1. Title page

Title: Vaccine Hesitancy and Betrayal Aversion

Authors: Abdelaziz Alsharawy[1ς], Esha Dwibedi[2], Jason Aimone[3] and Sheryl Ball [2,4]

[1] School of Public and International Affairs, Princeton University, Princeton, NJ, 08550, USA

[2] Department of Economics, Virginia Tech, Blacksburg, VA, 24061, USA

[3] Department of Economics, Baylor University, Waco, TX, 76798, USA

[4] School of Neuroscience, Virginia Tech, Blacksburg, VA, 24061, USA

[ς] The first author began working on this project as a PhD student in the Department of Economics at Virginia Tech, Blacksburg, VA, 24061, USA

Abbreviated title for the running head: Vaccine Hesitancy and Betrayal Aversion

Corresponding author: Sheryl Ball

Address: 3016 Pamplin Hall, Mail Code 0316, 880 West Campus Drive, Blacksburg, VA 24061

Phone: (540) 231-4349

Fax: 540-231-9288

Email: sball@vt.edu

2. Abstract and key terms

The determinants of vaccine hesitancy remain complex and context specific. Betrayal aversion occurs when an individual is hesitant to risk being betrayed in an environment involving trust. In this pre-registered vignette experiment, we show that betrayal aversion is not captured by current vaccine hesitancy measures despite representing a significant source of unwillingness to be vaccinated. Our survey instrument was administered to 888 United States residents via Amazon Mechanical Turk in March 2021. We find that over a third of participants have betrayal averse preferences, resulting in an 8-26% decline in vaccine acceptance, depending on the betrayal source. Interestingly, attributing betrayal risk to scientists or government results in the greatest declines in vaccine acceptance. We explore an exogenous message intervention and show that an otherwise effective message acts narrowly and fails to reduce betrayal aversion. Our results demonstrate the importance of betrayal aversion as a preference construct in the decision to vaccinate.

Key terms: Health Behavior, Persuasive Messages, COVID-19

3. Introduction

While vaccinations are important in protecting public health, many individuals are vaccine hesitant, delaying or refusing available vaccines. The 3C model [16,18] measures the prevalence of three factors of vaccine hesitancy: confidence (trust in safety and effectiveness), convenience (physical and psychological constraints,) and complacency (perceived risks of disease are low). Recently work suggests that individual differences in information processing and social preferences also influence hesitancy [4]. Nonetheless, the determinants of vaccine hesitancy remain complex and context specific [18].

Betrayal aversion is a decision-making preference that occurs when people avoid situations involving trust in order to avoid disutility associated with the negative emotions experienced when being betrayed [1,2,5,12]. A bad trust outcome carries with it both lost utility relative to the desired outcome and disutility from being betrayed, with betrayal aversion capturing aversion to negative emotional disutility related to the latter. Importantly, betrayal aversion is distinct from a preference for avoiding outcome risk in trust environments. When choosing between safety products like vaccines, betrayal averse individuals may accept lower levels of protection from the primary risk in order to avoid a relatively small secondary risk of being harmed by the safety product itself. People were significantly more likely to choose to purchase airbags and smoke alarms, and marginally more likely to choose vaccines when the potential risks of the products involved no betrayal [12]. Levels of betrayal aversion also decline in contexts that dampen emotional responses [9].

Current measures of vaccine hesitancy capture overall beliefs about the safety of vaccines without disassociating the source of the assumed risks [15,23,24], hence, are unable to determine whether betrayal aversion represents an additional, currently unmeasured, source of hesitancy. In

this work, we seek to establish the importance of betrayal aversion as a preference construct in the decision to vaccinate. We developed a pre-registered vignette experiment involving a highly infectious hypothetical disease to measure willingness to become vaccinated across different betrayal scenarios while holding the overall risk level constant. We hypothesized that willingness to become vaccinated is lower when there is an additional risk of betrayal (death due to side effects) compared to a non-betrayal risk. We found evidence confirming this hypothesis.

During pandemics, information about the disease and available vaccines develops rapidly. The nature of communication and spread of information can have significant influences on vaccine hesitancy [25,20,21,17,26,6,19]. Learning that President Obama had his daughters vaccinated, for example, was positively associated with willingness to get vaccinated against H1N1 influenza [25]. Recent studies on COVID-19 vaccination approval reported a positive impact of messaging stressing the importance of herd immunity [28] as well as endorsements by the director of the National Institute of Allergy and Infectious Diseases [6]. On the other hand, misinformation around COVID-19 caused a decline in vaccination intent [17]. In this study, we investigate whether a message that primes regret for not getting the vaccine increases willingness to vaccinate and, separately, whether it decreases betrayal aversion. Willingness to become vaccinated was greater among responders exposed to our messaging intervention. Contrary to our pre-registered hypothesis, the prevalence of betrayal aversion did not decline after message exposure. To identify the channels by which the messaging treatment operates on vaccination intent, we administered another pre-registered survey to another group of responders that reported the motives for getting the vaccine, and found exposure to the message treatment increased the intent to protect others.

