The Fake News Effect: Experimentally Identifying Motivated Reasoning Using Trust in News[†]

By Michael Thaler*

Motivated reasoning posits that people distort how they process information in the direction of beliefs they find attractive. This paper creates a novel experimental design to identify motivated reasoning from Bayesian updating when people have preconceived beliefs. It analyzes how subjects assess the veracity of information sources that tell them the median of their belief distribution is too high or too low. Bayesians infer nothing about the source veracity, but motivated beliefs are evoked. Evidence supports politically motivated reasoning about immigration, income mobility, crime, racial discrimination, gender, climate change, and gun laws. Motivated reasoning helps explain belief biases, polarization, and overconfidence. (JEL C91, D12, D72, D83, D91, L82)

On many topics, people disagree about the answers to factual questions, and their beliefs are often inaccurate in predictable directions. People have differing beliefs about questions related to income mobility, crime rates, and racial discrimination in labor markets; tend to be biased in the direction that is more representative of their political party's stances; and often overestimate their own political knowledge (e.g., Alesina, Stantcheva, and Teso 2018; Flynn, Nyhan, and Reifler 2017; Ortoleva and Snowberg 2015). As shown by Gerber and Huber (2009); Meeuwis et al. (2022); and Allcott et al. (2020), these partisan beliefs can affect consumer, financial, and public health behavior. Given the significance of these issues, why does such bias and belief polarization persist?

After receiving a piece of news, people immediately form posterior beliefs in ways that incorporate both their prior beliefs and how they perceive the informativeness of the news. If we only observe beliefs at a snapshot in time, two people's disagreement can be consistent with many explanations: for instance, they may have started with different beliefs, they may receive news from different sources, or they may have

differently distorted inference processes. The first two channels are often relevant in politicized settings. First, Democrats and Republicans often have different priors, leading to differences in posteriors; this can be consistent with Bayes' rule. Second, Democrats and Republicans often consume different media, and may find arguments akin to those from MSNBC and from Fox News differentially informative.[1]

This paper introduces a new tool for detecting a third channel, *motivated reasoning*, which posits that people distort their inference process in the direction of states they find more attractive. In many settings, we may think that people are motivated to hold beliefs that align with their political party's stances: Republicans are motivated to believe that increasing immigration is associated with higher crime rates; Democrats are motivated to believe that there is severe racial discrimination in labor markets; and both parties are motivated to believe that members of their party are more knowledgeable about these issues than are members of the opposing party. When people receive news, they are often reminded of what beliefs they currently hold and of beliefs they find more attractive to hold (*motives*). They then use these motives when making inferences from and about news sources.

While there is an intuition in the literature that motivated reasoning plays an important role in inference, particularly in political settings, designing an experiment to identify this mechanism has faced two major challenges. The first challenge is that it is difficult to identify motivated reasoning from Bayesian updating with sufficient statistical power when subjects' *unmotivated* biases outweigh motivated reasoning. As summarized by Benjamin (2019), designs in which people receive informative signals have often not been able to find statistically significant evidence for motivated reasoning: Möbius et al. (2022); Eil and Rao (2011); and Charness and Dave (2017) find that people update more from "good news" than "bad news," while Ertac (2011); Kuhnen (2014); Buser, Gerhards, and van der Weele (2018); Coutts (2019); and Barron (2021) do not find evidence supporting this hypothesis.[2] The typical design for these papers involves giving subjects partially informative news and testing for asymmetric updating from "Good," "Bad," and "Neutral" signals, a design that has consistently led to substantial underupdating (Benjamin 2019).[3]

This challenge has led some researchers to argue that the importance of motivated reasoning is dwarfed by unmotivated explanations like conservatism or "lazy thinking" (Pennycook and Rand 2019; Benjamin 2019; Tappin, Pennycook, and Rand 2020a, b). While such arguments may be valid in settings where there is both Bayesian inference and motivated reasoning, I bypass this issue by constructing an experimental paradigm that gives subjects messages in which there is *nothing* for a Bayesian to infer, but

---

[1] There is ample evidence consistent with these channels (e.g., Taber and Lodge 2006; Kahan et al. 2012; Nyhan and Reifler 2010, 2013; Nyhan, Reifler, and Ubel 2013).

[2] It is worth noting that there is more consistent evidence for motivated decision-making and memory. This includes information avoidance and moral wiggle room (Oster, Shoulson, and Dorsey 2013; Dana, Weber, and Kuang 2007; Gino, Norton, and Weber 2016), risk- and ambiguity-driven distortions (Exley 2016; Haisley and Weber 2010), and recall of past information (Zimmermann 2020; Chew, Huang, and Zhao 2020). However, these papers do not show that updating in response to *new* information is a driving channel.

[3] Another strand of literature has empirically and experimentally considered how varied incentives affect beliefs (e.g., Babcock, Loewenstein, et al. 1995; Babcock and Loewenstein 1997; Schwardmann, Tripodi, and van der Weele 2022). Using a different design, Exley and Kessler (2022) show that subjects behave differently in moral domains when they receive two pieces of equivalent information in differently complex ways. My paper focuses more on the inference process for new information.

motivated reasoning still predicts directional distortions. My design eliminates the role that Bayes' rule, conservatism, and confirmatory biases can play in inference, allowing for a cleaner identification of motivated reasoning. It therefore provides a more precise treatment-effect estimate, which enables me to find novel evidence identifying motivated reasoning in numerous previously hypothesized domains.

The second challenge is that in many domains of interest, such as politics, people enter into experiments with preconceived beliefs. Eliciting motivated reasoning in such settings is difficult; researchers typically either restrict tests to questions that have binary answers (e.g., Möbius et al. 2022), ask numerous questions in order to elicit entire prior and posterior belief distributions (e.g., Eil and Rao 2011), or do not fully rule out Bayesian updating (e.g., Sunstein et al. 2017). My design is able to identify motivated reasoning from Bayesian updating on questions that have numerical answers, while only requiring elicitation of one moment of subjects' belief distribution: their median. Researchers can use this design to test how people motivatedly reason about essentially any factual question with a numerical answer.

The design has two main steps. First, each subject is given a variety of factual questions with numerical answers. On each question, each subject selects a response that they think is equally likely to be above or below the correct answer; that is, the median of their belief distribution is elicited. Second, the subject is given one binary message that is chosen randomly from either a True News source or a Fake News source; the message tells her whether the answer was above or below her median. If the message is from True News, it is always accurate. If the message is from Fake News, it is always inaccurate. The subject is not told which source the message came from; instead, she is asked to make inferences about the source's veracity from the message she receives.

Since messages relate the true answer to subjects' subjective median, a Bayesian (she) would believe that it is equally likely for each source to report either message. That is, she has stated that she believes the answer is equally likely to be greater than or less than her median; so, she believes the likelihood that a True News source would report that the answer is greater than her median is one-half, and the likelihood that a Fake News source would report that the answer is less than her median is also equal to one-half. Therefore, the Bayesian would find a "greater than" message to be completely uninformative about the veracity of the news source. Likewise, she would find a "less than" message to be uninformative.

On the other hand, a subject who engages in motivated reasoning (he) will trust the news more if it sends a message that supports what he is more motivated to believe. If he engages in *politically* motivated reasoning, he will assess messages that align with the beliefs of his political party (*Pro-Party news*) to be more truthful, while assessing messages that misalign (*Anti-Party news*) to be less truthful.

I test for motivated reasoning using a within-subject experiment on Amazon Mechanical Turk with approximately 1,000 people in the United States. As a testament to the portability of the design, I run the test on nine different economic, political, and social questions (*politicized* topics), and on one question about own performance in the experiment. The list of topics and hypothesized motives is in Table 1.

The main finding from the experiment is that Bayesian updating is strongly rejected in favor of politically motivated reasoning on the politicized topics. While a Bayesian

TABLE 1—THE LIST OF TOPICS AND HYPOTHESIZED MOTIVES IN THE EXPERIMENT

Topic	Pro-Democrat motives	Pro-Republican motives
US crime	Got better under Obama	Got worse under Obama
Upward mobility	Low in US after tax cuts	High in US after tax cuts
Racial discrimination	Severe in labor market	Not severe in labor market
Gender	Girls better at math	Boys better at math
Refugees	Decreased violent crime	Increased violent crime
Climate change	Scientific consensus	No scientific consensus
Gun reform	Decreased homicides	Didn't decrease homicides
Media bias	Media not mostly Dems	Media mostly Dems
Party performance	Higher for Dems over Reps	Higher for Reps over Dems
Own performance	Higher for self over others	Higher for self over others

would believe that Pro-Party and Anti-Party news are equally likely to be True News on the politicized topics, subjects in the experiment believe that Pro-Party messages are 9 percentage points more likely than Anti-Party messages to come from the True News source. These effects are moderate but very precisely estimated (s.e. 0.6 percentage points; $p \approx 10^{-40}$). This test is precise enough to measure each topic individually; for eight of the nine politicized topics, the treatment effect is statistically significantly larger than zero ($p \leq 0.001$). To my knowledge, this experiment provides the first evidence for motivated reasoning on these questions that is not confounded by Bayesian updating or prior-confirming biases (like those in Festinger 1957; Lord, Ross, and Lepper 1979; Tetlock 1983; and Rabin and Schrag 1999). In recent work, Eyting (2022) has adapted this design to study motivated reasoning–based discrimination.[4] Treatment effects are larger for subjects who are more partisan. I also find consistent evidence that subjects motivatedly reason about their performance.

The new design could lead to potential confounds that are unique to this experiment, such as if subjects systematically misreport their median belief or misinterpret the experiment's definitions of True News and Fake News, but I find that results are unlikely to be explained by these alternative hypotheses.[5] To account for unmotivated mistakes, I also compare behavior on political topics to three neutral topics, and find that news veracity assessments on neutral topics are lower than assessments on Pro-Party news and higher than assessments on Anti-Party news.

The second main finding is that people's systematically biased beliefs about these topics is related to their motivated beliefs (as also discussed by Eil and Rao 2011). Since people who motivatedly reason about an issue will systematically distort their beliefs, we can partly infer what people's motives are by looking at their current beliefs. That is, the direction of one's *error* predicts the direction of one's *motive*. In the context of the experiment, this hypothesis predicts that people will give higher

[4] Other papers that find asymmetric responses to information on related topics include: Taber and Lodge 2006 [gun laws]; Alesina, Stantcheva, and Teso 2018 [upward mobility]; Cappelen, Haaland, and Tungodden (2018) [responses to taxation]; Haaland and Roth (2023) [racial labor market discrimination]; Sarsons (2017) and Kunda and Sinclair (2000) [gender and performance]; Alesina, Miano, and Stantcheva (2023), Haaland and Roth (2020), and Druckman, Peterson, and Slothuus (2013) [immigration]; Nyhan and Reifler 2013 and Nyhan, Reifler, and Ubel 2013 [perceptions of elected officials]; and Sunstein et al. 2017 [climate change]. Many findings from these papers may be due to motivated reasoning.

[5] As described in Section IVD and appendices, predictions are identical if subjects mistakenly believe Fake News sends random messages instead of always-false messages, and results are not driven by subjects who have skewed belief distributions and misreport their median.

veracity assessments to news that (falsely) reinforces their error compared to news that (truthfully) brings them closer to the correct answer. Indeed, subjects in the experiment trust the error-reinforcing Fake News more than the error-correcting True News, and only do so on topics where motivated reasoning is expected to play a role.

I also show how motivated reasoning can lead to apparent overprecision, arguably the most durable form of overconfidence (Moore and Healy 2008; Moore, Tenney, and Haran 2015). Motivated reasoning may provide a link between overprecision and partisanship, a relationship documented in Ortoleva and Snowberg (2015), because partisans' belief distributions are more miscalibrated. Miscalibrations are more severe with stronger motives, and this leads 50-percent confidence intervals to contain the true answer less than 50 percent of the time, especially for partisans.

Motivated reasoning not only affects how people trust or distrust news, but also impacts how people change their beliefs about the topics themselves. Despite being uninformative, the messages lead to further belief polarization: subjects are more likely to revise their beliefs away from the population mean than toward it. That is, informational content is not a necessary condition for belief polarization.[6] Politically motivated reasoning helps reconcile the notions that the ideological polarization of beliefs may be high, even if the ideological polarization of information acquisition is modest (Gentzkow and Shapiro 2011).[7] It can also explain why news providers may distort messages to appeal to their audience's motivated beliefs (Thaler 2023).

There are no sizable demographic heterogeneities in politically motivated reasoning, neither in direction nor magnitude, once party preference is controlled for. Differences in treatment effects are almost all statistically indistinguishable from zero, and estimates are precise enough to rule out modest effect sizes.[8] Motivated beliefs on these questions seem to be principally driven by politics.

