(Politically) Motivated Information Acquisition: A Minimal Formal Framework*

Egor Bronnikov

[*Check the latest version of the paper.*]

January, 2025

Abstract

Individuals with different identity-relevant beliefs often make divergent decisions even when presented with identical objective facts. This paper advances the theoretical understanding of motivated reasoning by developing a formal model that contrasts motivated information acquisition with a standard framework. We first explore how different types of decision makers—self- and unaware of their directional motives as well as information-sensitive motivated reasoners—make information acquisition choices based on their beliefs. Subsequently, we integrate the concept of separate decision utility to reevaluate our results in the context of actions. Finally, we derive several (experimentally) testable hypotheses based on the predictions of the formal model. Our paper contributes to the theoretical literature on motivated reasoning and information acquisition and connects to propaganda.

JEL classification: C60, C70, D83, D91

Keywords: Bayesian updating, motivated reasoning, information acquisition, (political) beliefs.

*I am grateful to Andrew T. Little, Carlo M. Horz, Elias Tsakas, Sebastian J. Goerg and Alexander Vostroknutov for their insightful discussions and fruitful comments. Bronnikov: School of Business and Economics, Maastricht University; Economics Department, University of Amsterdam; Mercatus Center, George Mason University; School of Science, Brīvā Universitāte; egor.bronnikov@maastrichtuniversity.nl.

1 Introduction

Several recent events—such as the COVID-19 pandemic or the full-scale still-ongoing Russian invasion of Ukraine—have exemplified the urgency and significance of gaining a much deeper understanding of behavioral patterns in information processing, information choices, and information incorporation.

Given an identical set of (even objective) facts to choose from, individuals with different identity-relevant—such as political, religious, and other self-related—beliefs will likely arrive at different decisions. A rapidly growing body of empirical research (discussed below) has been enriching the political science and economics literature with increasingly nuanced and detailed insights into various aspects of motivated reasoning (e.g., Bénabou, 2015; Bénabou and Tirole, 2016; Gino et al., 2016; Golman et al., 2016; Loewenstein and Molnar, 2018; Druckman and McGrath, 2019; Thaler, 2021; Molnar and Loewenstein, 2022; Zimmermann, 2020; Drobner, 2022; Gonzalez-Jimenez, 2022; Oprea and Yuksel, 2022; Burro and Castagnetti, 2022; Little, 2022a; Little et al., 2022; Sprengholz et al., 2023; Engelmann et al., 2024; Stoetzer and Zimmermann, 2024; Thaler, 2024a). However, from a theoretical perspective, the formal literature on motivated reasoning is still in its infancy.

While a perfectly standard approach to information processing serves as a valuable benchmark, the development of behavioral formal models that account for empirical regularities has become crucially important. A deeper understanding of how individuals manage information (and its complexity) is one of the key topics in the current literature in both political science (e.g., Marshall, 2019; Peterson and Iyengar, 2021; Ryan and Aziz, 2021; Little, 2022a,b) and economics (e.g., Charness et al., 2021; Enke et al., 2023a; Banovetz and Oprea, 2023; Enke, 2024).

In this paper, building on the theoretical framework introduced by Little (2022a), we develop a (minimalistic) formal model of motivated information acquisition, contrasting it with standard belief updating.

To do this, we consider a decision maker (DM [he/him/his]) who forms beliefs about the state of the world (or realizations of a random variable w) and has directional motives. When given the opportunity to observe feedback (a signal), the DM can choose whether or not to observe it.

Turning to the results, we first demonstrate that, for the DM with objective beliefs and directional motives, there is a unique solution to the belief updating problem (Theorem 1).

Second, we explore the DM's decision to acquire the signal in light of beliefs only. We start by analyzing the conditions under which the DM's—being a self-aware motivated reasoner—will choose to acquire information. For such a reasoner, obtaining information incurs a cost and the threshold for acquiring information increases with the DM's sensitivity (or degree of inclination) to directional motives (α) and decreases with the precision of the signal (λ). In other words, more precise signals are *worse* for the DM, as they make it more difficult to maintain the desired belief (Theorem 2).

Next, we investigate the requirements for a DM—who is an unaware motivated reasoner—to acquire information. For such a reasoner, there is no threshold for acquiring information: as long as the overall benefit of receiving information is not negative, they will always choose to do so (Theorem 3).

Proceeding, we show that as soon as the DM is sensitive to information, the condition for making an information acquisition choice will be identical, regardless of whether they are self-aware or unaware of their directional motives. For the information-sensitive motivated reasoner, the threshold for acquiring information increases with the DM's degree of inclination toward directional motives and their sensitivity to the asymmetry of information (δ), and decreases with the signal precision (Theorem 4).

