



# No association between numerical ability and politically motivated reasoning in a large US probability sample

Michael N. Stagnaro<sup>a,1</sup> , Ben M. Tappin<sup>a</sup>, and David G. Rand<sup>a</sup> 

Edited by Kenneth Wachter, University of California Berkeley, Berkeley, CA; received January 27, 2023; accepted June 14, 2023

The highly influential theory of “Motivated System 2 Reasoning” argues that analytical, deliberative (“System 2”) reasoning is hijacked by identity when considering ideologically charged issues—leading people who are more likely to engage in such reasoning to be more polarized, rather than more accurate. Here, we fail to replicate the key empirical support for this theory across five contentious issues, using a large gold-standard nationally representative probability sample of Americans. While participants were more accurate in evaluating a contingency table when the outcome aligned with their politics (even when controlling for prior beliefs), we find that participants with higher numeracy were more accurate in evaluating the contingency table, regardless of whether or not the table’s outcome aligned with their politics. These findings call for a reconsideration of the effect of identity on analytical reasoning.

motivated reasoning | polarization | decision-making | political psychology

How people reason about ideologically charged issues is hotly contested. The standard assumption that reasoning leads to more accurate beliefs has been challenged by the idea of (directionally) *motivated* reasoning (1). The general theory of motivated reasoning argues that, when an individual has a personal stake in reaching a particular conclusion, reasoning is clouded by bias and rendered less able to effectively use available evidence to form accurate beliefs. Instead, reasoning works as “a lawyer,” producing arguments and evaluating evidence to support the reasoner’s desired conclusion.

In recent years, an extension of the motivated reasoning thesis—known as “Motivated System 2 Reasoning” (MS2R; also called “Motivated Numeracy”) (2)—has become highly influential in academic and lay perceptions of motivated reasoning. This extended theory adopts a dual-process perspective on decision-making, in which individuals are assumed to vary in their reliance on intuitive, automatic (“System 1”) decision-making processes versus analytical, deliberative (“System 2”) reasoning. MS2R then argues that individuals who engage in more reasoning should be better at producing arguments and evidence to support their desired conclusion (and rejecting evidence that challenges their desired conclusion). As a result, on issues where there is strong ideological or cultural conflict, individuals who are most analytic are expected to be the most biased or polarized in their beliefs, as they are most effective at deploying their reasoning abilities to support the positions favored by their “side.”

The finding often described as providing the strongest empirical evidence for MS2R comes from a now-classic paradigm introduced by Kahan and colleagues (2, 3). Participants were provided with a description of a scientific study that examined either a politically charged topic (gun control) or a neutral topic (skin cream). Participants were then shown the results of this “study” in the form of a two-by-two contingency table and asked to interpret the table (i.e., to indicate what association was demonstrated in the table). The contingency table was constructed in such a way that participants needed to invest cognitive effort in numerical reasoning to reach the correct answer, and which answer was correct was randomized across participants.

When the correct answer was aligned with the participant’s identity commitments (or was nonpolitical), participants who scored higher on a numeracy test were more likely to interpret the contingency table correctly. Critically, however, when the correct answer ran counter to the participant’s identity commitments, the advantage conferred by numeracy disappeared: Those who scored higher on the numeracy test were no more likely to interpret the contingency table correctly than those who scored lower. As a result, political polarization in the interpretation of the data was highest among the most numerate individuals.

This surprising result has attracted a great deal of attention. It has been widely interpreted as evidence that people with superior reasoning abilities use them to defend their identities in the context of ideologically divisive topics—and, thus, that inaccurate beliefs about such topics are better explained by motivated reasoning than insufficient reasoning

Author affiliations: <sup>a</sup>Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142

Author contributions: M.N.S. and D.G.R. designed research; M.N.S. performed research; M.N.S. analyzed data; and M.N.S., B.M.T., and D.G.R. wrote the paper.

The authors declare no competing interest.

Copyright © 2023 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

<sup>1</sup>To whom correspondence may be addressed. Email: [mntag@mit.edu](mailto:mntag@mit.edu).

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2301491120/-/DCSupplemental>.

Published July 31, 2023.

(e.g., in the context of susceptibility to misinformation). Notably, however, this result could also be explained by differences in individuals' prior beliefs, where those who score higher on reasoning tasks are more deferential to their priors on a given topic rather than more deferential to their identities (4).