Lastly, to complement the experiment, I discuss a model of motivated reasoning in which people distort their inference when they receive new information. This model adds to the decades-old theoretical literature (such as Akerlof and Dickens 1982; Carrillo and Mariotti 2000; and Bénabou and Tirole 2002) by emphasizing the nonstrategic nature of the bias: agents in this model do not distort their beliefs for functional reasons.[9] They make inferences using a modified Bayes' rule, weighting priors and likelihoods as a Bayesian would, but act as if they receive an extra signal that puts more weight on higher-motive states.[10] This leads them to make inferences even when a Bayesian does not.

---

[6] Linking this finding to the polarization in trust in news relates to a growing literature that discusses the relationship between trust in news and political partisanship (NCG15; L13; DLM18).

[7] Gentzkow and Shapiro (2006) and Gentzkow, Wong, and Zhang (2018) provide alternative theoretical explanations with Bayesian agents who have different priors, but these models do not predict updating from uninformative signals.

[8] On the other hand, Thaler (2021) shows that performance-motivated reasoning has substantial heterogeneity by gender; only men systematically motivatedly reason to think their performance was higher, while women are Bayesian on average.

[9] This directly contrasts the model with optimal beliefs models such as Brunnermeier and Parker (2005) and Möbius et al. (2022). It is more similar to the selective memory models of Bénabou and Tirole (2002, 2011) and the Bayesian model of Mayraz (2019), but emphasizes distorting the processing of new information.

[10] This is a similar theory as Kahan (2016) but differs from related economic models of motivated cognition like Bénabou and Tirole (2011) and Möbius et al. (2022), in that it assumes directional distortions but not underreaction to information.

To further ensure that experimental results were unlikely to be driven by noise, I retested the main hypotheses using a preregistered replication one year later on a new sample. The findings successfully replicated; for details, see online Appendix D.

The rest of the paper proceeds as follows: Section I develops the model of motivated reasoning. Section II presents the experimental design and main hypotheses. Section III analyzes experimental results, including the evidence for motivated reasoning, treatment effect heterogeneity, and overprecision. Section IV concludes and proposes directions for future work. Other details are in the online Appendix.

I. Theory

A. *Defining and Identifying Motivated Reasoning*

This section outlines the model of motivated reasoning. Agents distort how they process new information by acting as if they receive an additional signal corresponding to their motivated beliefs. When signals are uninformative to a Bayesian, as will be the case in the experiment, this distortion can lead motivated reasoners to update in the direction that they find more attractive.

There are a set of agents i and a set of questions q. For each question, Nature draws a state θ_q from $\Theta_q \subseteq \mathbb{R}$. For each agent and question, Nature draws a source from $\mathcal{S}$: the source is either True (T_{iq}) or Fake (F_{iq}). Sources are independently drawn with $\Pr(T_{iq}) = p$, and agents receive data x_{iq} about the source veracity.

Agents engage in *motivated reasoning*, which means that they form their posterior by incorporating prior, likelihood, and a motivated beliefs term:

$$\underbrace{\Pr(T_{iq}|x_{iq})}_{\text{posterior}} \propto \underbrace{\Pr(T_{iq})}_{\text{prior}} \cdot \underbrace{\Pr(x_{iq}|T_{iq})}_{\text{likelihood}} \cdot \underbrace{M_{iq}(T_{iq},x_{iq})^{\varphi}}_{\text{mot. reasoning}}.$$

We take log odds ratios of both sides (Grether 1980) to attain an additive form:[11]

$$\text{(1)} \qquad \text{logit} \Pr(T_{iq}|x_{iq}) = \text{logit}\, p + \log\left(\frac{\Pr(x_{iq}|T_{iq})}{\Pr(x_{iq}|F_{iq})}\right)$$

$$+ \varphi\left[m_{iq}(T_{iq},x_{iq}) - m_{iq}(F_{iq},x_{iq})\right].$$

Motivated reasoners act as if they receive both the actual signal (x_{iq}) and a signal whose relative likelihood corresponds to how much they are motivated to believe the state is T_{iq} given x_{iq}. $m_{iq} : \mathcal{S} \to \mathbb{R}$ is called the *motive* function. The weight put on this signal is $\varphi \geq 0$, called *susceptibility*. $\varphi = 0$ corresponds to Bayes' rule; $\varphi > 0$ corresponds to motivated reasoning.

I now introduce an information structure in which the x_{iq} are binary and uninformative about the news source veracity, but affect m_{iq}. The x_{iq} are chosen such that $\Pr(x_{iq}|T_{iq}) = \Pr(x_{iq}|F_{iq}) = 1/2$. In such a setting, described below, a Bayesian

---

[11] Mayraz (2019) uses a similar formulation in which motivated beliefs affect priors instead of belief updating.

will set $\Pr(T_{iq}|x_{iq}) = p$, but a motivated reasoner will set $\operatorname{logit}\Pr(T_{iq}|x_{iq}) = \operatorname{logit} p + \varphi[m_{iq}(T_{iq},x_{iq}) - m_{iq}(F_{iq},x_{iq})]$. That is, if the motivated reasoner finds x_{iq} more attractive, then they will form a posterior that the message comes from a truthful source. Meanwhile, a Bayesian will not update from either message because the message is uninformative about the source.

In many environments, beliefs about $\Pr(x_{iq}|T_{iq})$ relate to $m_{iq}(T_{iq},x_{iq})$ since agents who are motivated to believe x_{iq} is true may currently believe that x_{iq} is true. As such, it is necessary to construct a signal structure that tailors messages to agents' *current* beliefs.

For each q, agents report a median belief μ_{iq} such that they believe θ_q is equally likely to be greater than μ_{iq} or less than μ_{iq}.[12] Nature then chooses one of two news sources from $\mathcal{S}$. The news source sends a message to the agent that either says that $\theta_q > \mu_{iq}$ or says that $\theta_q < \mu_{iq}$: $x_{iq}(\mu_{iq}) \in \{G(\mu_{iq}), L(\mu_{iq})\}$. The true source always sends the "truthful" message and the fake source always sends the "untruthful" message:

	$\theta_q > \mu_{iq}$	$\theta_q < \mu_{iq}$
True News sends	$G(\mu_{iq})$	$L(\mu_{iq})$
Fake News sends	$L(\mu_{iq})$	$G(\mu_{iq})$

After receiving $x_{iq}(\mu_{iq})$, agents report how likely they think it is that the message came from the true source: $a(x_{iq}(\mu_{iq}))$. They also report a revised median belief $\mu'_{iq}(x_{iq}(\mu_{iq}))$. Finally, the states θ_q and sources are revealed.

Agents may either be motivated to believe θ_q is higher than μ_{iq} or θ_q is lower than μ_{iq}. An agent with a high-θ_q motive has $m_{iq}(T_{iq},G(\mu_{iq})) = m_{iq}(F_{iq},L(\mu_{iq})) > m_{iq}(T_{iq},L(\mu_{iq})) = m_{iq}(F_{iq},G(\mu_{iq}))$, and vice versa for an agent with a low-θ_q motive.

We now consider how a Bayesian and a motivated reasoner update their beliefs about the news source. Given message $G(\mu_{iq})$, the Bayesian updates according to equation (1) with $\varphi = 0$:

$$\operatorname{logit} a_B(G(\mu_{iq})) = \operatorname{logit}\Pr(T(\mu_{iq})|G(\mu_{iq}))$$

$$= \operatorname{logit}\Pr(T(\mu_{iq})) + \log\left(\frac{\Pr(G(\mu_{iq})|T(\mu_{iq}))}{\Pr(G(\mu_{iq})|F(\mu_{iq}))}\right)$$

$$= \operatorname{logit} p + \log\left(\frac{\Pr(\theta_q > \mu_{iq})}{\Pr(\theta_q < \mu_{iq})}\right)$$

$$= \operatorname{logit} p.$$

Therefore: $a_B(G(\mu_{iq})) = p = a_B(L(\mu_{iq}))$.

---

[12] We assume that there is no atom at μ_{iq} and that $\Pr(\mu_{iq} = \theta_q) = 0$. That is, the agent believes that the answer has probability zero of being exactly equal to μ, and the probability is indeed zero.

Since the Bayesian has stated her median belief μ_{iq}, she thinks that both messages are equally likely ex ante and does not update in either direction.

Meanwhile, the motivated reasoner updates according to equation (1) with $\varphi > 0$:

$$
\begin{aligned}
\operatorname{logit} a\big(G(\mu_{iq})\big) &= \operatorname{logit} \Pr\big(T(\mu_{iq})\big) + \log\left(\frac{\Pr\big(G(\mu_{iq})\,|\,T(\mu_{iq})\big)}{\Pr\big(G(\mu_{iq})\,|\,F(\mu_{iq})\big)}\right) \\
&\quad + \varphi\Big[m_{iq}\big(T_{iq}, G(\mu_{iq})\big) - m_{iq}\big(F_{iq}, G(\mu_{iq})\big)\Big] \\
&= \operatorname{logit} p + \varphi\Big[m_{iq}\big(T_{iq}, G(\mu_{iq})\big) - m_{iq}\big(F_{iq}, G(\mu_{iq})\big)\Big].
\end{aligned}
$$

That is, we have the following:

PROPOSITION 1 (Identifying motivated reasoning using news veracity assessments): *The procedure above identifies motivated reasoning from Bayesian updating:*

- *With Bayesian updating ($\varphi = 0$):* $a_B\big(G(\mu_{iq})\big) = a_B\big(L(\mu_{iq})\big).$
- *With motivated reasoning ($\varphi > 0$):*

$$
a\big(G(\mu_{iq})\big) > a\big(L(\mu_{iq})\big) \Leftrightarrow m_{iq}\big(T_{iq}, G(\mu_{iq})\big) > m_{iq}\big(T_{iq}, L(\mu_{iq})\big).
$$

This design identifies whether agents are more motivated to believe that the true state is above their median belief μ_{iq} or to believe that the true state is below μ_{iq}.

More precisely, this design provides a joint test of $\varphi > 0$ and the direction of motives. In the political context, we will consider agents who are either of a Republican type or of a Democratic type, assume that $m_{Rq}\big(S_{Rq}, x(\mu_{Rq})\big) = -m_{Dq}\big(S_{Dq}, x(\mu_{Dq})\big)$, and then test whether, given these motives, $\varphi > 0$.

B. *Discussion*

Since signals are uninformative, this design also identifies motivated reasoning from several other non-Bayesian, models of inference. Consider the following generalized class of updating rules:

$$
\begin{aligned}
\operatorname{logit} a\big(G(\mu_{iq})\big) &= \zeta \operatorname{logit} p + \kappa \log\left(\frac{\Pr\big(G(\mu_{iq})\,|\,T(\mu_{iq})\big)}{\Pr\big(G(\mu_{iq})\,|\,F(\mu_{iq})\big)}\right) \\
&= \zeta \operatorname{logit} p = \operatorname{logit} a\big(L(\mu_{iq})\big).
\end{aligned}
$$

This class of updating rules includes a prior-confirming bias ($\zeta > 1$), conservatism ($\kappa < 1$), and base-rate neglect ($\zeta < 1$). In all these cases, agents form the same posterior after $G(\mu_{iq})$ and $L(\mu_{iq})$ (even if κ differs for good and bad news).

In addition, by taking Θ_q as the domain for the motive function, the same logic as in Section IA leads motivated reasoners to be more likely to revise their median

beliefs in the higher-motive direction. A Bayesian will have $\Pr(\mu'_{iq}(x_{iq}) > \mu_{iq}) = \Pr(\mu'_{iq}(x_{iq}) < \mu_{iq})$. However, a motivated reasoner who has $m_{iq}(\theta_q)$ increasing in θ_q will be more likely to update upwards: $\Pr(\mu'_{iq}(x_{iq}) > \mu_{iq}) > \Pr(\mu'_{iq}(x_{iq}) < \mu_{iq})$, and vice versa if $m_{iq}(\theta_q)$ is decreasing in θ_q.

Online Appendix C discusses extensions to the model that can generate further predictions. Agents who start out believing that μ_{iq} is larger, due to receiving signals prior to this game, will be motivated to believe that Θ_q is even larger than μ_{iq}, as current beliefs are reflective of motivated beliefs. With additional functional form assumptions, motivated reasoning also leads to overprecision and overconfidence.

II. Experimental Design

This section details the experiment introduced above. Section IIA outlines the experimental design, Section IIB discusses specific pages that subjects see, and Section IIC generates the main hypotheses. Section IID discusses the participant sample and further details of the data collected. Specific question wordings and screenshots for every page type, including instructions and scoring rules, are in the online Appendices.

A. *Summary*, *Timeline*, *and Topics*

To fix ideas, consider the following question, taken verbatim from the experiment:

> Some people believe that the Obama administration was too soft on crime and that violent crime increased during his presidency, while others believe that President Obama's pushes towards criminal justice reform and reducing incarceration did not increase violent crime.