Third, we explore the DM's decision to acquire the signal in the context of subsequent actions. Introducing a standard from the literature—separate decision utility, which accounts for the loss incurred when the DM's estimate differs from the real state of the world—we reconsider all three theorems in light of these new conditions. By equating the overall benefit to the satisfactory condition of the minimum solution for the utilitarian decision of whether to acquire the signal, we derive three corollaries. To begin with, the DM who is self-aware of their directional motive will acquire a signal with a loss parameter that increases with the extent of their motivated reasoning (Corollary 2). Next, the DM who is unaware of their directional motives will always acquire a signal (Corollary 3). Finally, the DM who is information-sensitive, regardless of whether they are self-aware or unaware of their directional motives, will acquire a signal with a loss parameter that increases with their inclination toward motivated reasoning and sensitivity to the type of information (Corollary 4).

Finally, in line with the insights provided by the formal model, we generate a range of hypotheses that can be tested empirically. These hypotheses are directly informed by the model's structure and allow for experimental validation of its theoretical implications.

From a more technical perspective, this paper builds on the theoretical framework introduced by Little (2022a) and establishes behavioral conditions for (motivated) information acquisition. Most importantly, we formalize the conditions for information acquisition across three types of decision-makers: those who are self-aware of their motivated reasoning, those who lack such awareness, and those who are sensitive to information. Additionally, we extend the solutions to key problems in Little (2022a), such as the DM's belief updating process and the (motivated) belief utility function $\widetilde{B}(f)$, from a finite to a continuous space.

Our paper contributes to several strands of the literature. First, we contribute to the theoretical literature on motivated reasoning (e.g., Gerber and Green, 1999; Rabin and Schrag, 1999; Bénabou and Tirole, 2002; Bracha and Brown, 2012; Bénabou, 2015; Kimbrough and Vostroknutov, 2022; Little, 2022a; Melnikoff and Strohminger, 2024). We extend the understanding of how directional motives influence belief updating by providing a formal model that differentiates between self-aware and unaware motivated reasoners. Our results reveal that even when the DM's beliefs are influenced by motivated reasoning, the conditions for acquiring information align consistently with the extent of their directional motives, thus deepening insights into how identity-relevant beliefs shape decision-making.

Second, we contribute to the theoretical literature on information acquisition (e.g., Barlevy and Veronesi, 2000; Persico, 2000; Bergemann and Välimäki, 2002; Gerardi and Yariv, 2008; Van Nieuwerburgh and Veldkamp, 2010; Liu, 2011; Colombo et al., 2014; Yang, 2015; Argenziano et al., 2016). Our results specify the behavioral conditions under which a DM with motivated reasoning decides to acquire information. By contrasting the decision to acquire information with belief updating and the associated costs, we provide a nuanced view of how self-awareness, unawareness of one's biases, and information sensitivity affect this decision, offering valuable insights into the behavioral aspects of information acquisition.

Finally, we connect to the theoretical literature within political economy, particularly in the context of propaganda (e.g., Gehlbach and Sonin, 2014; Gehlbach et al., 2016; Little, 2017; Barrera et al., 2020; Guriev and Treisman, 2020; Carter and Carter, 2023; Egorov and Sonin, 2024), with a special focus on behavioral models of it (e.g., Horz, 2021, 2023). We advance the understanding of how motivated reasoning affects choices of information—which is crucial to the procession of propaganda—by modeling how directional motives influence the decision to acquire information in a way that aligns with behavioral models of propaganda. Our paper enhances the comprehension of how different types of motivated reasoners interact with propaganda and how they are influenced by varying degrees of signal precision and information sensitivity.

The rest of the paper is organized as follows. Section 2 introduces the preliminaries of the formal model and presents the results of both standard and motivated belief updating. In Section 3, we explore the conditions under which the DM will (or will not) choose to acquire the signal based on the beliefs themselves. We begin by introducing the concept of belief utility and then consider three types of DM: the motivated reasoner who is self-aware of their directional motives (Subsection 3.1), the motivated reasoner who is unaware of their directional motives (Subsection 3.2), and the motivated reasoner who is sensitive to the asymmetry of information (Subsection 3.3). In Section 4, we investigate the conditions under which the DM will (or will not) choose to acquire the signal based on the actions. After presenting the formal model, we discuss the assumptions and advantages of the chosen approach (Subsection 5.1), formulate a series of testable hypotheses that directly stem from the formal model (Subsection 5.2), and offer further discussion on relevant ideas concerning prospective experimental design. Finally, Section 6 concludes the paper by highlighting several directions for further research. Appendices accompany the main body of the paper and provide the reader with proofs for all propositions, theorems, and corollaries presented in the paper (Appendix A), examples for the main theorems in Section 3 (Appendix B), and comparative statics graphs for Sections 3 and 4 (Appendices C and D, respectively).

2 Belief Updating

Consider a continuous random variable w that can take on values from the entire real line, $\Omega = (-\infty, \infty)$. Denote by $\mathscr{P}$ the collection of all probability density functions defined over this interval.