The COVID-19 pandemic has served as an illustration that many potential factors connect betrayal aversion and vaccine hesitancy. To explore some of these factors, in addition to our two

pre-registered studies, we designed an exploratory study (not pre-registered) to investigate betrayal aversion and vaccine hesitancy when the risk of betrayal is brought about by 1) a chance of the vaccine weakening the immune system, 2) a partisan government approving the vaccine rapidly, overlooking certain safety measures, or 3) scientists approving the vaccine rapidly, overlooking certain safety measures. Betrayal aversion was higher when a partisan government or scientists actively contribute to the associated risk. However, betrayal sensitivity was dampened when the political party of the government developing the vaccine was aligned with an individual's own preferences. This result demonstrates that political polarization can yield and/or amplify differences in vaccination intent [6,11,13,29].

4. Materials and Methods

We administered an online survey via Amazon Mechanical Turk (MTurk). The final sample includes 595 participants for our pre-registered hypotheses (https://osf.io/4peuy) (from registered studies 1 and 2) and 293 additional participants to test our exploratory hypotheses (study 3) (see Electronic Supplementary Material: Appendix A for MTurk exclusion criteria and Appendix B for sample characteristics in Table A1). This study was approved by the Institutional Review Board of a large public university in the United States.

Participants provided informed consent and received $2 compensation for the task, which took an average of 19 minutes to complete. After answering questions about our hypothetical disease scenario (described below and in Table 1) participants reported their real-world experiences with vaccines, including those related to COVID-19. In particular, the first group of surveys asked about participants' general attitudes toward vaccination (5 questions) and toward COVID-19 (3 questions). Other study measures included a survey of demographic characteristics and two widely

used vaccine hesitancy surveys adopted from 1) the Parental Attitudes About Childhood Vaccines questionnaire [23,24] (the 5-question PACV-short version [22,24]) and the Vaccine Confidence Index questionnaire (the 4-question VCI core survey [14,15], see Appendix A for indices' scoring). Note that data was collected in the United States during the COVID-19 pandemic between March 30th, 2021 and April 8th, 2021. In the United States, by March 30th, the cumulative number of reported COVID-19 cases surpassed 29 million and the number of the associated deaths was 543,003 [31]. By this time more than 130 million COVID-19 vaccines had been administered [30]. The vignette conditions and subsequent questionnaire are available in appendix section A3.

4.1. Study 1: betrayal aversion related vaccine hesitancy

In the vignette experiment, participants read a hypothetical scenario about a novel future disease described as both highly infectious and deadly (Table 1), with an available free and easy-to-take vaccine. To investigate betrayal aversion, we modify Koehler and Gershoff's (2003) safety product task (which did not explore general vaccine hesitancy and found marginal results on vaccine betrayal) to create a within-subject design with greater power to detect betrayal aversion to vaccines [9,12]. Participants were told that the unvaccinated face a 2% risk of death. Those who are vaccinated face a 1.01% risk of death, where 1% was due to the virus, while the explanation for the remaining 0.01% varies across treatments. By measuring changes in willingness to become vaccinated across treatments while holding the probability of death constant, we explore how the causes of betrayal aversion affect its severity.

[Table 1 here]

In our benchmark *Non-Betrayal* treatment, the 0.01% chance of death faced by those who get the vaccine was described as "problems unrelated to the vaccine." The cause of the 0.01% chance of death in our *Side-Effects* treatments involved active betrayal, and was described as

"vaccine-induced complications (side effects)." We asked participants (N=395) to indicate their willingness to become vaccinated (single 7-point Likert question with response alternatives ranging from "Definitely reject" at 1 to "Definitely accept" at 7) under both scenarios, with the order of scenarios presented counterbalanced across participants. By subtracting an individual's willingness to become vaccinated in the *Side-Effects* treatments from their willingness in the *Non-Betrayal* treatment we get an individual level measurement of betrayal aversion.