> This question asks how murder and manslaughter rates changed during the Obama administration. In 2008 (before Obama became president), the murder and manslaughter rate was 54 per million Americans.

> In 2016 (at the end of Obama's presidency), what was the per-million murder and manslaughter rate?

This is a question for which we will hypothesize that Republicans are motivated to believe the answer is higher, and Democrats are motivated to believe the answer is lower. The test of motivated reasoning involves the following steps:

(i) *Beliefs:* Subjects are asked to guess the answers to politicized questions like the one above. They are asked and incentivized to guess their median belief (such that they find it equally likely for the answer to be above or below their guess). They are also asked and incentivized for their interquartile range. Details on incentives are below.

(ii) *News:* On each question subjects receive one binary message from one of two randomly chosen news sources: True News and Fake News. The message from

True News is always correct, and the message from Fake News is always incorrect. This is the main (within-subject) treatment variation.

The message says either "The answer is *greater than* your previous guess of [previous guess]." or "The answer is *less than* your previous guess of [previous guess]." Note that the exact messages are different for subjects who make different initial guesses.

For the question above, "greater than" corresponds to pro-Republican news and "less than" to pro-Democratic news. For subjects who support the Republican Party more than the Democratic Party, "greater than" is coded as Pro-Party news and "less than" is coded as Anti-Party news, and vice versa for subjects who support the Democratic Party more.

(iii) *Assessment:* After receiving the message, subjects assess the probability that the message came from True News using a scale from 0/10 to 10/10, and are incentivized to state their true belief. I test for motivated reasoning by looking at the treatment effect of variation in news direction on news veracity assessments.

Recall that since subjects receive messages that compare the answer to their median, a Bayesian would not change her assessment based on the message. If she had a prior that $\Pr(\textit{True News}) = 1/2$ before seeing the message, she would form a posterior that $\Pr(\textit{True}|\text{"greater than"}) = \Pr(\textit{True}|\text{"less than"}) = 1/2$. We attribute systematic treatment effects of the messages on assessments to motivated reasoning. For instance, if Republicans tend to state $\Pr(\textit{True}|\text{"greater than"}) > \Pr(\textit{True}|\text{"less than"})$ and Democrats tend to state $\Pr(\textit{True}|\text{"greater than"}) < \Pr(\textit{True}|\text{"less than"})$ on the question above, this would be coded as politically motivated reasoning.

Subjects see 14 questions in the experiment, including the 10 topics in Table 1, 3 neutral questions, and 1 comprehension check.[13] The timing is as follows:
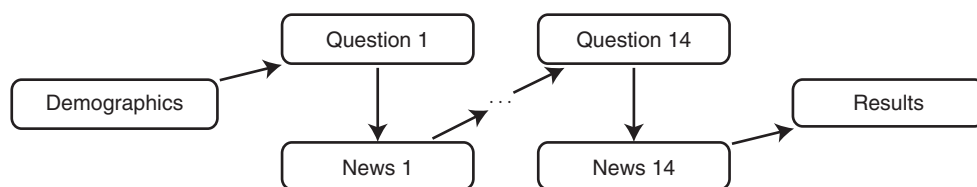
The demographics page includes questions about party ratings, party affiliation, ideology, gender, age, race and ethnicity, annual income, highest education level, state or territory of residence, and religion. Party ratings are used to categorize subjects' party preferences; subjects are asked to rate the Democratic and Republican parties using a 0–100 scale that is akin to the feeling thermometer used in the American National Election Studies.

---

[13] Neutral questions ask about the latitude and longitude of the center of the United States and about a random number drawn from 0–100. The comprehension check asks what the current year is.

The results page tells subjects what their overall performance was, what their score on each question and assessment was, and the correct answer to each question and assessment. Subjects are told that they will see this page at the beginning of the experiment, and they are forced to go through it before exiting the study and receiving payment. Forcing subjects to see the answers aims to limit subjects' scope for strategic self-deception.[14]

The order of Questions 1–12 is randomized between subjects, but Questions 13 and 14 are fixed.[15] Question 13 asks subjects how they scored on Questions 1–12 relative to 100 pilot subjects. Question 14 asks subjects to compare Democratic pilot subjects' performance to Republican pilot subjects' performance on Questions 1–12.[16]

B. *Pages and Scoring Rules*

Overall Scoring Rule.—At the end of the experiment, subjects earn a show-up fee of \$3 and either receive an additional \$10 bonus or no additional bonus. As described below, in each round of the experiment, subjects earn between 0–100 "points" based on their performance. These points correspond to the probability that the subject wins the bonus: a score of x points corresponds to an $x/10$ percent chance of winning the bonus.[17]

Questions Page.—On question pages, subjects are given the round number, the topic, the text of the question, and are asked to input three numbers about their initial beliefs:

- *My Guess*: This elicits the median of subjects' prior distribution.
- *My Lower Bound*: This elicits the twenty-fifth percentile of subjects' prior distribution.
- *My Upper Bound*: This elicits the seventy-fifth percentile of subjects' prior distribution.

The scoring rule for guesses is piecewise linear. Subjects earn $\max\{100 - |c - g|, 0\}$ points for a guess of g when the correct answer is c. Subjects are told that they will maximize expected points by stating the median of their belief distribution.

The scoring rule for bounds is piecewise linear with different slopes. For upper bound ub, subjects earn $\max\{100 - 3(c - ub), 0\}$ points if $c \geq ub$ and

---

$\max\{100 - (ub - c), 0\}$ points if $c \leq ub$. For lower bound lb, subjects earn $\max\{100 - (c - lb), 0\}$ points if $c \geq lb$ and $\max\{100 - 3(lb - c), 0\}$ points if $c \leq lb$. Subjects maximize expected points by setting ub to be the seventy-fifth percentile and lb to be the twenty-fifth percentile of their belief distribution. Subjects are restricted to give answers for which My Lower Bound $\leq$ My Guess $\leq$ My Upper Bound; if they do not, they see an error message.

News Assessments Page.—After submitting *My Guess*, subjects see a second page about the same question. At the top of the page is the text of the original question. Below the question, there is a message relating the correct answer to the number they submitted for *My Guess*. This message says either:

- "The answer is *greater than* your previous guess of [*My Guess*]." or
- "The answer is *less than* your previous guess of [*My Guess*]."

Subjects are told that True News always tells the truth and Fake News never tells the truth, and that sources are independent across questions. Below the message, subjects are asked: "Do you think this information is from True News or Fake News?" and choose one of 11 radio buttons that say "x/10 chance it's True News, $(10 - x)/10$ chance it's Fake News" from each $x = 0, 1, \ldots, 10$.

The scoring rule for assessments is quadratic. For assessment a, subjects earn $100\left[1 - (1 - a)^2\right]$ points if the source is True News and $100\left(1 - a^2\right)$ points if it is Fake News. Subjects maximize expected points by stating the closest multiple of 0.1 to their belief. They are given a table with the points earned as a function of each assessment and news type.

Occasionally, a subject will correctly guess the answer. If this happens, she skips the news assessment page and moves on to the next question.[18]

Half of subjects also see a "Second Guess" part of the news assessment page. For these subjects, below each news assessment question they are asked an additional question: "After seeing this message and assessing its truthfulness, what is your guess of the answer to the original question?" Subjects are given the same linear scoring rule as on the initial guess. They earn $\max\{100 - |c - g|, 0\}$ points for a guess of g when the correct answer is c. These second guesses will be used as robustness, as well as to test for belief polarization about how subjects change their guesses.

The other half of subjects see an additional page between Question 12 and News 12, on which they are given instructions and asked to submit a willingness-to-pay (WTP) for a message. Due to space constraints, this group is not discussed in the main text. For instructions and results, see the online Appendix.[19]

---

[18] This is not the case for the comprehension check question, where the message says "The answer is *equal/not equal* to your previous guess of [*My Guess*]."

[19] I find that subjects have a positive valuation of these messages, but that they do not differentially value messages on politicized and neutral topics; results suggest that subjects are aware that they treat signals as informative, but not that they expect to distort their updating process.

I randomly vary whether subjects are given a prior at the beginning of the experiment that $\Pr(\textit{True News}) = 0.5$ or not given any prior. One-third of subjects are given a prior. This randomization serves as a robustness check to ensure that results are not entirely driven by either updating over meta-priors over the probability of True News (as may be the case for the group who is not given a prior) or by anchoring to a 50–50 benchmark (as may be the case for the group who is given a prior). I find no statistically significant differences between the two groups on any main outcomes, so results in the main text of this paper pool all subjects. Results are robust to restricting the sample to individual groups, as shown in the online Appendix.

C. *Hypotheses*

The main hypothesis is that a news veracity assessment will be larger when it leads to a higher motive. This is a joint test that (i) people motivatedly reason, giving higher assessments to news in the direction of higher motives than to news in the direction of lower motives, and (ii) the predicted direction of motives is as in Table 1. News on neutral topics is assumed to not affect motives. In the context of the model, this means that $m_{iq}(\textit{True}_{iq}) = m_{iq}(\textit{Fake}_{iq})$ when topic q is neutral and is strictly greater or strictly less when q is politicized. For politicized topics, the level of partisanship affects the steepness of the motive function.

HYPOTHESIS 1 (Motivated reasoning with political motives):

- *Assessments of Pro-Party news are greater than assessments of Anti-Party news.*
- *Assessments of neutral topic news lie in between those of Anti-Party news and those of Pro-Party news.*
- *The gap in assessments of Pro-Party news and Anti-Party news increases in partisanship.*

The hypothesis for pro-performance news and anti-performance news is similar.

The next hypothesis is based on an observation that subjects may have different beliefs because these beliefs reflect past instances of motivated reasoning. When motives are unobservable, an experimenter can learn about subjects' motives by looking at their initial beliefs. Two subjects who motivatedly reason in different directions will hold different median beliefs: A motivated reasoner with an increasing motive function will be more likely to hold a median belief that is too high, and a motivated reasoner with a decreasing motive function will be more likely to hold a median belief that is too low. Therefore, a subject whose median is too high is more likely to have an increasing motive function as compared to a subject whose median is too low.

If the two subjects make news assessments using the procedure above, they will trust news that *reinforces* the error in their initial beliefs more than news that *mitigates* the error. This behavior occurs even though signals are designed so that their interpretation is distinct from their beliefs. In the experiment, this behavior implies that subjects will give higher assessments to error-reinforcing news compared to

error-mitigating news. Recalling that error-reinforcing news is Fake News and error-mitigating news is True News, we have the following:

HYPOTHESIS 2 (Motivated reasoning and trust in Fake News):

- *Assessments of Fake News are greater than assessments of True News on politicized topics but not on neutral topics.*
- *Assessments of Fake News are greater than assessments of True News, controlling for whether the news is Pro-Party or Anti-Party.*

We can use the second-guess group to construct an alternative test of motivated reasoning. By comparing subjects' first and second guesses, we can retest the main politically motivated reasoning prediction and also study a form of belief polarization.

First, as in Hypothesis 1, subjects are expected to more frequently adjust their guesses in the direction that favors their preferred party. Second, by a similar logic to 2, motivated reasoning would lead subjects to be more likely to adjust their guesses away from the population mean than toward it. We define *Follow Message* as the ternary variable that takes value:

- 1 if the message says G and the second guess is greater than μ, or if the message says L and the second guess is less than μ;
- 0 if the second guess equals μ; and
- -1 if the message says G and the second guess is less than μ, or if the message says L and the second guess is greater than μ.

Polarizing says G if μ is greater than the population mean guess or L if μ is less than the population mean guess. Anti-Polarizing news says the opposite.

HYPOTHESIS 3 (Motivated reasoning and second guesses):

- *Follow Message is larger for Pro-Party than for Anti-Party news.*
- *Follow Message is larger for polarizing than for anti-polarizing news.*

D. Data and Experiment Details

The experiment was conducted in June 2018 on Amazon's Mechanical Turk (MTurk) platform. MTurk is an online labor marketplace in which participants choose "Human Intelligence Tasks" to complete. MTurk has become a popular way to run economic experiments (e.g., Horton, Rand, and Zeckhauser 2011; Kuziemko et al. 2015), and Levay, Freese, and Druckman (2016) find that participants tend to have more diverse demographics than students in university laboratories with respect to politics. The experiment was coded using oTree, an open-source software based on the Django web application framework developed by Chen, Schonger, and Wickens (2016).

The study was offered to MTurk workers currently living in the United States. 1,387 subjects were recruited and answered at least one question, and 1,300 subjects

completed the study. Of these subjects, 987 (76 percent) passed simple attention and comprehension checks, and the rest are dropped from the analyses.[20]

When asked for their party ratings, 627 subjects (64 percent) give a higher rating to the Democratic Party; 270 (27 percent) give a higher rating to the Republican Party; and 90 (9 percent) give the same rating to each party.[21] These subjects are labeled as "Pro-Dem," "Pro-Rep," and "Neutral," respectively, and for most analyses I drop the neutral subjects.