Now consider a decision-maker (DM) who forms beliefs about the random variable w. The DM begins by establishing an initial belief, represented by a probability density function $f(w)$ such that $\int_{-\infty}^{\infty} f(w)\,dw = 1$. This function reflects the DM's initial subjective distribution over possible values of w. Given the opportunity to observe a signal, the DM faces a decision on whether to acquire or not acquire the additional information.

2.1 Standard Updating

Initially, let us assume that the prior belief about the random variable w is represented by a normal distribution with mean μ_0 and variance σ_0^2, that is $w \sim \mathcal{N}(\mu_0, \sigma_0^2)$. The signal, denoted by s, which may be observed, is expressed as $s = w + \epsilon$, where ϵ follows a normal distribution with mean 0 and variance σ_ϵ^2, denoted as $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. According to standard Bayesian theory, if the DM observes the signal s, the posterior distribution of w_0 given s is also normal. The result of standard belief updating is described in Remark 1.

Remark 1. *Suppose the objective belief follows a normal distribution $w \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Suppose also that there a signal $s = w + \epsilon$, where ϵ follows a normal distribution such that $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Then the posterior belief is $w_s \sim \mathcal{N}\left(\lambda s + (1-\lambda)\mu_0, (1-\lambda)\sigma_0^2\right)$ where $\lambda = \sigma_\epsilon^{-2}/(\sigma_\epsilon^{-2} + \sigma_0^{-2})$.*

Proof. See the proof of this remark in the Appendix (p. 33). ***Q.E.D.***

The posterior mean is a weighted average of the observed signal s and the prior mean μ_0. The weight λ reflects the relative precision of the signal compared to the prior belief. Specifically, λ increases with the precision of the signal (i.e., as σ_ϵ^{-2} becomes large relative to σ_0^{-2}) and decreases with the precision of the prior (i.e., as σ_0^{-2} becomes large relative to σ_ϵ^{-2}).

The posterior variance reflects the updated uncertainty about w after observing the signal. The posterior variance is reduced from the prior variance σ_0^2 based on the precision of the observed signal. The factor $1-\lambda$ represents the portion of the prior variance that is retained after incorporating the information from the signal.

Overall, more precise signals receive a higher weight in the posterior mean and contribute more to reducing the variance.

2.2 Motivated Updating

This subsection introduces the concept of belief updating in the context of motivated reasoning formally.

We will denote the objective density function as f and the density function reflecting a motivated belief—i.e., the belief formed as a result of motivated reasoning—as $\widetilde{f}$. Assume f_0 represents the density of the prior belief, and f_s denotes the density of the posterior belief. This differentiation allows for the analysis of how beliefs are altered due to cognitive bias(es).

Suppose there exists a *belief utility* function $B(f,f')$ that evaluates the utility of a belief f in relation to another belief f'. The belief utility function $B(f,f')$ is introduced to quantify how a belief f is valued in comparison to another belief f'. This function can incorporate various factors—that are outlined in the contemporary literature on motivated reasoning in both political science and economics[1]—such as alignment with personal goals (e.g., Gino et al., 2016; Loewenstein and Molnar, 2018; Molnar and Loewenstein, 2022), self-related (i.e., political, societal, religious, etc.) identity (e.g., Druckman and McGrath, 2019; Zimmermann, 2020; Drobner, 2022; Gonzalez-Jimenez, 2022; Oprea and Yuksel, 2022; Bronnikov, 2024b; Thaler, 2021, 2024a; Engelmann et al., 2024), consistency with prior world views (cognitive consonance) (e.g., Golman et al., 2016; Burro and Castagnetti, 2022; Little, 2022a; Little et al., 2022; Sprengholz et al., 2023; Bronnikov, 2024a), or emotional satisfaction (e.g., Bénabou, 2015; Bénabou and Tirole, 2016; Stoetzer and Zimmermann, 2024). The belief utility function $B(f,f')$ provides a framework for assessing how different belief states are evaluated in the decision-making process.

Additionally, consider that the DM may exhibit a tendency to either prioritize (or

[1]It is worth noting that while the stated categorization of various factors makes sense, it is rather difficult, if not impossible, to assign a particular paper—whether research or review—exclusively to one category. The majority of the cited papers refer to more than one category, and their assignment to a specific topic is to some extent arbitrary.

neglect) certain beliefs. In this context, forming motivated beliefs with density $\widetilde{f}$, the DM aims to maximize the belief utility function $B(f, f')$. Let $\mathscr{F}$ be the collection of all density functions on $\mathbb{R}$: this set encompasses all potential belief distributions that can be considered within the analysis.

$$\widetilde{f} = \arg\max_{f' \in \mathscr{F}} \left\{ B(f, f') \right\} \tag{1}$$

$$= \arg\max_{f' \in \mathscr{F}} \left\{ -D_{KL}(f' || f) + v(f') \right\} \tag{2}$$

where $D_{KL}(f'|f)$ is the Kullback-Leibler divergence (KLD) from f to f' (eq. 3), capturing the accuracy motives, and $v(f')$, reflecting the directional motives (eq. 4).