4.2. Study 2: Messaging, vaccine hesitancy and betrayal aversion

Our second study explores the influence of messaging on willingness to become vaccinated with and without the possibility of betrayal using Dan's Ariely's message suggestion (see Appendix A), which primes feelings of regret for not getting the vaccine [8]. Here is the exact wording of the message:

"*There is no chance you will regret getting the vaccine, but, if you don't get it then you may either get sick and might die or may get other people sick meaning that they might die, and you could regret it. Imagine how you would feel if you passed the virus to someone else. Just try to imagine how that would feel. Now tell us that you should not do a lot to prevent that terrible feeling of regret that you didn't get the vaccination earlier."*

The *Risk-Only* treatment (N=103) involves the same probabilities as above, 1.01% with and 2% chance of death without the vaccine, but no chance of betrayal. In the *Risk-Only w/ Message* treatment (N=97), we presented the message prior to participants making decisions in the *Risk-Only* treatment. By comparing the responses to five vaccine uptake questions in both treatments [17], the causal role of the regret message on vaccine hesitancy is identified while also investigating the motives behind vaccination intent and hesitancy. Participants reported their benchmark willingness to become vaccinated, their willingness to become vaccinated in order to protect

oneself, and separately their willingness to become vaccinated to protect others (family, friends, and at-risk groups). Responders also indicated their willingness to wait and see how the vaccine is working before getting vaccinated and their willingness to become vaccinated only if it was required by their work or school (7-point Likert questions).

4.3. Study 3: Source of betrayal and vaccine hesitancy

The last study focuses on our exploratory treatments, where we manipulated the source of betrayal (Table 1). In all of the three conditions with varied sources of betrayal, participants also made decisions in a *Non-Betrayal* condition as well (order counter-balanced across subjects). In the *Counter-Productivity* treatment (N=97), we elicited willingness to become vaccinated in an active betrayal condition where the 0.01% risk was caused by the vaccine lowering the recipient's general immunity making them more prone to catching the virus. Concerns about the vaccine overloading the immune system has been one of the most stated reasons behind vaccine hesitancy [27]. In the *Government* treatment (N=102), we elicited willingness to become vaccinated when the 0.01% risk of active betrayal is caused by having either a Democrat or a Republican government overseeing an accelerated vaccine approval process that overlooks potential safety concerns. Note that in this condition, participants were presented with three different conditions (order counter-balanced, see Appendix A): *Non-Betrayal*, Democrat Government, and Republican Government. The *Scientists* treatment (N=94) investigate betrayal aversion when the active betrayal is attributed to the behavior of scientists (working for pharmaceutical companies or for the government) (see Appendix A for additional details).

5. Results

5.1. Study 1 results: betrayal aversion related vaccine hesitancy

We first compare willingness to become vaccinated when the 0.01% additional risk was due to vaccine-induced complications (*Side-Effects* treatments) instead of unrelated problems (*Non-Betrayal* treatment). Participants reported a significantly lower willingness to vaccinate when the risk associated with vaccination was due to side effects ($M_{difference}$= 0.453; Wilcoxon signed-rank test, P<0.001) (Fig. 1A-B). Next, we compute a binary measure of vaccine hesitancy by collapsing the Likert responses for each treatment into two categories: vaccine hesitant (score< 5) or vaccine non-hesitant (score≥5). When the additional risk was caused by side-effects, 8.4% more people were identified as vaccine hesitant (McNemar test, P<0.001).

[Figure 1 here]

To derive an individual-level measure of betrayal aversion, we calculate the difference between willingness to become vaccinated in the *Non-Betrayal* and *Side-Effects* treatments. As shown in Fig. 1C, the distribution of betrayal aversion is skewed to the right with more than 30% of responders classified as betrayal averse (positive level of betrayal aversion). Taken together these results support our pre-registered hypothesis and indicates that betrayal aversion is likely influencing the willingness of individuals to vaccinate., This result is consistent with previous studies [12,5,1,2], demonstrating a substantial portion of the population is betrayal averse.

5.2. Study 2 results: Messaging, vaccine hesitancy and betrayal aversion

Similar to other studies that explored vaccine hesitancy and communication [25,20,21,17,19], participants exposed to our message treatment in the *Risk Only w/ Message* condition reported higher willingness to become vaccinated relative to the group that did not receive the message ($M_{difference}$= 0.668; Two-sample Wilcoxon rank-sum test, P=0.009) (Fig. 2A). The channel

through which the message appears to work is social – our regret messaging increased willingness to become vaccinated to protect others ($M_{difference}$ = 0.522; Two-sample Wilcoxon rank-sum test, P=0.022; Fig. 2A) rather than protecting oneself. After collapsing the Likert responses into a binary measure of vaccine hesitancy as before, exposure to the message reduced vaccine hesitancy by about 11.7% (One-sided Fisher's exact test, P=0.032). Confirming the social channel of the message, the likelihood of accepting the vaccine to protect others was 14.8% higher among participants exposed to the message compared to the control group (One-sided Fisher's exact test, P=0.008).