Two-thirds of subjects were randomly assigned to not receive a prior about the veracity of the news source, and one-third of subjects were told that True News and Fake News were equally likely. Independently, half of the subjects were randomly assigned to receive a WTP question and half were asked for their second guesses.

Each subject answered 13 questions; there are a total of 11,661 guesses to questions for the 897 nonneutral subjects. There are 11,443 news assessments. The discrepancy between these numbers is due to 143 subjects in the WTP group who did not receive a message in one round, and due to 75 (0.7 percent) correct guesses.[22] I drop these observations for news assessments, leaving 7,902 observations on politicized topics, 891 on the performance topic, and 2,650 on neutral topics.

The balance table for the Pro-Party/Anti-Party treatment is in Appendix B. The treatment is well-balanced, and overall shares of pro-party and Anti-Party news are nearly identical, suggesting that there was no differential attrition by treatment.

In order to validate the main results, I ran a replication experiment one year later. As discussed in the online Appendix, the replication confirms the results.[23]

III. Main Results

A. *Raw Data*

This subsection shows that the raw data support the main hypotheses, and the following subsection shows the relevant regressions. To validate that these questions are politicized, Appendix Table 4 compares initial guesses and finds systematic differences in median beliefs by party consistent with the directions predicted in Table 1.

In support of the first part of Hypothesis 1, I show that subjects trust the Pro-Party news more than they trust the Anti-Party news. Subjects' average assessment that the likelihood of Pro-Party news being from the True News source is 9.1 percentage

---

[20] The online Appendix shows that main results are robust to inclusion of these subjects. In order to pass these checks, subjects needed to perfectly answer the attention check question in Appendix A by giving a correct answer, correct bounds, and answering the news assessment with certainty. In addition, some questions had maximum and minimum possible answers (e.g., percentages, between 0 and 100). Subjects were dropped if any of their answers did not lie within these bounds.

[21] Levay, Freese, and Druckman (2016) also find that subjects on MTurk are mostly Democratic.

[22] The low frequency of correct guesses indicates that subjects were generally not looking up answers. It also suggests that the model's assumption of an atomless belief distribution is reasonable.

[23] The replication was identical in structure, but added several new political questions and did not include neutral questions, so it is not easy to pool results. All main hypotheses on political questions were preregistered, as were results on polarization and overprecision.
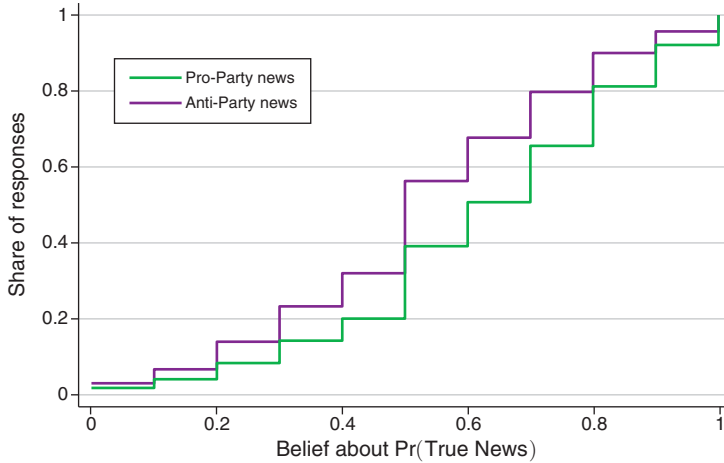
FIGURE 1. CDF OF ASSESSMENTS OF PRO-PARTY AND ANTI-PARTY NEWS

Notes: Pro-Party and Anti-Party news are defined in Table 1. This figure shows that subjects trust Pro-Party news more than Anti-Party news. The *x*-axis measures subjects' beliefs about Pr(*True News|Pro-/Anti-Party News*). The *y*-axis measures the share of respondents that give at most that high of an assessment. Bayesians would have the same trust in news for Pro-Party and Anti-Party news, and the residual is motivated reasoning.

points higher than their average assessment that Anti-Party news is from the True News source (s.e. 0.6 pp; $p < 10^{-40}$). Figure 1 shows the CDF of assessments for Pro-Party and Anti-Party news; the Pro-Party distribution first-order stochastically dominates the Anti-Party distribution. Appendix Figure 6 shows the same result for pro-performance and anti-performance news assessments.

Figure 2 shows the subject-demeaned assessments by news direction (Pro-Party; Anti-Party; news on neutral topics) and subject type (partisan and moderate, as defined by the absolute difference in party ratings). In support of all three parts of Hypothesis 1, subjects believe that Pro-Party news is more likely than news on neutral topics to be true, Anti-Party news is less likely than news on neutral topics to be true, and differences are larger for partisans. If neutral topics account for unmotivated mistakes in news assessments, then this figure shows that there is motivated reasoning in both positive and negative directions.

In support of Hypothesis 2, subjects trust the Fake News more than they trust the True News. Subjects' average assessment that the likelihood of True News is actually from the True News source is 6.0 percentage points lower than their average assessment that Fake News is from the True News source (s.e. 0.6 pp; $p < 10^{-20}$). Appendix Figure 7 shows the CDF of assessments for True News and Fake News on the political questions; the Fake News distribution first-order stochastically dominates the True News distribution.

This effect is not solely driven by whether the news is Pro-Party versus Anti-Party. Figure 3 shows the subject-demeaned assessments by news direction (Pro-Party; Anti-Party; neutral) and news veracity (True News; Fake News). Subjects give
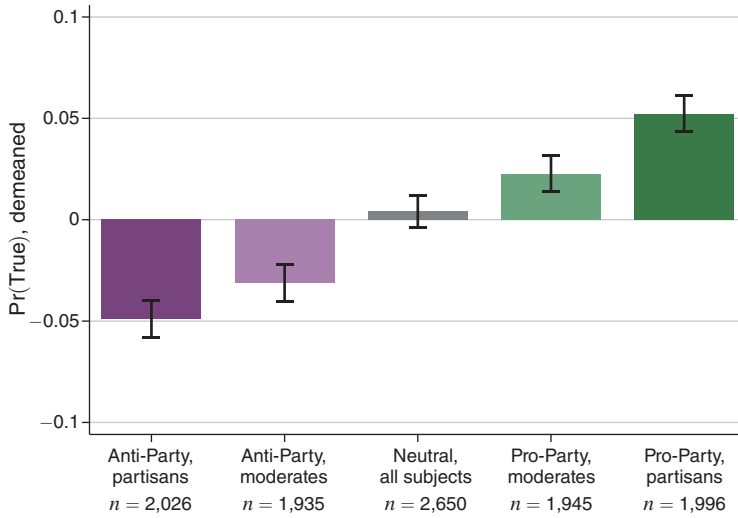
FIGURE 2. POLITICALLY MOTIVATED REASONING: NEWS DIRECTION AND PARTISANSHIP

Notes: The *y*-axis is motivated reasoning: stated $\Pr(True)$, demeaned at the subject level. News on partisan topics is classified as Pro-Party (Anti-Party) if it is more (less) representative of the subject's preferred political party, as defined in Table 1. A subject who is above the median value for abs($Republican\ Party\ rating - Democratic\ Party\ rating$) is classified as partisan; a subject who is not is classified as moderate. Error bars correspond to 95 percent confidence intervals.
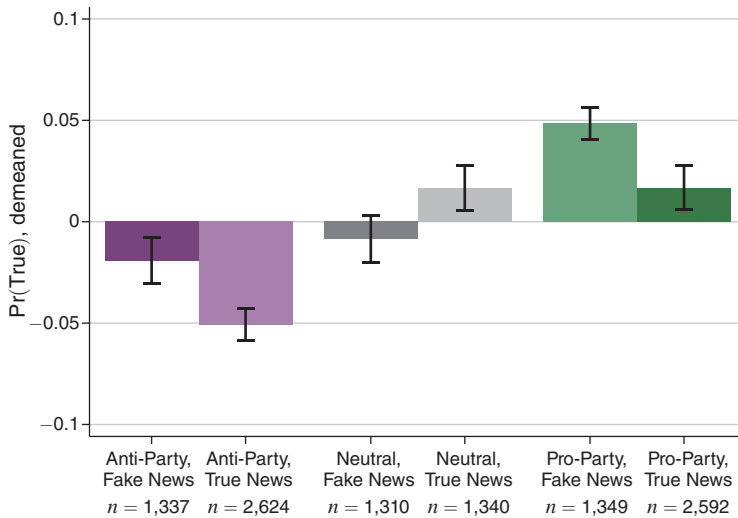
FIGURE 3. MOTIVATED REASONING AND ASSESSMENTS OF FAKE NEWS

Notes: The *y*-axis is motivated reasoning: stated $\Pr(True)$, demeaned at the subject level. News on partisan topics is classified as Pro-Party (Anti-Party) if it is more (less) representative of the subject's preferred political party, as defined in Table 1. Fake News sends messages that reinforce the direction of subjects' error; True News sends messages that mitigate subjects' error. Error bars correspond to 95 percent confidence intervals.

higher assessments to Fake News than to True News on politicized topics, but they do not do so on neutral topics.[24]

B. *Regression Specifications for News Assessments*

The primary regression specification is within subject.[25] It regresses assessments a on Pro-Party news for subject i, question topic q, and round r with fixed effects for i, q, and r, restricting to news that is Pro-Party or Anti-Party:[26]

$$a_{iqr} = \alpha + \beta \cdot \mathbf{1}\{Pro\text{-}Party\}_{iqr} + \gamma FE_i + \delta FE_q + \zeta FE_r + \epsilon_{iqr}.$$

This specification is used in Table 2, column 2. A similar alternative specification, replacing individual-level fixed effects with demographic controls is in column 1.

Hypothesis 1 claims that the Pro-Party/Anti-Party gap is increasing in partisanship, so column 3 interacts partisanship (the absolute difference in party ratings) with Pro-Party news. It also claims that motivated reasoning leads to both higher assessments for Pro-Party news and lower assessments for Anti-Party news; as such, column 4 includes indicators for both Pro-Party (versus neutral) news and Anti-Party (versus neutral) news. Hypothesis 2 posits that subjects will trust Fake News more than True News on politicized topics, so columns 5 and 6 regress assessments on a dummy for True News, controlling for and not controlling for Pro-Party news.

Hypotheses 1 and 2 are strongly supported. Assessments for Pro-Party news are higher than for Anti-Party news, and this effect increases in partisanship. There is evidence for motivated reasoning both towards Pro-Party and away from Anti-Party news, and Fake News assessments are higher than True News assessments.

Next, we look at each topic individually by regressing on the interaction of treatment and topic dummies. Figure 4 shows evidence for politically motivated reasoning on eight of the nine hypothesized topics ($p < 0.001$ on each). It also shows that people motivatedly reason towards believing they outperformed others ($p < 0.001$).

C. *Changing Guesses and Belief Polarization*

To further validate the "trust in news" measure of motivated reasoning, I also show that subjects engage in asymmetric updating about the questions themselves. Half of the subjects give a second guess to the initial question after receiving news, and I hypothesize that subjects are more likely to update in the Pro-Party direction than in the Anti-Party direction. This test is both useful as a robustness check as well as to better understand belief polarization on these questions.

---

[24] If anything, assessments are higher for True News than Fake News on neutral topics. Reflecting on the question may lead subjects to adjust *toward* the truth in the absence of motivated beliefs.

[25] At least one piece of Pro-Party and at least one piece of Anti-Party news were received by 99.4 percent of nonneutral subjects. Three subjects randomly never receive Pro-Party news and two never receive Anti-Party news.

[26] This model is slightly different than the one in Section II, as it uses fixed effects and a model that is linear in probabilities. These differences do not affect results; for instance, the online Appendix shows that results are qualitatively the same if we replace a_{iqr} with $\text{logit}(a_{iqr})$.