The KLD is a widely used metric for quantifying the difference between probability distributions[2] (eq. 3). From a dynamic perspective, the KLD is positive whenever $f'(w) \neq f(w)$ and it increases as the divergence between $f'(w)$ and $f(w)$ grows; conversely, when $f'(w) = f(w)$, the KLD equals zero. Although it cannot be claimed as omnipresent, the KLD is far from underrepresented in the literature on formal models related to belief updating[3].

$$D_{KL}(f' || f) = \int_{-\infty}^{\infty} f'(w) \log \left(\frac{f'(w)}{f(w)} \right) dw \tag{3}$$

The term $v(f')$ denotes the directional motive for the DM (eq. 4). If we permit the DM to prioritize certain aspects of their belief distribution, it becomes straightforward to identify subjects whose responses lack a Bayesian equivalent (Little, 2022a).

$$v(f') = \int_{-\infty}^{\infty} v(w) f'(w) dw \tag{4}$$

However, it may be more intuitive to assume that the DM derives satisfaction from believing that the state of the world takes on specific values, rather than from having a particular distribution over the states. For instance, a citizen might prefer to believe that the state of the world is either War or Peace, but it is less plausible for them to derive satisfaction from believing that the state of the world is War with a probability of 0.87.

Theorem 1. *For any objective belief f and directional motive v there is a unique solution*

[2] See, for instance, Reza (1994); Cover (1999) or, alternatively, *The Book of Statistical Proofs* by Soch et al. (2024) available online at https://statproofbook.github.io/D/kl.html.

[3] For a more thorough and broader discussion of KLD in the context of non-optimal beliefs and belief updating, see, e.g., Little (2022b) in political science, and Dominiak et al. (2021); Zhao (2022); Fudenberg et al. (2023); Pomatto et al. (2023) in economics.

to the DM's problem stated in eq. (1): $\widetilde{f}(w)$, where

$$\widetilde{f}(w) = \frac{f(w)\exp(v(w))}{\int_{-\infty}^{\infty} f(w')\exp(v(w))\,dw'}. \tag{5}$$

Proof. See the proof of this theorem in the Appendix (p. 35). ***Q.E.D.***

Suppose $v(\cdot)$ is linear in w, such that $v(w) = aw$, where $a \geqslant 0$ represents the DM's sensitivity to directional motives. In other words, the DM has a preference for believing that the state of the world is better (e.g., higher). A higher value of a indicates a stronger desire to believe the state is favorite, while a lower value of a reflects a weaker tendency to engage in motivated reasoning.

Corollary 1. *If the objective belief is normal with mean μ and variance σ^2, and if $v(\cdot)$ is linear in w: $v(w) = \alpha w$, then the motivated belief is distributed normally and can have mean $\mu + \frac{1}{2}\alpha\sigma^2$ and variance σ^2.*

Proof. See the proof of this corollary in the Appendix (p. 37). ***Q.E.D.***

Plugging the objective beliefs in for $\widetilde{f}$, the prior for motivated beliefs is given by

$$\widetilde{\mu}_0 = \mu_0 + \frac{1}{2}\alpha\sigma_0^2, \tag{6}$$

$$\widetilde{\sigma}_0^2 = \sigma_0^2, \tag{7}$$

and the posterior for motivated beliefs is given by

$$\widetilde{\mu}_s = \mu_s + \frac{1}{2}\alpha\sigma_s^2 \tag{8}$$

$$= \lambda s + (1-\lambda)\mu_0 + \frac{1}{2}\alpha(1-\lambda)\sigma_0^2, \tag{9}$$

$$\widetilde{\sigma}_s^2 = \sigma_s^2 \tag{10}$$

$$= (1-\lambda)\sigma_0^2, \tag{11}$$

where (as shown in subsection 2.1) $\lambda = \sigma_\epsilon^{-2}/(\sigma_\epsilon^{-2} + \sigma_0^{-2})$.

3 Choosing (Not) to Acquire the Signal Based on Beliefs

Assume that the DM, when deciding whether to acquire the signal, anticipates (either explicitly or implicitly) that the motivated belief will be formed as described above.

Firstly, suppose the DM decides to acquire the signal. In this case, the resulting utility is given by the belief utility with the posterior belief $f_s(\cdot)$, and there is an overall benefit (or cost) $b \in \mathbb{R}$ associated with acquiring the information. Now, suppose the DM chooses not to acquire the signal. Then the utility derived from the belief is given by the belief utility function $B(f, f')$ (Little, 2022a).