[Figure 2 here]

With evidence that messaging is effective at reducing overall vaccine hesitancy, we now explore whether the message also decrease betrayal aversion. This is measured by contrasting decisions in the *Side-Effects* treatment (N=198) to that from the *Side-Effects w/ Message* treatment (N=197) that involves displaying the treatment message prior to presenting participants with the *Non-Betrayal* and *Side-Effects* treatment frameworks. As shown in Fig. 2B (right panel), betrayal aversion, on average, was not statistically significantly lower for participants exposed to the message treatment ($M_{difference}$ = 0.084; Two-sample Wilcoxon rank-sum test; $P = 0.712$). Failure of the message to decrease betrayal aversion could arise either from similar changes of vaccine hesitancy in both treatments or from a lack of change of vaccine hesitancy in both the *Non-Betrayal* and *Side-Effects* treatments due to the message. Fig. 2B (left panel) shows that relative to the no message comparison treatment groups, participants exposed to the message treatment reported higher willingness to vaccinate under both *Non-Betrayal* ($M_{difference}$ = 0.351; Two-sample Wilcoxon rank-sum test, P= 0.028) and *Side-Effects* conditions ($M_{difference}$ = 0.435; Two-sample Wilcoxon rank-sum test, P=0.034). These results suggest that while regret messaging

is effective at reducing vaccine hesitancy, it is acting orthogonally to betrayal aversion related vaccine hesitancy. Consequently, this result seems to be consistent with the evidence from the *Risk-Only* and *Risk-Only w/ Message* treatments that suggested the messages were acting to increase willingness to become vaccinated to protect others. In particular, the effect of the message operates through an external/other focused channel while betrayal aversion would be expected to be an emotional/internal focused channel [9].

Our results exploring messaging's effects on vaccine hesitancy using the *Risk-Only* and *Risk-Only w/ Message* treatments are robust to exploring the data with a multiple linear regression analysis with vaccine acceptance as the dependent variable while controlling for individual characteristics (see Appendix A), measures of vaccine hesitancy and geographic region in fixed effects linear regression models (Effect on benchmark acceptance: $\beta_{message}$>0.661, P<0.050; Effect on acceptance to protect others: $\beta_{message}$>0.279, P<0.050; Appendix B Tables A2-A3). For these regression analyses, and subsequent tables in the Appendix B, we standardize all 7-Likert scale measures at the responder level, prior to calculating betrayal aversion, and use the z-scores in model estimation. Importantly, when betrayal aversion is modelled instead, the null effect of the message treatment on betrayal aversion persisted ($\beta_{message} = -0.059$ in (a), $\beta_{message} = -0.063$ in (b); $P > 0.100$) (Appendix B Table A4). Note that the coefficients for the commonly used vaccine hesitancy measures (PACV and VCI) are not significantly associated with betrayal aversion ($\beta_{PACV} = -0.023, P > 0.100$; $\beta_{\mathrm{VIC}} = -0.154, P > 0.100$) and seem to move in the opposite direction. We obtain an identical result when we include each of the four measures in the VCI separately in the regression (Larson et al., 2016) (available upon request). Moreover, the finding that the message increased willingness to become vaccinated in both treatments (*Non-Betrayal* and *Side-Effects*) holds in our regression analysis ($\beta_{message}$>0.251, P<0.050; Appendix

B Table A5). Even though the null result of the effect of the message on betrayal aversion related vaccine hesitancy is in conflict with our pre-registered hypothesis that vaccine hesitancy reducing messages would mitigate betrayal aversion, our findings suggest that the current measures of vaccine hesitancy, and potentially messaging interventions, seem to fall short in accounting for an important preference construct that influences the vaccination decision.

5.3. Study 3 results: Source of betrayal and vaccine hesitancy

We compared whether the level of betrayal aversion differs when active betrayal comes from side effects (*Side-Effects* treatment) or when due to the vaccine causing an individual to suffer weakened immunity that makes them more prone to catching the virus (*Counter-Productivity* treatment.) There was no discernable difference in betrayal aversion between these two types of active betrayals ($M_{difference}$= 0.135; Two-sample Wilcoxon rank-sum test, $P = 0.715$; Fig. 3A).

