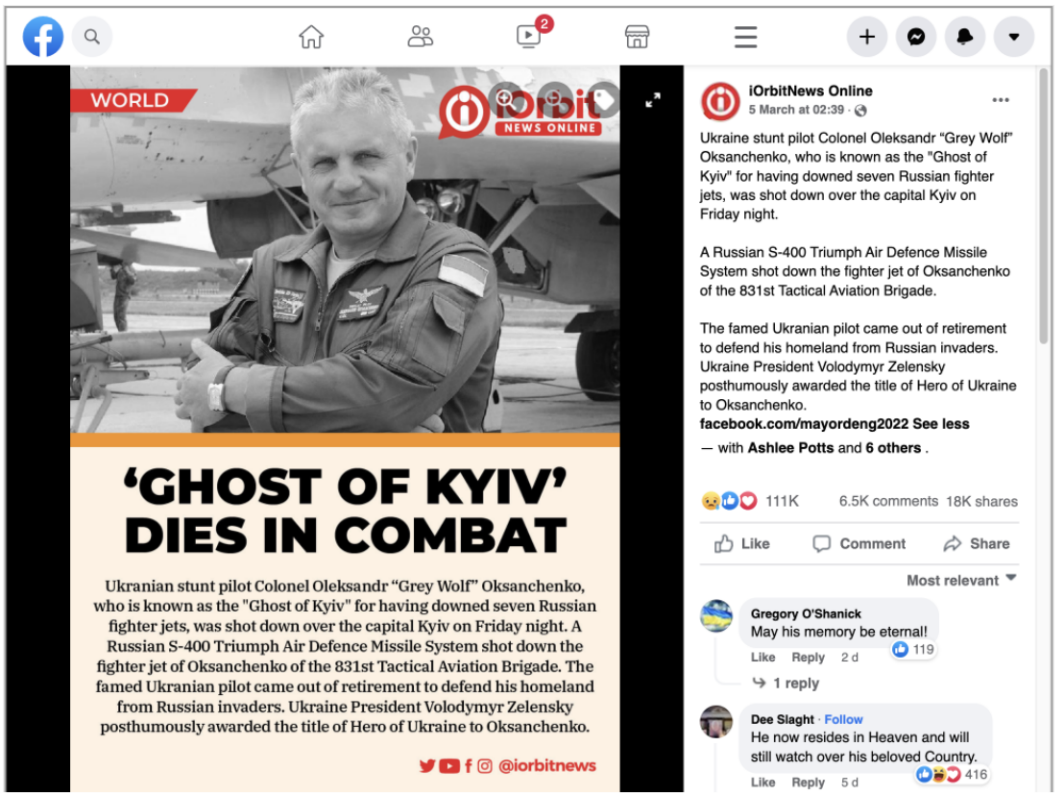


Figure A9: Ukraine post fact-checked as false

If you were to see the above article on Facebook, Instagram or other social media, how likely would you be to share it?

Figure A10: Post fact-checked as true

Appendix study

How well does our intervention perform compared to other interventions which target similar beliefs? A study conducted in September 2021 on Mechanical Turk (n=1,032) answers this question. After answering demographic and other important pre-treatment questions, participants were randomized to our treatment; an accuracy nudge; a news literacy intervention; or control. The nudge and literacy interventions mimicked prior interventions in this literature, with content taken directly from Pennycook et al. (2020) and Guess et al. (2020b), respectively.

The complete text of these interventions can be found in the appendix, while the Cialdini intervention was the same as described above. The outcomes were similar to that used in the earlier study. Once again, participants were asked to assess their belief in, and willingness to share, misinformation that denigrated Democrats, another that denigrated Republicans, and a third that was apolitical. The misinformation items themselves were the same as used in the prior study (and depicted in Figure A1). Following Pennycook et al. (2020), we restrict our analysis to those who report prior to treatment that they have Facebook accounts.

As with Study 1, we evaluate effects by creating indices of responses to all items. Tables A19 and A20 display results. On accuracy, our intervention is the only one to clear conventional thresholds of statistical significance, reducing belief in misinformation by .16 ($p < .05$). However, it is not statistically distinct from either of the other two interventions. No intervention affects sharing intent.