Beyond this conceptual issue, the replicability and generalizability of the key finding itself are unclear. While one independent team replicated the finding in an Australian sample (5), numerous others have failed to replicate it (6–11), and still others have found mixed results (12–14). Thus, the robustness of this central piece of evidence for MS2R is unclear.

The authors of the original study have argued that obtaining a large sample that accurately represents the distribution of “typical” US citizens is essential for identifying MS2R and have suggested that replication failures may be due to the use of small convenience samples rather than the probability sample of American participants used in the original work (3, 15).

Here, we assess this possibility by providing the first independent preregistered test of the *replicability* of the original finding using a gold-standard representative probability sample of Americans. Importantly, we conferred with one of the original authors from refs. 2 and 3 and confirmed that the AmeriSpeak panel provides sufficiently representative samples for this replication to satisfy earlier critiques regarding representativeness. Our sample is also twice as large as the original study, providing additional statistical power.

Furthermore, we assess the generalizability of the original finding by examining gun control as well as four additional topics that pre-testing indicates are highly polarizing for both Democrats and Republicans. Finally, we investigate the alternative explanation for the pattern observed in the original study, which involves numeracy being associated with greater deference to prior factual beliefs rather than identity (4). Preregistration available at <https://aspredicted.org/f7jw6.pdf>.

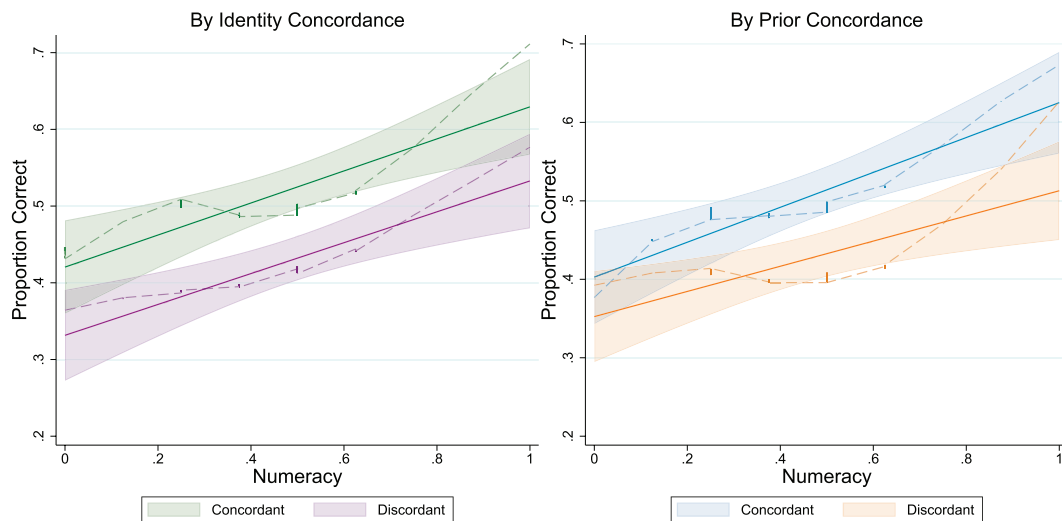
## Results

Results are shown in Fig. 1 (see *SI Appendix* for details of our modeling approach). Examining main effects, we find that when the

study's outcome was concordant with participants' political identity, they were 9.3 percentage points more likely to give the correct answer compared to when the outcome was discordant ( $\beta = 0.093$ , CI = [0.052, 0.135],  $P < 0.001$ ) and that participants with higher numeracy scores were more likely to give the correct answer ( $\beta = 0.111$ , CI = [0.035, 0.077],  $P < 0.001$ ). These results serve as manipulation checks on the identity manipulation and numeracy measure. (Note that all analyses use linear regression, and results are qualitatively equivalent when controlling for all demographics collected or when using a logistic model given the dichotomous outcome.)