[Figure 3 here]

On the other hand, betrayal aversion was higher in conditions involving either politicians or scientists actively contributing to the betrayal risk. For the *Government* (*Scientists*) treatment, the lowest willingness to become vaccinated between Democrat or Republican government (Pharmaceutical or Government scientists) was used to derive betrayal aversion. Relative to the active betrayal treatment (*Counter-Productivity)*, betrayal aversion to vaccination was significantly higher when betrayal was channeled through the government ($M_{difference}$= 0.481; Two-sample Wilcoxon rank-sum test, $P = 0.020$) or scientists ($M_{difference}$= 0.583; Two-sample Wilcoxon rank-sum test, $P = 0.006$; Fig. 3A). We confirm that betrayal aversion is greater in the government or scientists treatments relative to the *Side-Effects* treatment instead (*Government* $M_{difference}$ = 0.615; *Scientists* $M_{difference}$ = 0.717; Two-sample Wilcoxon rank-sum tests, $P < 0.001$). In addition, we confirm these results using regression analysis that controls for individual

characteristics, measures of vaccine hesitancy and region fixed effects ($\beta_{Government} > 0.346, P < 0.010$; $\beta_{\text{Scientists}} > 0.388$, $P < 0.010$; Appendix B Table A6).

When the additional risk of betrayal was caused by *Counter-Productivity*, vaccine acceptance was reduced by about 12.4% (McNemar test, P=0.008). This decline is again comparable to the 8.4% decline in vaccine acceptance reported in *Study 1* for betrayal caused by *Side-Effects*. On the other hand, vaccine acceptance was reduced more steeply by 20.6% and 25.6% when the additional risk of betrayal was caused by the government or the scientists (McNemar tests, P<0.001), respectively.

Participants reported lower willingness to vaccinate when the betrayal source was a Republican led government compared to a Democrat led government ($M_{difference}$ = 0.275; $P = 0.041$; Wilcoxon signed-rank test, $P = 0.017$). Interestingly, betrayal aversion to a vaccine developed by a Democrat or Republican led government is modulated by political orientation. Here, responders who self-identified as Democrats were more betrayal averse to a vaccine from a Republican led government compared to a Democrat led government ($M_{difference}$ = 0.771; Wilcoxon signed-rank test, $P < 0.001$) (Fig. 3B). Analogously, Republicans were more sensitive to betrayal when the vaccine risk was brought about by a Democrat led government instead of a Republican led government ($M_{difference}$ = 0.720; Wilcoxon signed-rank test, $P = 0.014$). Willingness to become vaccinated for self-identified Republicans in the *Non-Betrayal* scenario was not significantly different from that in the scenario involving the Republican government as a proxy cause of betrayal ($M_{difference}$ = 0.120; Wilcoxon signed-rank test, $P = 0.420$). On the other hand, betrayal aversion persisted for self-identified Democrats who reported higher willingness to become vaccinated in the *Non-Betrayal* scenario relative to that involving the Democrat government as a proxy cause of betrayal ($M_{difference}$ = 0.604; Wilcoxon signed-rank

test, $P = 0.009$). Responders who self-identified as neither Republicans nor Democrats were marginally more betrayal averse when the source of betrayal involved a Republican led government instead of a Democrat led government ($\mu_{difference}$ = 0.310; Wilcoxon signed-rank test: $P = 0.056$).

Willingness to vaccinate did not differ significantly when the betrayal source involved either pharmaceutical company scientists or government employed scientists ($M_{difference}$ = 0.043; Wilcoxon signed-rank test, $P = 0.461$). Moreover, betrayal aversion to vaccines developed rapidly by pharmaceutical company or government employed scientists were not differentiated by political affiliation ($M_{difference}$ < 0.104; Wilcoxon signed-rank tests, $P > 0.206$). Interestingly, however, we find that Republican responders, relative to Democrat responders, were significantly more betrayal averse to vaccines from government employed scientists ($M_{difference}$ = 0.871; Two-sample Wilcoxon rank-sum test, $P = 0.015$) and marginally more averse to betrayal from vaccines developed by pharmaceutical company employed scientists ($M_{difference}$ = 0.819; Two-sample Wilcoxon rank-sum test, $P = 0.058$).

6. Discussion

Betrayal aversion is an important preference that influences decision-making in situations where trust can be broken. Despite their potential to reduce the overall risk of harm, safety products with small chances of causing the very same harm they are expected to prevent are often less preferred [2,9,12]. In this study, we demonstrate that betrayal aversion is an important preference construct in the decision to vaccinate and is one not accounted for by widely used vaccine hesitancy measures.