Table A19: Appendix Study: Effects on Belief Accuracy

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.16	0.08	-2.02	0.04	-0.32	-0.00	835.00
2	News Literacy	-0.14	0.08	-1.74	0.08	-0.29	0.02	835.00
3	Accuracy Nudge	-0.08	0.08	-1.10	0.27	-0.24	0.07	835.00
4	(Intercept)	2.89	0.05	55.98	0.00	2.79	2.99	835.00

Table A20: Appendix Study: Effects on Sharing Intent

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.17	0.12	-1.45	0.15	-0.40	0.06	835.00
2	News Literacy	-0.12	0.12	-1.03	0.30	-0.35	0.11	835.00
3	Accuracy Nudge	-0.13	0.11	-1.16	0.25	-0.35	0.09	835.00
4	(Intercept)	3.45	0.08	44.08	0.00	3.29	3.60	835.00

In tables A21 and A22, we investigate whether effects of the Cialdini-inspired intervention differ by partisanship and CRT, using the same approach as in Study 1. Again, we find no evidence that it does.

These results offer further evidence that, at the least a) our intervention can improve belief accuracy, and b) these improvements are indistinguishable from improvements generated by other interventions.

Table A21: Appendix Study: Belief Accuracy and Partisanship

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.00	0.08	-0.03	0.97	-0.16	0.15	835.00
2	Republican	0.27	0.07	3.83	0.00	0.13	0.40	835.00
3	Treatment:Republican	-0.28	0.16	-1.81	0.07	-0.59	0.02	835.00
4	(Intercept)	2.74	0.04	71.63	0.00	2.66	2.81	835.00

Table A22: Appendix Study: Belief Accuracy and CRT

	term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
1	Treatment	-0.16	0.10	-1.61	0.11	-0.36	0.04	1000.00
2	CRT High	-0.93	0.07	-12.36	0.00	-1.07	-0.78	1000.00
3	CRT Low	0.24	0.06	4.19	0.00	0.13	0.35	1000.00
4	Treatment:CRT High	-0.09	0.16	-0.56	0.58	-0.41	0.23	1000.00
5	Treatment:CRT Low	0.19	0.12	1.62	0.11	-0.04	0.43	1000.00
6	(Intercept)	2.89	0.05	60.72	0.00	2.79	2.98	1000.00

5 Pre-registration documents

Study 1 pre-registration

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Pivotal Message (#82485)

Created: 12/08/2021 02:40 PM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
H1: A message which communicates to subjects that they are pivotal in solving the misinformation crisis will reduce intent-to-share false social media posts
H2: A message which communicates to subjects that they are pivotal in solving the misinformation crisis will reduce belief in false social media posts

RQ1: Will the effects of a message which communicates to subjects that they are pivotal in solving the misinformation crisis vary by partisanship?
RQ2: Will the effects of a message which communicates to subjects that they are pivotal in solving the misinformation crisis vary by CRT-2 responses?

3) Describe the key dependent variable(s) specifying how they will be measured.
Participants will be shown 6 social media posts in random order. After each post, in random order, we will evaluate intent-to-share and belief accuracy.

To evaluate intent-to-share, participants will be asked: "If you were to see the above article on Facebook, Instagram or other social media, how likely would you be to share it?" Responses will be available on a 1-6 scale, ranging from "Extremely unlikely" to "Extremely likely."

To evaluate belief accuracy, participants will be asked: "To the best of your knowledge, how accurate is this statement?" A statement summarizing the post will follow. Responses will be available on a 1-5 scale, ranging from "Not at all accurate" to "Very accurate."

We will then scale responses into a Intent-to-Share Index and a Belief Accuracy Index.

Only responses from the three posts fact-checked as "Fake" by PolitiFact will be included in the indices and subsequent analyses. Specifically, we will aggregate and evaluate responses to a post about Hunter Biden; a post about Rand Paul; and a post about Japanese cemeteries. The others are included to reduce demand effects.

4) How many and which conditions will participants be assigned to?
Participants will be randomly assigned with equal probability to one of the two following conditions:
-A control condition, in which they answer only outcome questions for each post;
-A Pivotal respondent condition, in which they see the message which communicates to subjects that they are pivotal in solving the misinformation crisis, and then answer outcome questions for each post

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
To test H1, we will regress the Intent-to-Share Index on treatment assignment with HC2 robust standard errors.
To test H2, we will regress the Belief Accuracy Index on treatment assignment with HC2 robust standard errors.

To test RQ1, we will estimate variants of the models used to test H1 and H2, but include an indicator for respondents who identify with or lean toward the Republican Party and an interaction term between each indicator in the model and the Republican indicator.

To test RQ2, we we will include variants of the models used to test H1 and H2, but include indicators for respondents who score in the top tercile on CRT-2, in the middle tercile, and the bottom tercile, and interaction terms between those indicators and treatment assignment.