Consider a flattered DM holding a motivated belief $\widetilde{f}(\cdot)$, while their Bayesian belief is $f(\cdot)$. The belief utility associated with the motivated belief is then expressed as the value of the objective function evaluated at the optimal motivated belief.

Proposition 1. *The objective function at the optimal motivated belief is given by:*

$$\widetilde{B}(f) = B(f, f') \tag{12}$$

$$= \log\left(\int e^{v(w)} f(w) dw\right). \tag{13}$$

Proof. See the proof of this proposition in the Appendix (p. 37). ***Q.E.D.***

The belief utility represents the logarithm of the expected utility value of $e^{v(w)}$ with respect to the objective density $f(\cdot)$. If the objective belief $f(\cdot)$ is normally distributed with mean μ and variance σ^2, and the function $v(\cdot)$ is linear in w, such that $v(w) = \alpha w$, then $v(w)$ is a normally distributed random variable[4] with mean $\alpha\mu$ and variance $\alpha^2\sigma^2$.

Since $\widetilde{B}(f)$ is the logarithm of the moment-generating function of this random variable evaluated at $t = 1$, and since the moment-generating function of a normal variable[5] is $M(t, \mu, \sigma) = \exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)$, we have:

$$\widetilde{B}(f) = M(t = 1, \mu, \sigma) \tag{14}$$

$$= \alpha\mu + \frac{1}{2}\alpha^2\sigma^2. \tag{15}$$

Unsurprisingly, the belief utility increases with higher objective beliefs, as $a > 0$. Perhaps less obviously, the belief utility at the optimal motivated belief also increases with higher variance, as this provides the DM with more flexibility to believe whatever he prefers.

The utility for retaining the prior motivated belief and selecting the optimal motivated belief is given by:

$$\widetilde{B}(f_0) = \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2, \tag{16}$$

[4]One can easily see it: if $W \sim \mathcal{N}\left(\mu, \sigma^2\right)$ and $V(W) = \varphi W$, then $\varphi W \sim \mathcal{N}\left(\varphi\mu, \varphi^2\sigma^2\right)$ since $\mathbb{E}[V(W)] = \mathbb{E}[\varphi W] = \varphi\mathbb{E}[W] = \varphi\mu$ and $\mathbb{V}[V(W)] = \mathbb{V}[\varphi W] = \varphi^2\mathbb{V}[W] = \varphi^2\sigma^2$.

[5]This is a classical theorem: if a random variable X is following a normal distribution: $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then the moment-generating function of X is $M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$. One can refer to many great sources, for instance, Bulmer (1979); Kobayashi et al. (2011); Pitman (2012) or, alternatively, to *The Book of Statistical Proofs* by Soch et al. (2024) available online at https://statproofbook.github.io/P/norm-mgf.html.

where $f_0(\cdot)$ is the prior belief with mean μ_0 and variance σ_0^2.

Since the DM is not sure about the mean of their posterior belief, some uncertainty is involved. If the DM acquires the signal, the resulting belief utility, conditional on the revelation s, is:

$$\widetilde{B}(f_s) = \alpha\mu_s + \frac{1}{2}\alpha^2\sigma_s^2 \tag{17}$$

$$= \alpha\left(\lambda s + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2, \tag{18}$$

where $\lambda = \sigma_\epsilon^{-2}/(\sigma_\epsilon^{-2} + \sigma_0^{-2})$.

Since this expression (eq. 18) is linear in s, the expected belief utility across realizations of s can be written as:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda\mathbb{E}_s[s] + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2. \tag{19}$$

Ultimately—reasoning from the utilitarian perspective—the DM will acquire the signal if and only if:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] + b \geqslant \widetilde{B}(f_0). \tag{20}$$

One advantage of incorporating the benefit b into the model in such a simple way is that it allows to capture both the objective value of information (b^{obj}) and the psychological (subjective) value of information (b^{psy}) as a convex combination: $b = \gamma b^{\text{obj}} + (1-\gamma)b^{\text{psy}}$, where γ reflects the weight the decision-maker (DM) assigns to each aspect. This framework allows for three types of DM: (i) a fully standard agent (when $\gamma = 1$), who is solely concerned with the objective value of information and disregards any psychological effects, (ii) a DM who is driven entirely by the subjective value of information (when $\gamma = 0$), and (iii) a DM who values both the objective and psychological aspects of information, but neither exclusively ($\gamma \in (0,1)$).

Since $\mathbb{E}_s[s]$ plays a pivotal role in eq. (20), a crucial question arises: does the DM, when forming beliefs about the revelation of the signal, rely on the objective belief or the motivated belief? In other words, is the DM aware that their beliefs are motivated when deciding whether to acquire the information?