In addition, the observed level of betrayal aversion depended on the source of betrayal. In comparison to the levels observed when the source of betrayal is a vaccine side effect, betrayal

aversion was amplified when the government or scientists may be at fault. This finding may be due to stronger emotional responses when potential betrayals involve institutions or personnel that were expected to prevent harm. Thus, the active involvement of the government or scientists in breaking the trust may have compounded betrayal aversion to the safety product itself. Interestingly, Democrats (Republicans) were more sensitive to betrayal by a Republican (Democrat) government. Thus, betrayal aversion seems to further amplify political polarization in vaccination decisions [3,6,11,13,29].

We explore a messaging intervention that increased overall willingness to vaccinate but did not ameliorate betrayal aversion. The messaging intervention primes feelings of regret for not getting the vaccine via phrases like "*Imagine how you would feel if you passed the virus to someone else."* The message operated through other regarding preferences, such as altruism, enhancing receivers' willingness to vaccinate to protect friends, family members and at-risk groups. Since the message was not targeted at reducing concerns about betrayal, we were not surprised that it failed to decrease hesitancy through that channel. In fact, in a sense the message's ineffectiveness in reducing betrayal aversion underscores the importance of treating betrayal aversion as a unique preference construct so that it is not overlooked in health communications and behavioral interventions.

Vignette experiments are a valuable first step in a multi-method research agenda where many observations are needed and circumstances preclude other data collection strategies (e.g. ethical issues are present) [7]. At the same time, we acknowledge that research using survey methods may have reduced external validity when compared with some other approaches because participants do not experience the scenario and decisions are not consequential. Nevertheless, an external validity check revealed a positive and significant correlation between willingness to get COVID-

19 vaccination and willingness to become vaccinated in the vignette experiment ($\beta_{Willingness\ to\ vaccinate\ in\ vignette} > 0.299, P < 0.050$; Appendix B Table A7).

The results of this study motivate future research on vaccine betrayal aversion. First, betrayal aversion might be added to measures of vaccine hesitancy to augment their accuracy. Next, research on interventions to increase vaccination rates should address the betrayal aversion channel of vaccine hesitancy, which we show is less malleable to some messaging. Emotion regulation interventions [10], however, may be a potential candidate to mitigate betrayal aversion to vaccination. Finally, we believe that, with additional research and perhaps including a field or randomized controlled study, results from this study may shed light on how to reduce COVID-19-related vaccine hesitancy.

7. Acknowledgments

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Funding provided by the National Science Foundation (grant #1541105).

Informed consent

This study was approved by the Institutional Review Board of a large public university in the United States and participants provided informed consent

Author Contributions

All authors contributed to conception and design of the study. First and second authors collected the data and organized the database. First author performed the statistical analysis and wrote the first draft of the manuscript. All authors wrote sections of the manuscript, contributed to manuscript revision, and read and approved the submitted version.

Data and Code Availability Statement

Data reported in this study and code generating the results will be made available via the Open Science Framework upon publication: https://osf.io/4peuy.

Contributions that do not justify authorship

Not applicable

Technical help

Not applicable

8. References

1. Aimone, J. A., and D. Houser. What you don't know won't hurt you: a laboratory analysis of betrayal aversion. *Experimental Economics* 15:571–588, 2012.
2. Aimone, J., S. Ball, and B. King-Casas. The betrayal aversion elicitation task: an individual level betrayal aversion measure. *PloS one* 10:e0137491, 2015.
3. Betsch, C., L. Korn, and R. Böhm. Reply to Weisel: From polarization to vaccination and back. *Proceedings of the National Academy of Sciences* 118:, 2021.
4. Betsch, C., P. Schmid, D. Heinemeier, L. Korn, C. Holtmann, and R. Böhm. Beyond confidence: Development of a measure assessing the 5C psychological antecedents of vaccination. *PloS one* 13:e0208601, 2018.
5. Bohnet, I., F. Greig, B. Herrmann, and R. Zeckhauser. Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *American Economic Review* 98:294–310, 2008.
6. Bokemper, S. E., G. A. Huber, A. S. Gerber, E. K. James, and S. B. Omer. Timing of COVID-19 vaccine approval and endorsement by public figures. *Vaccine* 39:825–829, 2021.
7. Erfanian, F., R. Latifnejad Roudsari, A. Heydari, and M. Noghani Dokht Bahmani. A Narrative on the Use of Vignette: Its Advantages and Drawbacks. *Journal of Midwifery and Reproductive Health* 8:2134–2145, 2020.
8. Ferreri, E. To Convince Vaccine Skeptics, Use Empathy, Information and a Re-Start, Experts Say. , 2020.at <https://today.duke.edu/2020/12/convince-vaccine-skeptics-use-empathy-information-and-re-start-experts-say>
9. Gershoff, A. D., and J. J. Koehler. Safety first? The role of emotion in safety product betrayal aversion. *Journal of Consumer Research* 38:140–150, 2011.