TABLE 2—MOTIVATED REASONING AND PERCEIVED TRUTHFULNESS OF NEWS

	(1)	(2)	(3)	(4)	(5)	(6)
Pro-Party News	0.092	0.088	0.041	0.037		0.077
	(0.006)	(0.006)	(0.012)	(0.006)		(0.006)
Partisanship × *Pro-Party*			0.099			
			(0.022)			
Anti-Party News				−0.048		
				(0.007)		
True News					−0.059	−0.034
					(0.006)	(0.006)
Question FE	Yes	Yes	Yes	No	Yes	Yes
Round FE	Yes	Yes	Yes	Yes	Yes	Yes
Subject FE	No	Yes	Yes	Yes	Yes	Yes
Subject controls	Yes	No	No	No	No	No
Neutral news	No	No	No	Yes	No	No
Observations	7,902	7,902	7,902	10,552	7,902	7,902
R^2	0.05	0.25	0.25	0.21	0.23	0.25
Mean	0.574	0.574	0.574	0.575	0.574	0.574

Notes: OLS, errors clustered at subject level. Neutral news indicates that Pro-Party/Anti-Party news assessments are compared to assessments on neutral topics. These classifications are in Table 1. Controls: race, gender, log(income), years of education, religion, and whether state voted for Trump or Clinton in 2016. Partisanship is abs(*Republican Party rating* − *Democratic Party rating*). In column 3, "Partisanship" is also interacted with round and "Question FE." Standard errors in parentheses.
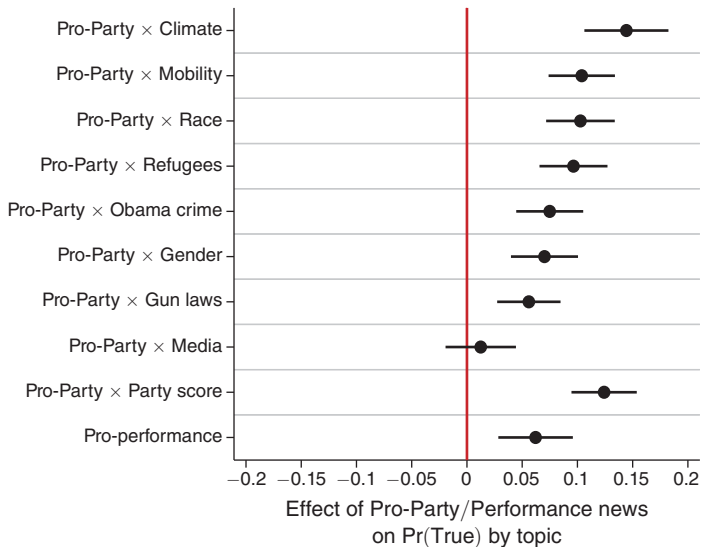
FIGURE 4. MOTIVATED REASONING BY TOPIC

Notes: OLS regression coefficients, errors clustered at subject level. FE included for subject, round number, and topic. Pro-Party (versus Anti-Party) news is defined in Table 1. Error bars correspond to 95 percent confidence intervals.

Column 1 of Table 3 shows that subjects are more likely to update their median in the direction of Pro-Party messages than they are from Anti-Party messages. Column 2 shows that on politicized topics, subjects are also more likely to change

TABLE 3—MOTIVATED REASONING AND FOLLOWING THE MESSAGE SENT

	(1)	(2)	(3)	(4)	(5)	(6)
Pro-Party news	0.122		0.115	0.018		0.021
	(0.019)		(0.020)	(0.017)		(0.018)
Polarizing news		0.060	0.034		−0.008	−0.013
		(0.018)	(0.018)		(0.016)	(0.016)
Pr(True)				1.126	1.136	1.128
				(0.059)	(0.058)	(0.059)
Question FE	Yes	Yes	Yes	Yes	Yes	Yes
Round FE	Yes	Yes	Yes	Yes	Yes	Yes
Subject FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,085	4,085	4,085	4,085	4,085	4,085
R^2	0.28	0.28	0.28	0.45	0.45	0.45
Mean	0.658	0.658	0.658	0.658	0.658	0.658

Notes: OLS, errors clustered at subject level. Only subjects from the second-guess group. Only Pro-Party/Anti-Party news observations, as defined in Table 1. *Polarizing news* is news that tells subjects that, compared to their initial guess, the answer is in the opposite direction from the population mean. Dependent variable is 1 if subjects change their guess upward when the message says "Greater Than" or downward when the message says "Less Than," −1 if they change their guess in the opposite direction, and 0 if they do not change their guess. Standard errors in parentheses.

their guesses in the direction of a polarizing message (one that tells them their guess is farther away from the mean) than from an anti-polarizing message.

Columns 4–6 of Table 3 show that discrepancies in both motivated posteriors and belief polarization can be explained by differences in news assessments. After controlling for assessments, guess changes are not statistically significantly affected by Pro-Party/Anti-Party messages, nor are they statistically significantly affected by polarizing messages, and the point estimates are close to zero. Evidence suggests that subjects change their beliefs to follow Pro-Party news more exactly because they trust that news source more.

More broadly, these results give a stark prediction about how people change their beliefs. They show that, in environments where signals remind people of their motivated beliefs, informational content is not a necessary condition for belief polarization.

D. *Alternative Explanations and Robustness Checks*

There are features in this design that may lead subjects to behave in ways that are consistent with motivated reasoning but are also consistent with unmotivated confounding hypotheses. This subsection discusses a number of potential confounds, and shows how it is possible to test these alternative hypotheses. I argue that it is unlikely that these confounds can fully explain the results above, strengthening the interpretation that motivated reasoning is what is being identified in the experiment.

Misunderstanding Medians and Skewed Belief Distributions.—It is reasonable to expect that subjects do not fully understand the concept of a median. For instance, they may answer with their mean belief instead. This would not directionally impact the news assessment results in a systematic direction, unless the prior distribution

were notably skewed. We can use where the initial guess μ_q lies in subjects' confidence intervals as a proxy for skewness, and see that the main results hold for subjects whose distributions are unskewed.

On the politicized questions, 32 percent of subjects' guesses are exactly halfway between their upper and lower bounds. Appendix B uses the same specification as the main regression but interacts Pro-Party news, Anti-Party news, and True News with a dummy for having "unskewed" priors. The treatment effects are qualitatively and quantitatively similar, indicating that skewness does not directionally affect results.

The Independence of News Sources.—We have so far assumed that subjects treat news sources as being drawn from independent distributions. While subjects are explicitly told to do this in the instructions, it is useful to show that they are not using previous pieces of news to update about current pieces of news.

In Appendix B, I modify the main regression table to account for the relative number of Pro-Party and Anti-Party news in previous rounds. The effect of previous rounds' Pro-Party and Anti-Party news have no effect on the current round's assessment, and the main treatment effects are unchanged, suggesting that subjects indeed treat news sources as independent.

Misunderstanding "Fake News".—Using the terminology "Fake News" can lead to greater subject engagement, but may also evoke other commonly used definitions (such as fake articles, as in Allcott and Gentzkow 2017). This subsection shows that results cannot be explained by two plausible misinterpretations of Fake News.

First, suppose that subjects believe that messages from Fake News are equally likely to send true and false messages, instead of *always* sending false messages. In this experiment, no predictions about assessments would change. A Bayesian would still have an ex ante prior that Pro-Party and Anti-Party messages are equally likely, and would not infer anything about $\Pr(True)$ given either message. A motivated reasoner who is motivated to believe that the answer is large would still infer that $\Pr(True|Pro\text{-}Party) > \Pr(True|Anti\text{-}Party)$. The same argument can be made if subjects believe Fake News randomly sends false messages with any interior probability.

A more complicated situation involves subjects who believe that messages from Fake News are actually from a news source that is biased against their party. That is, suppose that subjects believe that Fake News is more likely to report Anti-Party news given a Pro-Party truth than Pro-Party news given an Anti-Party truth.

To test this, we can again look at how subjects change their guesses. In particular, suppose that subjects were Bayesian but used this asymmetrically wrong definition of Fake News. Then, they would find Pro-Party "Fake News" messages to be more informative than Anti-Party "Fake News" messages, since "Fake News" is expected to usually send the Anti-Party message. (The quotes here indicate that these subjects are using the wrong definition.) Such subjects would then update more from Pro-Party than Anti-Party news, *conditional on their assessment of* $\Pr(True\ News)$. In Section IIIC we see that subjects are similarly likely to update from Pro-Party and Anti-Party news after controlling for their assessments. While the data are too

imprecise to rule out the existence of subjects who treat Fake News as biased, this story is insufficient for explaining the main effects.

Incorrect Initial Guesses.—While it can sometimes be in subjects' best interests to strategically misreport their median in order to earn more points on news assessment questions, I find no evidence that subjects do this.[27] In Round 1 of the experiment, subjects do not yet know that they will be seeing a news assessment page. If subjects were strategically mis-guessing to earn more assessment points, they would perform worse in Round 1 than in subsequent rounds on assessments and better in Round 1 than in subsequent rounds on guesses.

There are no statistically significant differences in assessment scores in Round 1. Subjects score 67.2 points (s.e. 0.9) in Round 1 and 66.4 points (s.e. 0.3) in Rounds 2–12; the difference is 0.8 points (s.e. 1.0; $p = 0.383$).[28] There are also no statistically significant differences in guess scores in Round 1. Subjects score 76.2 points (s.e. 1.0) in Round 1 and 75.9 points (s.e. 0.2) in Rounds 2–12; the difference is 0.3 points (s.e. 1.0; $p = 0.757$).

Non-strategic forms of incorrect initial guesses are more complicated to rule out. If there is symmetric noise in guesses, such that the probability that a subject is equally likely to state her Q quantile and her $1 - Q$ quantile for $Q \neq 1/2$, then the main results do not change. Results are also not consistent with subjects biasing their initial guesses towards the population mean. While such behavior can explain why subjects trust error-reinforcing news more than error-mitigating news on politicized, and why they trust Pro-Party news more than Anti-Party news, it incorrectly predicts the same pattern on neutral topics.

The one form of misreporting that can be consistent with both Bayesian updating and results from the experiment involves subjects systematically misreporting initial guesses in a way that is biased in the *opposite* direction from their party. One potential reason for such a bias is that subjects do not sufficiently think about the question; and, given more time, they update towards their actual (more Pro-Party) belief. It is possible that seeing the second screen causes subjects to think harder about the original question, and thinking harder leads to more Pro-Party beliefs.[29] However, thinking harder does not seem to be asymmetric, as I find no evidence that subjects prefer to spend more time thinking about good news.[30]

Expressive Preferences.—Bursztyn et al. (2020) provides recent evidence showing that people in experiments may forgo payment in order to make political statements. In this experiment, if subjects have a preference for stating Pro-Party signals, then

---

[27] For instance, if a subject believes that $\Theta \sim U[0, 20]$, then she will expect to earn more points on the news assessment question by guessing 20 than by guessing 10 since she is confident the correct answer is not greater than 20. She would expect to earn 90 points on the question and 100 points on news, as opposed to earning 95 points on the question and 75 points on news.

[28] I exclude scoring on Rounds 13–14 since the questions are not randomly assigned in those rounds; the result is similar if they are included. I also exclude scoring on comprehension check questions.

[29] The psychology behind this explanation overlaps with this theory of motivated reasoning, as the second page evokes the motive, and further work could better elucidate what qualifies as a signal.

[30] The mean time spent on the assessment page with Pro-Party news is 14.6 seconds (s.e. 0.3 seconds), and the mean time spent on the assessment page with Anti-Party news is 14.8 seconds (s.e. 0.3 seconds).

both their initial guesses and their news assessments will be biased in the Pro-Party direction, consistent with the data. However, if they are Bayesian, how they *change* their guesses will not be directional, since they have already stated their preferred belief.

In Table 3, subjects are more likely to update their guesses in the Pro-Party direction than in the Anti-Party direction, even though they are equally likely to receive Pro-Party and Anti-Party news. This is consistent with subjects sincerely trusting the Pro-Party news more; it is not consistent with expressive Bayesian updating.

Other Robustness Tests.—As discussed in Section IID, subjects either were told that $\Pr(True\ News) = 1/2$ in the instructions or were not told this. In the online Appendix, I restrict the regressions from Section IIIB to subjects in each randomization arm, and find similar results in the two arms. Likewise, seeing the WTP page or having a second guess has little effect.

It is possible that subjects learn over the course of the experiment that they motivatedly reason and debias themselves. I interact the main effect with dummies for each round number, and do not find evidence for this. In each round, subjects give larger assessments to Pro-Party than to Anti-Party news.

E. *Heterogeneity in Motivated Reasoning*

There are two types of heterogeneity to consider: heterogeneity in the direction of motivated reasoning, and heterogeneity in its magnitude.

First, we consider the direction of heterogeneity. To do this, Figure 5 shows the coefficients from the regression of news assessments on the interaction of the political direction of the news (pro-Rep versus pro-Dem) and party preferences, as well as on binarized observable demographics. Nonpolitical demographics are race, gender, income, age, education, whether the subject's state voted for Trump or Clinton in 2016, and religious affiliation.

After controlling for party preference, none of the other demographics have a statistically significant effect on the direction of motivated reasoning. All coefficients are between plus and minus 0.03, a magnitude less than one-third than for politics.