Again, as stated above, we will only analyze responses to the three posts that have been fact-checked as "False" by PolitiFact in our analyses. Specifically, we will aggregate and evaluate responses to a post about Hunter Biden; a post about Rand Paul; and a post about Japanese cemeteries.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will only analyze responses by participants who, prior to treatment, reply that they have a Facebook account and/or a Twitter account and/or an Instagram account. These questions will be worded as follows: "Do you have a [Facebook/Twitter/Instagram] account?" with Yes/No response options. Only those who answer "Yes" to one of these questions will be included in the final sample.

We will only analyze responses by participants who, prior to treatment, answer an attention check correctly. The attention check will be worded as follows:

"'People are very busy these days and many do not have time to follow what goes on in the government. We are testing whether people read questions. To show that you've read this much, answer both "extremely interested" and "very interested.'" Only those who select "extremely interested" and "very interested" will be included in the final sample.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
3,200 participants will be recruited. As described above, the final sample will be determined by pre-treatment responses to an attention check and whether participants report having a Facebook, Twitter or Instagram account.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
-We will report results without covariates, contingent on conducting balance tests for the following covariates: Party ID; sex; race; education; income; CRT. If we are dissatisfied with balance, we may report results that account for the covariates that are imbalanced.

-Significance threshold: $p < 0.05$ (two-sided).

-The message which communicates to subjects that they are pivotal in solving the misinformation crisis reads as follows (without the question marks and in bold text):

"Misinformation is a serious problem. Around the world, people go on social media to spread fake news and tell lies. Unfortunately, some people believe those lies.

You have the power to make a difference. When you see fake news, you should call it out for what it is: FAKE."

-We may also investigate effects on the non-false posts. Specifically, they are posts about the Texas power grid; correlations between marijuana and psychotic disorders; and the percentage of Wisconsites who support fair electoral maps. All posts were fact-checked as "True" by PolitiFact. Our investigation of this particular matter would be preliminary. Once again, we are only testing our hypotheses and questions on the posts fact-checked as "False" by PolitiFact.

Study 2 pre-registration

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Cialdini 2022--Realistic Environment (#91033)

Created: 03/15/2022 03:25 PM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
H: Exposure to a message emphasizing each individual's ability to help solve the misinformation crisis will reduce willingness to share misinformation on social media.

3) Describe the key dependent variable(s) specifying how they will be measured.
We will construct an index of willingness to share fake items by taking the mean willingness of three fake items included in our set of 16. The three fake items are: a social media post alleging that Ivanka Trump has been arrested; a social media post asserting that, under Joe Biden, the U.S. has stopped gas production; and a social media post alleging that a famed fighter pilot has died in the Russia-Ukraine war.

Following each item, participants are asked: "If you were to see the above article on Facebook, Instagram or other social media, how likely would you be to share it?" Answers range from "Extremely unlikely" to "extremely likely" on a 1-6 scale. To create the index that will serve as our independent variable, we will take the mean response to the three fake items.

4) How many and which conditions will participants be assigned to?
Three conditions:
Treatment message: Participants see a treatment message.
Placebo message: Participants see a placebo message.
Control: Participants do not see either of the above.

The treatment message is: "Misinformation is a serious problem. Around the world, people go on social media to spread fake news and tell lies. Unfortunately, some people believe those lies. You have the power to make a difference. When you see fake news, you should call it out for what it is: FAKE. "

The placebo message is: "Five sauces for the modern cook

By Samrin Nosrat

Travis Lett often steals. Of course, the only person this pensive chef ever steals from is himself. At his Los Angeles, USA restaurant, "We're constantly appropriating elements from dishes we've done in the past to create new combinations," he said.

There's a lesson here: To improve your cooking, learn how to make and use sauce like a professional.

Five basic types of sauces appear over and over again on menus and in cookbooks that feature the kind of vegetable-heavy, flavor-dense food that cooks and eaters favor today: yogurt sauce, pepper sauce, herb sauce, tahini sauce and pesto. Master each one, and you'll immediately have access to the dozens of variations that descend from them, too.

Think of them as the new mother sauces, an updated version of the five mother sauces of French cuisine. Armed with one of these five sauces, the home cook can go on and cook what he or she is most comfortable cooking. The right sauce will transform the distinct elements of a dish into a unified statement of taste."

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will regress the willingness-to-share fake items index on treatment assignment (alpha = .05). We will estimate this model without covariates. In the appendix, we will present results from a model that includes the following covariates: Party ID; household income; education; race; Biden approval; gender; state.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will exclude people who do not report having either a Facebook or an Instagram or a Twitter account prior to treatment. Prior to treatment, we will ask

an attention check question; participants who fail it will not move forward in the survey.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
1500

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
1. As exploratory matters, we will consider the interaction between treatment effects and responses to the Cognitive Reflection Test and Party ID. To evaluate these, we will interact treatment assignment with the bottom tercile of CRT responses (for CRT) and Republican identity (for Party ID).
2. To assess treatment effects on an item that has been fact-checked as "true," we will estimate the effects of willingness to share one post in particular. The post says: "And just like that, Southwest Airlines announce they will no longer terminate employees over the mandate."
3. To assess spillover of treatment effects on non-fake, non-fact-checked items, we will compile an index of willingness to share of the 12 remaining non-fake, non-fact-checked items, and estimate the same models as used to evaluate our primary hypothesis. We will compare the coefficient of the treatment effect from this model with the coefficient of the treatment effect from the model that estimates effects on willingness to share fake items.