3.1 Self-aware Motivated Reasoner

Let us start by assuming that the DM is aware of their directional motive and uses the objective belief. Thus, we have $\mathbb{E}_s[s] = \mu_0$, and the expected utility of the belief with the posterior is given by:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda\mu_0 + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{21}$$

$$= \alpha\mu_0 + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2. \tag{22}$$

Compared to the belief utility with the prior (eq. 16), the first term in the expected belief utility with the posterior for the unaware motivated reasoner (eq. 22) remains the same, while the second term is reduced by a factor equal to the complement of the signal precision. This is simplified and formalized in Theorem 2.

Theorem 2. *The decision-maker, who is self-aware of his own directional motive, will acquire a signal s with an overall benefit b if and only if:*

$$b \geqslant \frac{1}{2}\alpha^2\lambda\sigma_0^2 > 0. \tag{23}$$

This threshold is strictly positive, and so information will never be acquired if costly $(b < 0)$.

Proof. See the proof of this theorem in the Appendix (p. 38). ***Q.E.D.***

The first feature of comparative statics is that acquiring information is costly because the DM expects their objective mean belief to remain the same as it was prior to acquiring the information[6]. However, the DM also incurs a cost for having a more precise belief, as it makes it harder to maintain the desired belief. In this sense, more precise information signals are *worse* because λ increases with the precision of the signal.

Another feature of comparative statics is that the threshold for information acquisition increases with α, meaning that more motivated subjects are less likely to seek information. This reflects the idea that stronger motivated reasoning leads to a lower propensity to acquire new information, as it could potentially contradict the subject's desired beliefs.

Overall, for the self-aware motivated reasoner, the threshold for acquiring information increases with the DM's extent of inclination toward directional motives (α) and decreases with the precision of the signal (λ).

3.2 Unaware Motivated Reasoner

Now suppose the DM is unaware of their directional motive when choosing whether to acquire information, meaning $\mathbb{E}_s[s] = \widetilde{\mu}_0$. In this case, the DM is motivated by the belief

[6] Remarkably, this is a standard property of conditional expectations and a defining property of a martingale: $\mathbb{E}[X_{t+1} \mid \mathscr{F}_t] = X_t \ \forall t$, where $\{X_t\}_{t\geq 0}$ is a sequence of random variables and $\{\mathscr{F}_t\}_{t\geq 0}$ is a filtration.

that their belief will become more precise upon observing the new signal, while still being influenced by their original motivated belief. Since $\widetilde{\mu}_0 = \mu_0 + \frac{1}{2}\alpha\sigma_0^2$, the expected belief utility with the posterior is now:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda\widetilde{\mu}_0 + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{24}$$

$$= \alpha\left(\lambda\left(\mu_0 + \frac{1}{2}\alpha\sigma_0^2\right) + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{25}$$

$$= \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2. \tag{26}$$

Compared to the belief utility with the prior (eq. 16), the expected belief utility with the posterior for the unaware motivated reasoner (eq. 26) is completely identical. This is simplified and formalized in Theorem 3.

Theorem 3. *The decision-maker, who is unaware of his own directional motive, will acquire a signal s with an overall benefit b if and only if:*

$$b \geqslant 0. \tag{27}$$

Proof. See the proof of this theorem in the Appendix (p. 39). ***Q.E.D.***

Overall, for the unaware motivated reasoner, there is no threshold for acquiring information. Therefore, regardless of the extent of their directional motives or the precision of the signal, such a DM will always acquire the signal as long as the overall benefit of receiving information is not negative.

3.3 Information-Sensitive Motivated Reasoner

Now let us consider the situation in which the DM is sensitive to the type of information he receives.

Formally, we allow the signal to be perceived as "good" or "bad", that is $s \in \{\text{Good}, \text{Bad}\}$. We model the DM's sensitivity to the asymmetry of information by the absolute difference between the value (or utility) of a favorable signal and an unfavorable signal by $\delta = |o^+ - o^-|$, where $o^+ = o(s = \text{Good})$ and $o^- = o(s = \text{Bad})$ represent utility from 'good news' and 'bad news,' respectively.

Given this, we have $\mathbb{E}_s[s] = \mu - \delta$, where $\mu \in \{\mu_0, \widetilde{\mu}_0\}$ and where δ represents the net difference between the values of the 'positive' and 'negative' signals, which is subtracted from the specific belief. Since the asymmetry of belief updating is context-dependent, this approach allows us *not* to specify whether the DM reacts more to 'good news' than to 'bad news' or vice versa (see more on this in Section 5).