In the online Appendix, we consider the magnitude of motivated reasoning, acknowledging that this design is unable to disentangle magnitude of bias and strength of motive. Partisans have a larger treatment effect than moderates, but demographics besides partisanship do not notably affect the magnitude of the bias; all effects are between $+/-0.03$.[31]

These results suggest that the magnitude of bias of motivated reasoning is broadly similar across demographics and that while the direction of motivated

---

[31] There is suggestive evidence that pro-Dem subjects motivatedly reason more than pro-Rep subjects, though this may simply be due to the non-representativeness, conditional on party, of the MTurk sample. For instance, only 76 percent of Republicans in this sample approved of President Trump's performance; in a Gallup poll conducted contemporaneously (from May 28–July 1), 88 percent of Republicans approved of his performance (Gallup 2018).
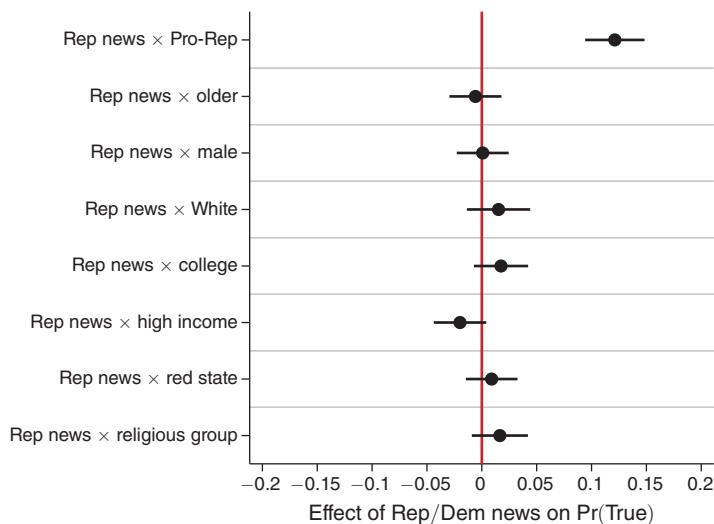
FIGURE 5. HETEROGENEITY IN THE PARTISAN DIRECTION OF MOTIVATED REASONING

Notes: This figure plots the relative treatment effect of seeing pro-Rep versus pro-Dem news on subjects' news assessments by binary demographics. These are OLS regression coefficients, errors clustered at subject level. FE included for subject, round number, and topic. Only Pro-Party/Anti-Party news observations, as defined in Table 1. pro-Rep: higher rating for Republican than Democratic party. Older: above the median age in the experiment. High income: above median income in the experiment. Red state: state voted for Trump in 2016. Religious: subject affiliates with any religion.

beliefs is heterogeneous by party, it is not very different across nonpolitical demographics.

F. *Motivated Reasoning, Initial Beliefs, and Overprecision*

We now consider two other relationships between initial beliefs and motivated reasoning: how much motivated reasoning can explain people's differences in beliefs, and how the bias relates to overprecise confidence intervals.

First, the data show that variation in this experiment's measure of motivated beliefs can explain a sizable fraction of variation in actual beliefs about these questions. I look at the relationship between motives and beliefs by correlating answers to politicized questions with differences in assessments between pro-Rep and pro-Dem news. For each politicized question, subjects' initial guesses are winsorized (at the 5 percent level), normalized, and signed; positive numbers correspond to more pro-Rep. Next, for each subject, these normalized guesses are averaged (and renormalized) to give a measure of how pro-Rep her beliefs are. I correlate this value with the normalized average difference between pro-Rep news assessments and pro-Dem news assessments.

Using R^2, variation in news assessments explains 13 percent of the variation in beliefs. By comparison, all of the nonpolitical demographics collected in this experiment—age, gender, race, education, logged income, whether one is religious, and whether one is from a state that voted for Trump or Clinton in 2016—explain

7 percent of the variance in beliefs.[32] This suggests that my measure of motivated reasoning performs at least as well as oft-discussed demographics at predicting beliefs.

Second, motivated reasoning can help explain a particular form of overprecision due to miscalibrated confidence intervals. In particular, we posit that motivated reasoners formed directionally biased belief distributions, thereby leading them to overestimate the probability that the answers are within their confidence interval. Such a story would imply that overprecision is stronger for politicized than for neutral topics, and stronger for partisans than for moderates:

HYPOTHESIS 4 (Overprecision and partisanship):

- *On politicized and performance questions, subjects' 50 percent confidence intervals contain the correct answer less than 50 percent of the time.*
- *On politicized questions, the likelihood that subjects' confidence intervals contain the correct answer decreases in their partisanship.*

In support of Hypothesis 4, subjects are overprecise in their beliefs about questions that evoke motivated beliefs. On politicized topics, subjects' confidence intervals contain the correct answer 46.6 percent of the time (s.e. 0.6 percent); this is statistically significantly less than 50 percent ($p < 0.001$). Overprecision on these topics is primarily driven by partisans, whose intervals contain the correct answer 44.2 percent of the time (s.e. 0.9 percent). Moderates' intervals contain the correct answer 48.8 percent of the time (s.e. 0.8 percent). Partisans' overprecision is statistically significantly larger than moderates' ($p < 0.001$). On the performance question, subjects' confidence intervals contain the correct answer 42.0 percent of the time (s.e. 1.6 percent), which is statistically significantly less than 50 percent ($p < 0.001$).

Overprecision is stronger for motivated questions than neutral questions, suggesting that these results are not simply driven by a bias toward overly narrow confidence intervals. On the "Random Number" question, which asks subjects to guess what a random number drawn uniformly from 0 to 100 will be, confidence intervals contain the correct answer 54.6 percent of the time, which indicates mild *underprecision*.[33]

Finally, overprecision is correlated with over-trusting error-reinforcing Fake News, consistent with the notion that current beliefs are reflective of motivated beliefs. Subjects who are overprecise on a question assess Fake News to be 3.3 pp more likely to be truthful (s.e. 0.8 pp; $p < 0.001$) compared to underprecise subjects, and subjects who are overprecise on a question assess True News to be 2.4 pp less likely to be truthful (s.e. 0.7 pp; $p = 0.001$) compared to underprecise subjects.

G. *Discussion*

The experimental results strongly favor the hypothesis of motivated reasoning with politically motivated beliefs over the null hypothesis of Bayesian updating.

---

[32] These are unadjusted values. Adjusted R^2 is 13 percent for assessments and 6 percent for demographics.
[33] On the other neutral questions, subjects also exhibit moderate underprecision.

Motivated reasoning also fits the data better than unmotivated patterns of over- or under-inference. Subjects significantly overtrust Pro-Party news and "Fake News" in an environment with uninformative signals, real monetary stakes, and little room for self-deception.

These results show that political motives are important to consider, in addition to prior beliefs, when predicting how people respond to new information. In fact, motivated reasoning can partly explain one why people update further in the direction of their prior than a Bayesian would. When priors reflect motivated beliefs, a method that detects prior-confirming biases in belief updating may be detecting motivated reasoning instead (e.g., Eil and Rao 2011).

The results in Section IIIC relate to the effect of motivated reasoning on political polarization. Not only do subjects polarize in beliefs about the veracity of news, they polarize in their beliefs about the questions themselves, despite receiving uninformative signals. Gentzkow and Shapiro (2011) find only modest differences in the media that liberals and conservatives consume, and motivated reasoning can help explain why people polarize even if they consume similar media outlets. Results also suggest that the beliefs people are motivated by are not necessarily optimistic ones, but are rather ones that support their political identity. In follow-up work, I show that absent politics, "good news" about the world plays a negligible role in motivated reasoning (Thaler 2024a). Furthermore, these results indicate that beliefs people find attractive are *even further apart* than their current beliefs. One reason that people do not already hold their most-preferred beliefs is that motivated reasoners are still affected by information; the amount of distortion in updating is constrained by actual informational content. Motivated reasoners who receive precise signals would in fact become *less* polarized.

IV. Conclusion

This paper has shown that distortion of new information in favorable directions—motivated reasoning—plays a substantial role in people's assessment of the veracity of news and helps explain why people form inaccurate and polarized beliefs about the world around them. It developed a novel experimental paradigm that is able to identify the channel of motivated reasoning from Bayesian updating and other channels across a variety of settings. Results demonstrated the role that *politicall*y motivated reasoning plays in how people form beliefs about applied economic issues like income mobility, crime, and immigration. They also showed how this bias can lead to further belief polarization, overprecision, and an excess trust in Fake News.

There are several avenues for future work with this design. For instance, the design can be used by applied researchers who are interested in detecting motivated reasoning. Researchers, working in many different contexts, can plug in any factual question with a real-valued answer to determine how people motivatedly reason about that question. Motives can be compared across varied contexts and populations. Conceptually, this design could be used to further unpack the relationship between motives and belief-based utility, a relationship that is presently ambiguous in the literature (Barron 2021; Thaler 2024a).

The notion of a susceptibility parameter also suggests a lever for debiasing people. This design enables us to identify and estimate the magnitude of motivated reasoning; interventions whose objective is debiasing can then use this estimate to test the efficacy of the treatment. One approach is to estimate susceptibility for a treatment group and a control group. Having an approach to reducing motivated reasoning would be a valuable way for researchers to combat polarization and biased beliefs about important issues, especially in a highly politicized society.

Appendix A. Study Materials: Exact Question Wordings

Crime Under Obama

Some people believe that the Obama administration was too soft on crime and that violent crime increased during his presidency, while others believe that President Obama's pushes towards criminal justice reform and reducing incarceration did not increase violent crime.

This question asks how murder and manslaughter rates changed during the Obama administration. In 2008 (before Obama became president), the murder and manslaughter rate was 54 per million Americans.

In 2016 (at the end of Obama's presidency), what was the per-million murder and manslaughter rate?

Correct answer: 53.

Source linked on results page: http://bit.ly/us-crime-rate

Upward Mobility

In 2017, Donald Trump signed into law the largest tax reform bill since Ronald Reagan's 1981 and 1986 bills. Some people believe that Reagan's reforms accelerated economic growth and allowed lower-income Americans to reap the benefits of lower taxes, while other people believe that this decreased the government's spending to help lower-income Americans get ahead.

This question asks whether children who grew up in low-income families during Reagan's tenure were able to benefit from his tax reforms.

Of Americans who were born in the lowest-income (bottom 20 percent) families from 1980–1985, what percent rose out of the lowest-income group as adults?

(Please guess between 0 and 100.)

Correct answer: 64.9.

Source linked on results page: http://bit.ly/us-upward-mobility. (*page* 47)

Racial Discrimination

In the United States, white Americans have higher salaries than black Americans on average. Some people attribute these differences in income to differences in education, training, and culture, while others attribute them more to racial discrimination.

In a study, researchers sent fictitious resumes to respond to thousands of help-wanted ads in newspapers. The resumes sent had identical skills and education, but the researchers gave half of the (fake) applicants stereotypically White names such as Emily Walsh and Greg Baker, and gave the other half of the

applicants stereotypically Black names such as Lakisha Washington and Jamal Jones.

9.65 percent of the applicants with White-sounding names received a call back. What percent of the applicants with Black-sounding names received a call back?

(Please guess between 0 and 100.)

Correct answer: 6.45.

Source linked on results page: http://bit.ly/labor-market-discrimination.

Gender and Math GPA

In the United States, men are more likely to enter into mathematics and math-related fields. Some people attribute this to gender differences in interest in or ability in math, while others attribute it to other factors like gender discrimination.

This question asks whether high school boys and girls differ substantially in how well they do in math classes. A major testing service analyzed data on high school seniors and compared the average GPA for male and female students in various subjects.

Male students averaged a 3.04 GPA (out of 4.00) in math classes. What GPA did female students average in math classes?

(Please guess between 0.00 and 4.00.)

Correct answer: 3.15.

Source linked on results page: http://bit.ly/gender-hs-gpa

Refugees and Violent Crime

Some people believe that the United States has a responsibility to accept refugees into the country, while others believe that an open-doors refugee policy will be taken advantage of by criminals and put Americans at risk.

In 2015, German leader Angela Merkel announced an open-doors policy that allowed all Syrian refugees who had entered Europe to take up residence in Germany. From 2015–17, nearly one million Syrians moved to Germany. This question asks about the effect of Germany's open-doors refugee policy on violent crime rates.

In 2014 (before the influx of refugees), the violent crime rate in Germany was 224.0 per hundred-thousand people.

In 2017 (after the entrance of refugees), what was the violent crime rate in Germany per hundred-thousand people?

Correct answer: 228.2.

Sources linked on results page: Main site: http://bit.ly/germany-crime-main-site. *2014 and 2015 data:* http://bit.ly/germany-crime-2014-2015. *2016 and 2017 data:* http://bit.ly/germany-crime-2016-2017.

Climate Change

Some people believe that there is a scientific consensus that human activity is causing global warming and that we should have stricter environmental regulations, while others believe that scientists are not in agreement about the existence or cause of global warming and think that stricter environmental regulations will sacrifice jobs without much environmental gain.

This question asks about whether most scientists think that global warming is caused by humans. A major nonpartisan polling company surveyed thousands of scientists about the existence and cause of global warming.

What percent of these scientists believed that "Climate change is mostly due to human activity"?

(Please guess between 0 and 100.)

Correct answer: 87.

Source linked on results page: http://bit.ly/scientists-climate-change.