Now we turn to the key test of MS2R: Does the relationship between reasoning ability and answer correctness disappear when the correct outcome is challenging to the subject's identity? The answer here is a clear “No”: When the interaction between outcome concordance and numerical ability is added to the model, we find a precisely estimated null interaction ( $\beta = 0.002$ , CI = [-0.04, 0.0043],  $P = 0.94$ ). Numeracy shows a highly significant positive relationship with answer correctness regardless of whether the outcome is concordant ( $\beta = 0.111$ , CI = [0.099, 0.317],  $P < 0.001$ ) or discordant ( $\beta = 0.110$ , CI = [0.095, 0.307],  $P < 0.001$ ), see Fig. 1A. There is also no significant three-way interaction when including the subject's political ideology in the model ( $\beta = 0.025$ , CI = [-0.024, 0.061],  $P = 0.387$ ), such that there is no significant interaction between outcome concordance and numerical ability for liberal Democrats ( $\beta = 0.006$ ,  $P = 0.879$ ) or conservative Republicans ( $\beta = -0.001$ ,  $P = 0.999$ ). Results are qualitatively equivalent when also including a quadratic term for numeracy.

There was little variation in the outcome-by-numeracy interaction across the five topics. Most notably, when looking specifically at gun control (as in the original study), we find a nonsignificant interaction between outcome concordance and numerical ability that is in the opposite direction from the prediction ( $\beta = -0.11$ ,  $P = 0.089$ ); among subjects who were shown the gun control scenario, those below the median on numeracy showed a significant effect of outcome concordance,  $\beta = 0.192$ ,  $P = 0.001$ , while those above the median on numeracy did not,  $\beta = 0.019$ ,  $P = 0.806$ . The only topic (out of five) to show an effect qualitatively similar to the original was for taxation, and even here, the interaction was not



**Fig. 1.** More numerate participants perform better at interpreting the contingency table, regardless of whether the correct answer aligns with their identity or prior beliefs. Shown is probability of giving the correct answer on the contingency table as a function of numeracy score, split by how the correct answer aligns with identity (A; green = concordant, purple = discordant) or prior beliefs (B; blue = concordant, orange = discordant). Solid lines indicate linear regression lines (with 95% CIs), and dotted lines indicate local polynomial regression locally estimated scatterplot smoothing line (LOESS) to show underlying data patterns.

significant ( $P = 0.069$ ). The interaction was also not significant for any of the other topics (all  $ps > 0.3$ ). In sum, we do not find any support for the key prediction of the MS2R account.

Finally, although we did not observe the expected interaction between outcome concordance and numeracy, we nonetheless examine the consequences of controlling for prior beliefs. A model that adds the outcome's alignment with the subject's prior beliefs about the contingency table's topic and does not include any interactions, shows little difference in the main effects of outcome concordance ( $\beta = 0.081$ ,  $P < 0.001$ ) and numeracy ( $\beta = 0.098$ ,  $P < 0.001$ ), while also showing that people are 6 percentage points more likely to respond correctly when the correct answer aligns with their prior beliefs ( $\beta = 0.061$ ,  $P = 0.014$ ). When adding the concordance-numeracy and priors-numeracy interactions to the model, we continue to find no significant interaction between outcome concordance and numeracy ( $\beta = -0.02$ ,  $P = 0.548$ ) and also find no significant change in the effect of numeracy based on how well the outcome aligns with the subject's prior beliefs ( $\beta = 0.007$ ,  $P = 0.801$ ).

## Discussion

In sum, we fail to replicate the key empirical support for the influential theory of motivated system two reasoning (MS2R) in a large, representative probability sample. Instead of reasoning ability magnifying political differences, we found evidence that people who were better at reasoning were more accurate, even for highly polarized issues. Importantly, however, we also found a robust effect of identity alignment that was unaffected by either numeracy or prior beliefs. These results raise serious questions about the key empirical evidence cited as support for MS2R theory (while nonetheless supporting the existence of some form of motivated processing) and thereby call for a reconsideration of the role of analytical, deliberative reasoning in ideologically charged contexts.