Since $\mathbb{E}_s[s] = \mu - \delta$, where $\mu \in \{\mu_0, \widetilde{\mu}_0\}$, the expected belief utility with the posterior for the self-aware (SA) information-sensitive motivated reasoner is:

$$\mathbb{E}_s^{SA}\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda(\mu_0 - \delta) + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2\sigma_0^2 \tag{28}$$

$$= \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 - \alpha\lambda\delta. \tag{29}$$

At the same time, since $\widetilde{\mu}_0 = \mu_0 + \frac{1}{2}\alpha\sigma_0^2$, the expected belief utility with the posterior for the unaware (UA) information-sensitive motivated reasoner is:

$$\mathbb{E}_s^{UA}\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda(\widetilde{\mu}_0 - \delta) + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{30}$$

$$= \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 - \alpha\lambda\delta. \tag{31}$$

That is the expected belief utility with the posterior for the self-aware information-sensitive motivated reasoner (eq. 29) is identical to the expected belief utility with the posterior for the unaware information-sensitive motivated reasoner (eq. 31).

Compared to the belief utility with the prior (eq. 16), the expected belief utility for the (both self-aware and unaware) information-sensitive motivated reasoner differs by the subtraction of the term $\alpha\lambda\delta$ (eq. 31). This is simplified and formalized in Theorem 4.

Theorem 4. *If the decision-maker is sensitive to information—disregarding whether he is self-aware or unaware of his directional motives when making an information acquisition choice—then signals s are acquired if and only if:*

$$b \geqslant \alpha\lambda\delta. \tag{32}$$

Proof. See the proof of this theorem in the Appendix (p. 39). ***Q.E.D.***

The initial aspect of comparative statics is that obtaining information comes at a cost because the DM anticipates that their objective mean belief will not change from what it was before acquiring the information. Additionally, the DM faces a cost for having a more accurate belief, as this precision (λ) makes it more challenging to uphold the desired belief.

Another feature of comparative statics is that the threshold for information acquisition increases with δ, meaning that subjects who are more sensitive to the type of information are less likely to seek it. This reflects the idea that the greater the (absolute) difference between the value of favorable and unfavorable signals, the lower the decision maker's (DM's) propensity to acquire new information.

Finally, the information acquisition threshold increases with α, meaning that more motivated subjects are less likely to seek information. This reflects the idea that stronger motivated reasoning leads to a lower propensity to acquire new information, as it could potentially contradict the subject's desired beliefs.

Overall, for the information-sensitive motivated reasoner, the threshold for acquiring information decreases with the precision of the signal (λ), increases with the DM's extent of inclination toward directional motives (α) and increases with the DM's sensitivity to the asymmetry of information (δ).

4 Choosing (Not) to Acquire the Signal Based on Actions

Suppose now the DM can anticipate making better decisions with more information. Consider the DM taking an action a with a separate decision utility:

$$u(a,w) = -\ell(a-w)^2, \tag{33}$$

where a is the action chosen by the DM, w is the true state of the world, and $\ell > 0$ represents the loss parameter, capturing the cost of making an incorrect decision. This is a standard introduction of the loss operationalized in such a way that the real state of the world is subtracted from the estimate under a special weight (see, e.g., Savage, 1972; Raiffa and Schlaifer, 2000; DeGroot, 2005/1970).

Since w is normally distributed and represents a normal distribution with mean μ_s and variance σ_s^2, it is easy to see[7] that the expected loss is expressed as

$$\mathbb{E}_s\left[\mathbb{E}_w\left[-\ell(a-w)^2\right]\right] = -\ell(a-\mu_s)^2 - \ell\sigma_s^2. \tag{34}$$

When taking an action equal to the—either objective or motivated—posterior belief, the expectation becomes equal to $-\ell\sigma_s^2$. It is worth noticing, that in the case of action being equal to the posterior belief, whether the objective or motivated posterior belief is considered, this does not affect the result[8], since $\tilde{\sigma}_s^2 = \sigma_s^2$.

The DM will choose to acquire the signal if the expected decision utility after acquiring the signal, i.e., $\mathbb{E}_w[u(\mu_0,w)]$, is greater than or equal to the decision utility without the signal, i.e., $\mathbb{E}_w[u(\mu_s,w)]$, adjusted for any overall benefit b:

$$\mathbb{E}_w[u(\mu_s,w)] + b \geqslant \mathbb{E}_w[u(\mu_0,w)]. \tag{35}$$

[7]One can easily see that, first, $\mathbb{E}_s\left[\mathbb{E}_w\left[-\ell(a-w)^2\right]\right] = \mathbb{E}_w\left[-\ell(a-w)^2\right]$, and, second, $\mathbb{E}_w[-\ell(a-w)^2] = -\ell\mathbb{E}_w\left[a^2-2aw+w^2\right] = -\ell\left(a^2-2a\mu_s+\mathbb{E}_w\left[w^2\right]\right) = -\ell\left(a^2-2a\mu_s+\mathbb{V}_w[w]+(\mathbb{E}_w[w])^2\right) = -\ell\left(a^2-2a\mu_s+\sigma_s^2+\mu_s^2\right) = -\ell\left((a-\mu_s)^2+\sigma_s^2\right)$.