Gun Reform

The United States has a homicide rate that is much higher than other wealthy countries. Some people attribute this to the prevalence of guns and favor stricter gun laws, while others believe that stricter gun laws will limit Americans' Second Amendment rights without reducing homicides very much.

After a mass shooting in 1996, Australia passed a massive gun control law called the National Firearms Agreement (NFA). The law illegalized, bought back, and destroyed almost one million firearms by 1997, mandated that all non-destroyed firearms be registered, and required a lengthy waiting period for firearm sales.

Democrats and Republicans have each pointed to the NFA as evidence for/ against stricter gun laws. This question asks about the effect of the NFA on the homicide rate in Australia.

In the five years before the NFA (1991-1996), there were 319.8 homicides per year in Australia. In the five years after the NFA (1998-2003), how many homicides were there per year in Australia?

Correct answer: 318.6.

Sources linked on results page: http://bit.ly/australia-homicide-rate (*Suicides declined substantially, however. For details:* http://bit.ly/impact-australia-gun-laws.)

Media Bias

Some people believe that the media is unfairly biased towards Democrats, while some believe it is balanced, and others believe it is biased towards Republicans.

This question asks whether journalists are more likely to be Democrats than Republicans.

A representative sample of journalists were asked about their party affiliation. Of those either affiliated with either the Democratic or Republican Party, what percent of journalists are Republicans?

(Please guess between 0 and 100.)

Correct answer: 19.8.

Source linked on results page: http://bit.ly/journalist-political-affiliation

Democrats' Relative Performance

This question asks whether you think Democrats or Republicans did better on this study about political and US knowledge. I've compared the average points scored by Democrats and Republicans among 100 participants (not including yourself).

The Republicans scored 70.83 points on average.

How many points do you think the Democrats scored on average?

(Please guess between 0 and 100)
Correct answer: 72.44.

Republicans' Relative Performance

This question asks whether you think Democrats or Republicans did better on this study about political and US knowledge. I've compared the average points scored by Democrats and Republicans among 100 participants (not including yourself).

The Democrats scored 72.44 points on average.

How many points do you think the Republicans scored on average?

(Please guess between 0 and 100)

Correct answer: 70.83.

Own Relative Performance

How well do you think you performed on this study about political and US knowledge? I've compared the average points you scored for all questions (prior to this one) to that of 100 other participants.

How many of the 100 do you think you scored higher than?

(Please guess between 0 and 100.)

Correct answer: Depends on participant's performance.

Random Number

A computer will randomly generate a number between 0 and 100. What number do you think the computer chose?

(As a reminder, it is in your best interest to guess an answer that is close to the computer's choice, even if you don't perfectly guess it.)

Correct answer: Randomly generated for each participant.

Latitude of Center of the United States

The US National Geodetic Survey approximated the geographic center of the continental United States. (This excludes Alaska and Hawaii, and US territories.)

How many degrees North is this geographic center?

(Please guess between 0 and 90. The continental US lies in the Northern Hemisphere, the Equator is 0 degrees North, and the North Pole is 90 degrees North.)

Correct answer: 39.833.

Source linked on results page: http://bit.ly/center-of-the-us

Longitude of Center of the United States

The US National Geodetic Survey approximated the geographic center of the continental United States. (This excludes Alaska and Hawaii, and US territories.)

How many degrees West is this geographic center?

(Please guess between 0 and 180. The continental US lies in the Western Hemisphere, which ranges from 0 degrees West to 180 degrees West.)

Correct answer: 98.583.

Source linked on results page: http://bit.ly/center-of-the-us

Comprehension Check: Current Year

In 1776 our fathers brought forth, upon this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.

What is the year right now?

This is not a trick question and the first sentence is irrelevant; this is a comprehension check to make sure you are paying attention. For this question, your lower and upper bounds should be equal to your guess if you know what year it currently is.

Correct answer: 2018.

Source linked on results page: http://bit.ly/what-year-is-it

APPENDIX B. ADDITIONAL EXPERIMENTAL RESULTS

TABLE 4—PRIOR BELIEFS BY PARTY

	Pro-Rep	Pro-Dem	Difference	Answer
Obama crime guess	55.907	49.560	6.348	53
	(0.765)	(0.391)	(0.858)	
Mobility guess	30.185	22.152	8.034	64.9
	(1.048)	(0.611)	(1.211)	
Race guess	12.349	8.051	4.298	6.45
	(0.874)	(0.436)	(0.975)	
Gender guess	3.059	3.086	−0.027	3.15
	(0.015)	(0.008)	(0.017)	
Refugees guess	287.640	239.004	48.637	228.2
	(5.894)	(2.353)	(6.335)	
Climate guess	75.226	85.366	−10.140	87
	(1.056)	(0.572)	(1.200)	
Gun laws guess	230.013	184.478	45.535	318.6
	(5.950)	(3.914)	(7.113)	
Media guess	36.656	41.850	−5.195	19.8
	(1.211)	(0.599)	(1.349)	
Rep. Score guess	71.563	61.933	9.630	70.83
	(0.787)	(0.614)	(0.997)	
Dem. Score guess	64.671	73.277	−8.606	72.44
	(0.771)	(0.415)	(0.875)	
Observations	2,430	5,643	8,073	

Notes: OLS, robust standard errors. Guesses are winsorized at the 5 percent level. Third column represents mean pro-Rep guess minus mean pro-Dem guess. The sign of every coefficient points in the predicted motive direction from Table 1. Standard errors in parentheses.

TABLE 5—BALANCE TABLE

	Anti-Party news	Pro-Party news	Anti versus pro	p-value
Partisanship	0.484	0.478	0.007	0.312
	(0.005)	(0.005)	(0.007)	
Rep. versus Dem.	−0.237	−0.236	−0.001	0.937
	(0.008)	(0.008)	(0.011)	
Male	0.532	0.534	−0.002	0.881
	(0.008)	(0.008)	(0.011)	
Age	35.261	35.400	−0.139	0.573
	(0.175)	(0.173)	(0.246)	
Education	14.716	14.765	−0.049	0.242
	(0.029)	(0.030)	(0.042)	
log($income$)	10.725	10.748	−0.024	0.182
	(0.012)	(0.013)	(0.018)	
White	0.752	0.760	−0.008	0.430
	(0.007)	(0.007)	(0.010)	
Black	0.075	0.081	−0.006	0.303
	(0.004)	(0.004)	(0.006)	
Hispanic	0.066	0.062	0.004	0.499
	(0.004)	(0.004)	(0.006)	
Religious	0.443	0.457	−0.014	0.214
	(0.008)	(0.008)	(0.011)	
Red State	0.567	0.558	0.009	0.431
	(0.008)	(0.008)	(0.011)	
WTP elicited	0.490	0.476	0.014	0.213
	(0.008)	(0.008)	(0.011)	
Told 1/2 true	0.333	0.344	−0.011	0.309
	(0.007)	(0.008)	(0.011)	
N	3,961	3,941	7,902	

Notes: Standard errors in parentheses. Rep. versus Dem. is the rating of the Republican party minus the rating of the Democratic party and is between −1 and 1. Partisanship is the absolute difference in these ratings. Education is in years. Religious is 1 if subject in any religious group. Red State is 1 if state voted for Trump in 2016 election. WTP elicited is 1 if subject in the WTP group and 0 if in the second-guess group. Told 1/2 True is 1 if subject is told that Pr($True\ News$) is 1/2 and 0 if subject is not.

Table 6—The Effect of News Direction, Actual Veracity, and Skewed Prior Distributions on Perceived Veracity

	(1)	(2)	(3)	(4)
Unskewed	−0.010	−0.011	−0.011	0.007
	(0.009)	(0.009)	(0.010)	(0.013)
Pro-Party News	0.083	0.037	0.028	0.075
	(0.007)	(0.014)	(0.008)	(0.008)
Unskewed × Pro-Party	0.016	0.015	0.018	0.007
	(0.012)	(0.019)	(0.013)	(0.013)
Partisanship × Pro-Party		0.097		
		(0.024)		
Unskewed × Partisanship × Pro-Party		0.006		
		(0.032)		
Anti-Party news			−0.052	
			(0.008)	
Unskewed × Anti-Party			0.001	
			(0.013)	
True News				−0.026
				(0.007)
Unskewed × True News				−0.026
				(0.013)
Question FE	Yes	Yes	No	Yes
Round FE	Yes	Yes	Yes	Yes
Subject FE	Yes	Yes	Yes	Yes
Neutral news	No	No	Yes	No
Observations	7,882	7,882	10,499	7,882
R^2	0.25	0.25	0.21	0.25
Mean	0.574	0.574	0.575	0.574

Notes: OLS, errors clustered at subject level. "Neutral news" indicates that Pro-Party/Anti-Party compared to neutral news, as defined in Table 1. Controls: race, gender, log(income), education (in years), religion, whether state voted for Trump or Clinton in 2016. Partisanship is absolute difference between Republican and Democratic ratings. In column 2, partisanship is also interacted with round and "Question FE". Unskewed is 1 if initial guess is exactly halfway between lower/upper bounds and 0 otherwise. Observations dropped if lower bound equals upper bound. Standard errors in parentheses.

TABLE 7—THE EFFECT OF NEWS DIRECTION, ACTUAL VERACITY, AND PREVIOUS NEWS DIRECTIONS ON PERCEIVED VERACITY

	(1)	(2)	(3)	(4)
Previous Pro-Party	−0.001	−0.001	−0.002	−0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Pro-Party news	0.087	0.040	0.036	0.076
	(0.006)	(0.012)	(0.006)	(0.007)
Partisanship × Pro-Party		0.099		
		(0.022)		
Anti-Party news			−0.048	
			(0.007)	
True News				−0.034
				(0.006)
Question FE	Yes	Yes	No	Yes
Round FE	Yes	Yes	Yes	Yes
Subject FE	Yes	Yes	Yes	Yes
Neutral news	No	No	Yes	No
Observations	7,902	7,902	10,552	7,902
R^2	0.25	0.25	0.21	0.25
Mean	0.574	0.574	0.575	0.574

Notes: OLS, errors clustered at subject level. Neutral news indicates that Pro-Party/Anti-Party compared to "Neutral news," as defined in Table 1. Controls: race, gender, log(income), education (in years), religion, whether state voted for Trump or Clinton in 2016. Partisanship is the absolute difference between Republican and Democratic ratings. In column 2, partisanship is also interacted with round and "Question FE". "Previous Pro-Party" is the number of all previous pieces of news that are Pro-Party minus the number that are anti-party. Standard errors in parentheses.
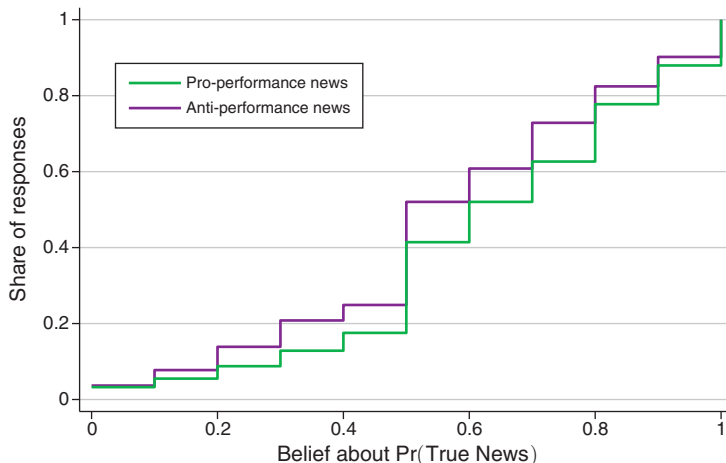

FIGURE 6. CDF OF PERCEIVED VERACITY FOR PRO-PERFORMANCE AND ANTI-PERFORMANCE NEWS

Notes: Pro-performance and anti-performance news are defined in Table 1. This figure shows that subjects trust Pro-performance news more than anti-performance news. The *x*-axis measures subjects' beliefs about $\Pr(True \mid Pro\text{-}/Anti\text{-}Performance\ News)$. The *y*-axis measures the share of respondents that give at most that high of an assessment. Bayesians would have the same trust in news for pro-performance and anti-performance news, and the residual is motivated reasoning.
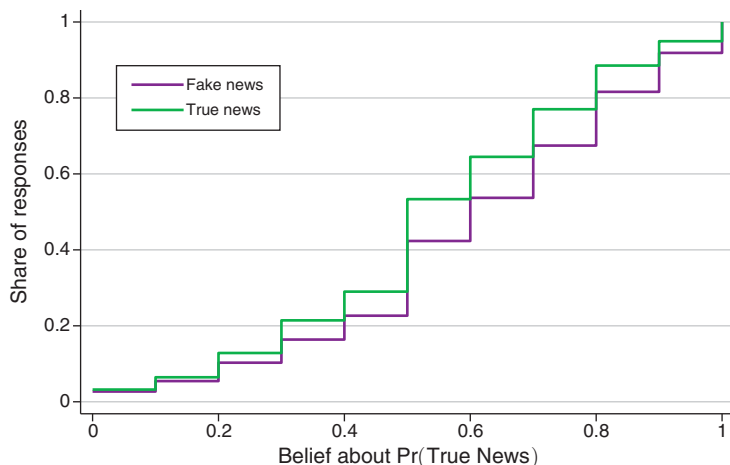
FIGURE 7. CDF OF PERCEIVED VERACITY FOR TRUE NEWS AND FAKE NEWS

Notes: Only political topics included. This figure shows that subjects trust Fake News more than True News. The *x*-axis measures subjects' beliefs about $\Pr(True\,|\,actual\;True/Fake\,News)$. The *y*-axis measures the share of respondents that give at most that high of an assessment. Bayesians would have the same trust in news for True and Fake News, and the residual is motivated reasoning.