10. Gross, J. J. Emotion regulation: Current status and future prospects. *Psychological inquiry* 26:1–26, 2015.
11. Hamel, L., A. Kirzinger, C. Muñana, and M. Brodie. KFF COVID-19 Vaccine Monitor: December 2020. 2020.at <https://www.kff.org/coronavirus-covid-19/report/kff-covid-19-vaccine-monitor-december-2020/>
12. Koehler, J. J., and A. D. Gershoff. Betrayal aversion: When agents of protection become agents of harm. *Organizational Behavior and Human Decision Processes* 90:244–261, 2003.
13. Kreps, S., S. Prasad, J. S. Brownstein, Y. Hswen, B. T. Garibaldi, B. Zhang, and D. L. Kriner. Factors associated with US adults' likelihood of accepting COVID-19 vaccination. *JAMA network open* 3:e2025594–e2025594, 2020.
14. Larson, H., A. de Figueiredo, E. Karafillakis, and M. Rawal. State of vaccine confidence in the EU 2018. *Luxembourg: Publications Office of the European Union* 10:241099, 2018.
15. Larson, H. J., A. De Figueiredo, Z. Xiahong, W. S. Schulz, P. Verger, I. G. Johnston, A. R. Cook, and N. S. Jones. The state of vaccine confidence 2016: global insights through a 67-country survey. *EBioMedicine* 12:295–301, 2016.
16. Larson, H. J., C. Jarrett, E. Eckersberger, D. M. Smith, and P. Paterson. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012. *Vaccine* 32:2150–2159, 2014.
17. Loomba, S., A. de Figueiredo, S. J. Piatek, K. de Graaf, and H. J. Larson. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature human behaviour* 5:337–348, 2021.
18. MacDonald, N. E. Vaccine hesitancy: Definition, scope and determinants. *Vaccine* 33:4161–4164, 2015.

19. Milkman, K. L. *et al.* A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. *Proc Natl Acad Sci USA* 118:e2101165118, 2021.
20. Nyhan, B., J. Reifler, S. Richey, and G. L. Freed. Effective messages in vaccine promotion: a randomized trial. *Pediatrics* 133:e835–e842, 2014.
21. Odone, A., A. Ferrari, F. Spagnoli, S. Visciarelli, A. Shefer, C. Pasquarella, and C. Signorelli. Effectiveness of interventions that apply new media to improve vaccine uptake and vaccine coverage: a systematic review. *Human vaccines & immunotherapeutics* 11:72–82, 2015.
22. Oladejo, O., K. Allen, A. Amin, P. M. Frew, R. A. Bednarczyk, and S. B. Omer. Comparative analysis of the Parent Attitudes about Childhood Vaccines (PACV) short scale and the five categories of vaccine acceptance identified by Gust et al. *Vaccine* 34:4964–4968, 2016.
23. Opel, D. J., J. A. Taylor, R. Mangione-Smith, C. Solomon, C. Zhao, S. Catz, and D. Martin. Validity and reliability of a survey to identify vaccine-hesitant parents. *Vaccine* 29:6598–6605, 2011.
24. Opel, D. J., J. A. Taylor, C. Zhou, S. Catz, M. Myaing, and R. Mangione-Smith. The relationship between parent attitudes about childhood vaccines survey scores and future child immunization status: a validation study. *JAMA pediatrics* 167:1065–1071, 2013.
25. Quinn, S. C., J. Parmer, V. S. Freimuth, K. M. Hilyard, D. Musa, and K. H. Kim. Exploring communication, trust in government, and vaccination intention later in the 2009 H1N1 pandemic: results of a national survey. *Biosecurity and bioterrorism: biodefense strategy, practice, and science* 11:96–106, 2013.