Study 3 pre-registration

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Cialdini v Accuracy Nudge (#104537)

Created: 08/11/2022 01:12 PM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
This is both a replication and extension of a previous pre-registration ("Cialdini 2022--Realistic Environment" (#91033)). We will be evaluating the same hypothesis, while also evaluating a new intervention.

H1: Exposure to a message emphasizing each individual's ability to help solve the misinformation crisis will reduce willingness to share misinformation on social media.

RQ1: Will effects from the message emphasizing each individual's ability to help solve the misinformation crisis be distinct from effects attributable to an accuracy nudge?

3) Describe the key dependent variable(s) specifying how they will be measured.
We will construct an index of willingness to share fake items by taking the mean willingness of three fake items included in our set of 16. The three fake items are: a social media post alleging that Ivanka Trump has been arrested; a social media post asserting that, under Joe Biden, the U.S. has stopped gas production; and a social media post alleging that a famed fighter pilot has died in the Russia-Ukraine war.

Following each item, participants are asked: "If you were to see the above article on Facebook, Instagram or other social media, how likely would you be to share it?" Answers range from "Extremely unlikely" to "extremely likely" on a 1-6 scale. To create the index that will serve as our independent variable, we will take the mean response to the three fake items.

4) How many and which conditions will participants be assigned to?
Four conditions: Pivotal message: Participants see a pivotal message. (This condition was called the "treatment" in pre-registration #91033. It is identical to that condition.) Placebo message: Participants see a placebo message. Accuracy nudge: Participants see an accuracy nudge. Control: Participants do not see any of the above.

All conditions but the accuracy nudge are identical those described in pre-reg #91033.
The accuracy nudge is as follows. It consists of three screens. On the first screen, participants see the following: "First, we would like to pretest an actual news headline from 2016 and 2017 for future studies. We are interested in whether people think it is accurate or not. We only need you to give your opinion about the accuracy of a single headline. We will then continue on to the primary task. Note: The image may take a moment to load."

Then, participants see, in random order, headlines from a neutral fake implausible story and a neutral real implausible story, and are asked to share their assessment of each story's accuracy. The neutral fake implausible story headline is "Because of the lack of men, Iceland Gives $5,000 Per Month to Immigrants Who Marry Icelandic Women!" The neutral real implausible story headline is "Woman who had ovary frozen in childhood gives birth."

Under each headline, participants are asked: "To the best of your knowledge, how accurate is the claim in the above headline?" Response options are "Not at all accurate/Not very accurate/Somewhat accurate/Very accurate." Note that these questions are part of the nudge condition and are not outcome measures.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
To evaluate H1, we will regress the willingness-to-share fake items index on treatment assignment. The hypothesis will be evaluated by inspecting the significance level for the coefficient for assignment to the pivotal message (alpha = .05). The control condition will be omitted. We will estimate this model without covariates. In the appendix, we will present results from a model that includes the following covariates: Party ID; household income; education; race; Biden approval; gender; state.

To evaluate RQ1, we will compare the coefficients for the pivotal message and the accuracy nudge. Depending on results, we may also use the two one-sided test procedure.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will exclude people who do not report having either a Facebook or an Instagram or a Twitter account prior to treatment. Prior to treatment, we will ask an attention check question; participants who fail it will not move forward in the survey.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
5000

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
1. As exploratory matters, we will consider the interaction between treatment effects and responses to the Cognitive Reflection Test and Party ID. To evaluate these, we will interact treatment assignment with the bottom tercile of CRT responses (for CRT) and Republican identity (for Party ID).
2. To assess treatment effects on an item that has been fact-checked as "true," we will estimate the effects of willingness to share one post in particular. The post says: "And just like that, Southwest Airlines announce they will no longer terminate employees over the mandate."
3. To assess spillover of treatment effects on non-fake, non-fact-checked items, we will compile an index of willingness to share of the 12 remaining non-fake, non-fact-checked items, and estimate the same models as used to evaluate our primary hypothesis. We will compare the coefficient of the treatment effect from this model with the coefficient of the treatment effect from the model that estimates effects on willingness to share fake items.