REFERENCES

Akerlof, George A., and William T. Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." American Economic Review 72 (3): 307–19.

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva. 2023. "Immigration and Redistribution." Review of Economic Studies 90 (1): 1–39.

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." American Economic Review 108 (2): 521–54.

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." Journal of Economic Perspectives 31 (2): 211–36.

Allcott, Hunt, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang. 2020. "Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic." Journal of Public Economics 191: 104254.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy. 2018. "Incentives in Experiments: A Theoretical Analysis." Journal of Political Economy 126 (4): 1472–1503.

Babcock, Linda, and George Loewenstein. 1997. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." Journal of Economic Perspectives 11 (1): 109–26.

Babcock, Linda, George Loewenstein, Samuel Issacharoff, and Colin Camerer. 1995. "Biased Judgments of Fairness in Bargaining." American Economic Review 85 (5): 1337–43.

Barron, Kai. 2021. "Belief Updating: Does the 'Good-News, Bad-News' Asymmetry Extend to Purely Financial Domains?" Experimental Economics 24: 31–58.

Bénabou, Roland, and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." Quarterly Journal of Economics 117 (3): 871–915.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." Quarterly Journal of Economics 126 (2): 805–55.

Benjamin, Daniel. 2019. "Errors in Probabilistic Reasoning and Judgment Biases." In Handbook of Behavioral Economics, Vol. 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 69–186. Amsterdam: Elsevier.

Brunnermeier, Markus K., and Jonathan A. Parker. 2005. "Optimal Expectations." American Economic Review 95 (4): 1092–1118.

Bursztyn, Leonardo, Michael Callen, Bruno Ferman, Saad Gulzar, Ali Hasanain, and Noam Yuchtman. 2020. "Political Identity: Experimental Evidence on Anti-Americanism in Pakistan." Journal of the European Economic Association 18 (5): 2532–60.

Buser, Thomas, Leonie Gerhards, and Joël van der Weele. 2018. "Responsiveness to Feedback as a Personal Trait." Journal of Risk and Uncertainty 56: 165–92.

Cappelen, Alexander W., Ingar K.Haaland, and Bertil Tungodden. 2018. "Beliefs about Behavioral Responses to Taxation." Unpublished.

Carrillo, Juan D., and Thomas Mariotti. 2000. "Strategic Ignorance as a Self-Disciplining Device." Review of Economic Studies 67 (3): 529–44.

Charness, Gary, and Chetan Dave. 2017. "Confirmation Bias with Motivated Beliefs." Games and Economic Behavior 104: 1–23.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree – An Open-Source Platform for Laboratory, Online, and Field Experiments." Journal of Behavioral and Experimental Finance 9: 88–97.

Chew, Soo Hong, Wei Huang, and Xiaojian Zhao. 2020. "Motivated False Memory." Journal of Political Economy 128 (10): 3913–3939.

Coutts, Alexander. 2019. "Good News and Bad News Are Still News: Experimental Evidence on Belief Updating." Experimental Economics 22 (2): 369–95.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." Economic Theory 33: 67–80.

Druckman, James N., Erik Peterson, and Rune Slothuus. 2013. "How Elite Partisan Polarization Affects Public Opinion Formation." American Political Science Review 107 (1): 57–79.

Druckman, James N., Matthew S. Levendusky, and Audrey McLain. 2018. "No Need to Watch: How the Effects of Partisan Media Can Spread via Interpersonal Discussions." American Journal of Political Science 62 (1): 99–112.

Eil, David, and Justin M. Rao. 2011. "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself." American Economic Journal: Microeconomics 3 (2): 114–38.

Ertac, Seda. 2011. "Does Self-Relevance Affect Information Processing? Experimental Evidence on the Response to Performance and Non-performance Feedback." Journal of Economic Behavior and Organization 80 (3): 532–45.

Exley, Christine L. 2016. "Excusing Selfishness in Charitable Giving: The Role of Risk." Review of Economic Studies 83 (2): 587–628.

Exley, Christine L., and Judd B. Kessler. 2022. "Motivated Errors." Unpublished.

Eyting, Markus. 2022. "Why Do We Discriminate? The Role of Motivated Reasoning." Unpublished.

Festinger, Leon. 1957. A Theory of Cognitive Dissonance. Redwood City, CA: Stanford University Press.

Flynn, D. J., Brendan Nyhan, and Jason Reifler. 2017. "The Nature and Origins of Misperceptions: Understanding False and Unsupporting Beliefs About Politics." Advances in Political Psychology 38 (S1): 127–50.

Gallup. 2018. "Presidential Job Approval Center." https://news.gallup.com/interactives/507569/presidential-job-approval-center.aspx (accessed March 2, 2024).

Gentzkow, Matthew, and Jesse M. Shapiro. 2006. "Media Bias and Reputation." Journal of Political Economy 114 (2): 280–316.

Gentzkow, Matthew, and Jesse M. Shapiro. 2011. "Ideological Segregation Online and Offline." Quarterly Journal of Economics 126 (4): 1799–1839.

Gentzkow, Matthew, Michael B. Wong, and Allen T. Zhang. 2018. "Ideological Bias and Trust in Information Sources." Unpublished.

Gerber, Alan S., and Gregory A. Huber. 2009. "Partisanship and Economic Behavior: Do Partisan Differences in Economic Forecasts Predict Real Economic Behavior?" American Political Science Review 103 (3): 407–26.

Gino, Francesca, Michael I. Norton, and Roberto A. Weber. 2016. "Motivated Bayesians: Feeling Moral While Acting Egoistically." Journal of Economic Perspectives 30 (3): 189–212.

Grether, David M. 1980. "Bayes Rule as a Descriptive Model: The Representativeness Heuristic." Quarterly Journal of Economics 95 (3): 537–57.

Haaland, Ingar, and Christopher Roth. 2020. "Labor Market Concerns and Support for Immigration." Journal of Public Economics 191: 104256.

Haaland, Ingar, and Christopher Roth. 2023. "Beliefs about Racial Discrimination and Support for Pro-Black Policies." Review of Economics and Statistics 105 (1): 40–53.

Haisley, Emily C., and Roberto A. Weber. 2010. "Self-Serving Interpretations of Ambiguity in Other-Regarding Behavior." Games and Economic Behavior 68 (2): 614–25.

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The Online Laboratory: Conducting Experiments in a Real Labor Market." Experimental Economics 14: 399–425.

Kahan, Dan M. 2016. "The Politically Motivated Reasoning Paradigm, Part 1: What Politically Motivated Reasoning Is and How to Measure It." In Emerging Trends in Social and Behavioral Sciences, edited by Robert Scott and Stephen Kosslyn, 1–16. New York, NY: John Wiley and Sons.

Kahan, Dan M., David A. Hoffman, Donald Braman, Danieli Evans, and Jeffrey J. Rachlinski. 2012. "They Saw a Protest: Cognitive Illiberalism and the Speech-Conduct Distinction." Stanford Law Review 64 (4): 851–906.

Kuhnen, Camelia M. 2014. "Asymmetric Learning from Financial Information." Journal of Finance 70 (5): 2029–62.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." Psychological Bulletin 108 (3): 480–98.

Kunda, Ziva, and Lisa Sinclair. 2000. "Motivated Stereotyping of Women: She's Fine If She Praised Me but Incompetent If She Criticized Me." Personality and Social Psychology Bulletin 26 (11): 1329–42.

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." American Economic Review 105 (4): 1478–1508.

Levay, Kevin E., Jeremy Freese, and James N. Druckman. 2016. "The Demographic and Political Composition of Mechanical Turk Samples." SAGE Open 6 (1).

Levendusky, Matthew S. 2013. "Why Do Partisan Media Polarize Viewers?" American Journal of Political Science 57 (3): 611–23.

Lord, Charles G., Lee Ross, and Mark R. Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." Journal of Personality and Social Psychology 37 (11): 2098–2109.

Mayraz, Guy. 2019. "Priors and Desires — A Bayesian Model of Wishful Thinking and Cognitive Dissonance." Unpublished.

Meeuwis, Maarten, Jonathan A. Parker, Antoinette Schoar, and Duncan Simester. 2022. "Belief Disagreement and Portfolio Choice." Journal of Finance 77 (6): 3191–3247.

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya Rosenblat. 2022. "Managing Self-Confidence: Theory and Experimental Evidence." Management Science 68 (11): 7793–7817.

Moore, D. A., and P. J. Healy. 2008. "The Trouble with Overconfidence." Psychological Review 115 (2): 502–17.

Moore, Don A., Elizabeth R. Tenney, and Uriel Haran. 2015. "Overprecision in Judgment." In Wiley Blackwell Handbook of Judgment and Decision Making, edited by Gideon Keren and George Wu, 182–209. New York, NY: John Wiley and Sons.

Nisbet, Erik C., Kathryn E. Cooper, and R. Kelly Garrett. 2015. "The Partisan Brain: How Dissonant Science Messages Lead Conservatives and Liberals to (Dis)Trust Science." ANNALS of the American Academy of Political and Social Science 658 (1): 36–66.

Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." Political Behavior 32: 303–30.

Nyhan, Brendan, and Jason Reifler. 2013. Which Corrections Work? Research Results and Practice Recommendations. Washington, DC: New America Foundation.

Nyhan, Brendan, Jason Reifler, and Peter A. Ubel. 2013. "The Hazards of Correcting Myths about Health Care Reform." Medical Care 51 (2): 127–32.

Ortoleva, Pietro, and Erik Snowberg. 2015. "Overconfidence in Political Behavior." American Economic Review 105 (2): 504–35.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey. 2013. "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease." American Economic Review 103 (2): 804–30.

Pennycook, Gordon, and David G. Rand. 2019. "Lazy, Not Biased: Susceptibility to Partisan Fake News is Better Explained by Lack of Reasoning Than by Motivated Reasoning." Cognition 188: 39–50.

Rabin, Matthew, and Joel L. Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." Quarterly Journal of Economics 144 (1): 37–82.

Sarsons, Heather. 2017. "Interpreting Signals in the Labor Market: Evidence from Medical Referrals." Unpublished.

Schwardmann, Peter, Egon Tripodi, and Joël van der Weele. 2022. "Self-Persuasion: Evidence from International Debating Competitions." American Economic Review 112 (4): 1118–46.

Sunstein, Cass R., Sebastian Bobadilla-Suarez, Stephanie C. Lazzaro, and Tali Sharot. 2017. "How People Update Beliefs about Climate Change: Good News and Bad News." Cornell Law Review 102 (6): 1431–1443.

Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." American Journal of Political Science 50 (3): 755–69.

Tappin, Ben M., Gordon Pennycook, and David G. Rand. 2020a. "Bayesian or Biased? Analytic Thinking and Political Belief Updating." Cognition 204: 104375.

Tappin, Ben M., Gordon Pennycook, and David G. Rand. 2020b. "Thinking Clearly about Causal Inferences of Politically Motivated Reasoning: Why Paradigmatic Study Designs Often Undermine Causal Inference." Current Opinion in Behavioral Science 34: 81–87.

Tetlock, Philip E. 1983. "Accountability and the Perseverance of First Impressions." Social Psychology Quarterly 46 (4): 285–92.

Thaler, Michael. 2019. "Debiasing Motivated Reasoning Through Learning: Evidence from an Online Experiment (Registration)." AEA RCT Registry. https://doi.org/10.1257/rct.4401.

Thaler, Michael. 2021. "Gender Differences in Motivated Reasoning." Journal of Economic Behavior and Organization 191: 501–18.

Thaler, Michael. 2023. "The Supply of Motivated Beliefs." Unpublished.

Thaler, Michael. 2024a. "Good News is Not a Sufficient Condition for Motivated Reasoning." CESifo Working Paper No. 10915.

Thaler, Michael. 2024b. "Replication data for: The Fake News Effect: Experimentally Identifying Motivated Reasoning Using Trust in News." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E183845V1.

Zimmermann, Florian. 2020. "The Dynamics of Motivated Beliefs." American Economic Review 110 (2): 337–61.