1. Z. Kunda, The case for motivated reasoning. *Psychol. Bull.* **108**, 480–498 (1990).
2. D. M. Kahan, E. Peters, E. C. Dawson, P. Slovic, Motivated numeracy and enlightened self-government. *Behav. Public Policy* **1**, 54–86 (2017).
3. D. M. Kahan, E. Peters, Rumors of the "Nonreplication" of the "Motivated Numeracy Effect" are greatly exaggerated. Yale Law & Economics Research Paper 584 (2017).
4. B. M. Tappin, G. Pennycook, D. G. Rand, Rethinking the link between cognitive sophistication and politically motivated reasoning. *J. Exp. Psychol. Gen.* **150**, 1095 (2021).
5. M. S. Nurse, W. J. Grant, I'll see it when I believe it: Motivated numeracy in perceptions of climate change risk. *Environ. Commun.* **14**, 184–201 (2020).
6. C. Ballarini, S. A. Sloman, "Reasons and the 'motivated numeracy effect'" in *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (2017), pp. 1580–1585.
7. S. G. Baker, N. Patel, C. Von Gunten, K. D. Valentine, L. D. Scherer, Interpreting politically-charged numerical information: The influence of numeracy and problem difficulty on response accuracy. *Judgm. Decis. Mak.* **15**, 203 (2020).
8. P. Connor, E. Sullivan, M. Alfano, N. Tintarev, Motivated numeracy and active reasoning in a Western European sample. *Behav. Public Policy*, 10.1017/bpp.2020.32 (2020).

## Methods

We collected  $n = 2,355$  participants from the AmeriSpeak panel via an award from the Time-Sharing Experiments for the Social Sciences group (sample size was determined by the size of the award). AmeriSpeak is a nationally representative, probability-based panel based on NORC's National Sample Frame. It is used by gold-standard survey projects in the United States and represents the level of typicality lacking in previous work. The study was conducted online in January of 2022 and took participants roughly 11 min on average to complete. The sample was probability matched and roughly reflected the overall demographics of the US population: 51.6% female, 55.7% White, non-Hispanic, median age 44 y old. Participants provided informed consent, and this study was deemed exempt by the MIT Committee on the Use of Humans as Experimental Subjects (protocol # E-3386).

Participants were randomized to one of five political topics (gun control, health care, illegal immigration, police accountability, or increased taxation). They first indicated their position on this topic, then completed a 9-item measure of numerical ability. Next, they read a description of a study that compared cities that did versus did not enact a particular policy relevant to the political issue. The results of the study were presented in a two-by-two contingency table, and participants were asked whether the results indicated that the policy increased or decreased the relevant outcome (e.g., did gun control increase or decrease gun deaths). The rows of the contingency table (and thus the correct answer) were randomized between subjects. Further, the numbers used in each cell of the contingency table were the same as those used in the original study. See *SI Appendix* for exact wording for all items.

**Data, Materials, and Software Availability.** CSV data have been deposited in OSF (DOI: [10.17605/OSF.IO/Q2VCM](https://doi.org/10.17605/OSF.IO/Q2VCM)) (16).

**ACKNOWLEDGMENTS.** We gratefully acknowledge funding from the Time-sharing Experiments (TESS) for the Social Sciences program, the Office of Naval Research (N00014-21-1-2573), the National Science Foundation (FAIN 2047152), the Alfred P. Sloan Foundation (Grant #2021-16891), and the John Templeton Foundation (Grant #61779).

9. T. Lind, A. Erlandsson, D. Västfjäll, G. Tinghög, Motivated reasoning when assessing the effects of refugee intake. *Behav. Public Policy* **6**, 213–236 (2022).
10. A. Maguire, E. Persson, D. Västfjäll, G. Tinghög, COVID-19 and politically motivated reasoning. *Med. Decis. Mak.* **42**, 1078–1086 (2022).
11. E. Chung, P. Govindan, A. O. Pechenkina, The effect of incentives on motivated numeracy amidst COVID-19. *J. Exp. Pol. Sci.*, 10.1017/XPS.2022.32 (2022).
12. E. Persson, D. Andersson, L. Koppel, D. Västfjäll, G. Tinghög, A preregistered replication of motivated numeracy. *Cognition* **214**, 104768 (2021).
13. C. Strömbäck, D. Andersson, D. Västfjäll, G. Tinghög, Motivated reasoning, fast and slow. *Behav. Public Policy*, 10.1017/bpp.2021.34 (2021).
14. B. Guay, C. D. Johnston, Ideological asymmetries and the determinants of politically motivated reasoning. *Am. J. Pol. Sci.* **66**, 285–301 (2022).
15. D. M. Kahan, "The politically motivated reasoning paradigm" in *Emerging Trends in Social & Behavioral Sciences* (2015).
16. M. N. Stagnaro, B. Tappin, D. G. Rand, Unmotivated numeracy and self governance data. OSF. <https://osf.io/q2vcm/>. Deposited 11 July 2023.