26. Romaniuc, R., A. Guido, N. Mai, E. Spiegelman, and A. Sutan. Increasing Vaccine Acceptance and Uptake: A Review of the Evidence. *Available at SSRN 3839654* , 2021.
27. Salmon, D. A., L. H. Moulton, S. B. Omer, M. P. DeHart, S. Stokley, and N. A. Halsey. Factors associated with refusal of childhood vaccines among parents of school-aged children: a case-control study. *Archives of pediatrics & adolescent medicine* 159:470–476, 2005.
28. Trueblood, J. S., A. B. Sussman, and D. O'Leary. The Role of General Risk Preferences in Messaging About COVID-19 Vaccine Take-Up. *Available at SSRN 3649654* , 2020.
29. Weisel, O. Vaccination as a social contract: The case of COVID-19 and US political partisanship. *Proceedings of the National Academy of Sciences* 118:, 2021.
30. COVID Data Tracker Weekly Review for March 26, 2021. Centers for Disease Control and Prevention: 2021.at <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/past-reports/03262021.html>
31. COVID-19 Weekly Epidemiological Update, March 30 2021. World Health Organization: 2021.at <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---31-march-2021>

9. Tables, figure legends

Table 1: Caption: List of treatment conditions.

<table>
<tr><th colspan="3">Treatments
(Explanations for the 1.01% chance of Death after Vaccine)</th></tr>
<tr><th>Treatment</th><th>Undivided or Divided Probability</th><th>Observations</th></tr>
<tr><td>Risk-Only</td><td rowspan="2">Undivided probability
"a 1.01% chance that people treated with the vaccine will contract the virus and die as a result."</td><td>103</td></tr>
<tr><td>Risk-Only w/ Message</td><td>97</td></tr>
<tr><td></td><td>Divided Probability
“a 1% chance that people treated with the vaccine will contract the virus and die as a result.”
and "an additional one chance in 10,000 (0.01%) that someone who is treated with the vaccine will die due to…</td><td></td></tr>
<tr><td>Non-Betrayal (Benchmark)</td><td>... problems unrelated to the vaccine."</td><td>This treatment is run counterbalanced with all other treatments below</td></tr>
<tr><td>Side-Effects</td><td>…vaccine-induced complications (side effects)."</td><td>198</td></tr>
<tr><td>Side-Effects w/ Message</td><td>…vaccine-induced complications (side effects)."</td><td>197</td></tr>
</table>

		(The message was also displayed in the Non-Betrayal benchmark)
Counter-Productivity	*…the vaccine lowering the recipient's immunity making them more prone to catching the virus.*"	97
Government	*…a [XYZ} government in charge that approved the vaccine too rapidly, overlooking certain safety concerns*."	102 ("XYZ" read "Democrat" or "Republican", order counterbalanced)
Scientists	*…scientists who work for [XYZ} developing the vaccine too rapidly, overlooking certain safety concerns*."	94 ("XYZ" read "the government" or "pharmaceutical companies", order counterbalanced)

Figure 1: Willingness to vaccinate and betrayal aversion. (A) Distribution of willingness to vaccinate responses across Non-Betrayal and Side-Effects treatments. (B) Average willingness to vaccinate across betrayal treatments. (C) Betrayal aversion distribution (Difference in willingness to vaccinate across treatments). $N_{Side-Effects_{pooled}}$=395 (Side-Effects + Side-Effects w/ Message) Errors bars denote 95% confidence intervals. Two-sided paired t-tests: *** P<.01, ** P<.05, * P<.1, N.S. Not significant.

Figure 2: Message exposure, willingness to vaccinate, and betrayal aversion. (A) Willingness to get the vaccine, and associated motives with undivided probabilities across message and no message treatments (see Table 1; $N_{\text{Risk-Only}}$=103; $N_{\text{Risk-Only w/ message}}$ =97). (B) Willingness to get the vaccine with divided probabilities across message treatments (see Table 1; $N_{\text{Side-Effects}}$=198; $N_{\text{Side-Effects w/ message}}$=197) and the corresponding average betrayal aversion across message treatments. Errors bars denote 95% confidence intervals. Two-sided t-tests: *** p<.01, ** p<.05, * p<.1, N.S. Non-significant.

Figure 3: Betrayal aversion across betrayal sources. (A) Betrayal aversion across betrayal source conditions ($N_{Side-Effects(pooled)}$=395; $N_{Counter-Productivitiy}$=97; $N_{Government}$=102; $N_{Scientists}$=94). For Government (Scientists), the lowest willingness to get the vaccine between Democrat or Republican governments (Pharmaceutical or Government scientists) is used to calculate betrayal aversion. (B) Betrayal aversion by political orientation for the two betrayal treatments: Government (N_{Other}=29; $N_{Republican}$=25; $N_{Democrat}$=48) and Scientists (N_{Other}=27; $N_{Republican}$=19; $N_{Democrat}$=48). Errors bars denote 95% confidence intervals. Two-sided t-tests (paired tests in (B) only): *** p<.01, ** p<.05, * p<.1, N.S. Non-significant.