[8]One can easily see that: $\mathbb{E}_w[u(\mu_s,w)] = \mathbb{E}_w[-\ell(\mu_s-w)^2] = -\ell\sigma_s^2 = \mathbb{E}_w[-\ell(\tilde{\mu}_s-w)^2] = \mathbb{E}_w[u(\tilde{\mu}_s,w)]$.

Since both $\mathbb{E}_w[u(\mu_0, w)] < 0$ and $\mathbb{E}_w[u(\mu_s, w)] < 0$, we equation becomes

$$-\ell\sigma_s^2 - b \leqslant -\ell\sigma_0^2 \tag{36}$$

$$b \geqslant \ell\lambda\sigma_0^2. \tag{37}$$

If we let b be equal to the reduction in variance with the new information and, at the same time, the minimum satisfactory condition for eq. (37), then $b = \ell\lambda\sigma_0^2$. This allows us to derive three corollaries from Theorems 2, 3, and 4 in the following way.

4.1 Self-aware Motivated Reasoner

First, let us (re)consider the DM who is self-aware of his own directional motive. This case is captured in Corollary 2.

Corollary 2. *The decision-maker, who is self-aware of his own directional motive, will acquire a signal s with the loss parameter ℓ if and only if:*

$$\ell \geqslant \frac{1}{2}\alpha^2. \tag{38}$$

Proof. See the proof of this corollary in the Appendix (p. 40). ***Q.E.D.***

The main feature of comparative statics in Corollary 2 is that the loss parameter increases with α, which implies that more motivated subjects are less likely to seek information. This reflects the idea that stronger motivated reasoning leads to a lower propensity to acquire new information, as it could potentially contradict the subject's desired beliefs.

Corollary 2, which demonstrates that both the cost and benefit are scaled by the reduction in variance, carries a natural and intuitive insight: individuals will acquire information when the benefit of making better decisions outweighs the cost associated with the potential for less pleasant beliefs. However, the DM underestimates the value of information, as it also leads to less distorted beliefs and, consequently, better decision-making.

4.2 Unaware Motivated Reasoner

Second, let us turn to the DM who is unaware of his own directional motive. This is formalized in Corollary 3.

Corollary 3. *The decision-maker, who is unaware of his own directional motives, will always acquire a signal s:*

$$\ell > 0. \tag{39}$$

Proof. See the proof of this corollary in the Appendix (p. 40). ***Q.E.D.***

For the unaware motivated reasoner, there is no threshold for the loss parameter ℓ. Therefore, regardless of the extent of their directional motives or the precision of the signal, such a DM will always acquire the signal.

Simultaneously, a notable feature of the modeling approach employed is that, when forming beliefs, individuals do not account for the fact that this will result in poorer choices during the decision-making stage.

4.3 Information-Sensitive Motivated Reasoner

Finally, let us consider the DM who is information-sensitive. This case is formalized in Corollary 4.

Corollary 4. *The decision-maker, who is information-sensitive, disregarding whether he is self-aware or unaware of his directional motives when making an information acquisition choice, will acquire a signal s with the loss parameter ℓ if and only if:*

$$\ell \geqslant \alpha\delta\frac{1}{\sigma_0^2}. \tag{40}$$

Proof. See the proof of this corollary in the Appendix (p. 40). ***Q.E.D.***

The first feature of comparative statics in Corollary 4 is that the threshold for the loss parameter ℓ increases with α, which implies that more motivated subjects are less likely to seek information.

It also increases with the DM's sensitivity to the asymmetry of information δ, which means motivated reasoners who are more sensitive to the difference between good and bad news are less likely to choose to receive information.

Finally, the loss parameter ℓ decreases in σ_0^2, that is a larger σ_0^2 reduces the threshold for acquiring information. This means that as the decision-maker becomes less certain about their prior beliefs (higher uncertainty), they are more willing to acquire new information because the required loss parameter ℓ decreases. In other words, the more a DM is confident in his prior beliefs (lower σ_0^2), the less he feels the need for new information.

5 Discussion

Here, we provide a brief discussion of several aspects of modeling motivated reasoning, as well as the experimental settings for testing the model.

5.1 Modelling Motivated Reasoning

So far, neither the political science nor economics literature appears to have reached a consensus on which approach is more solid and robust. While alternative formalizations are also feasible (a relevant discussion can be found in, e.g., Bracha and Brown, 2012; Mayraz, 2019; Little, 2022a), the modeling approach to motivated reasoning employed in this paper offers several advantages (see below).

This paper presents a minimal formal behavioral model of information acquisition. More precisely, this paper derives conditions under which motivated information acquisition occurs, building on the theoretical foundations introduced by Little (2022a). A key contribution is the formalization of conditions for information acquisition among three categories of decision-makers: those who are aware of their motivated reasoning, those who are unaware, and those with information sensitivity. Furthermore, we generalize the solutions to pivotal challenges outlined in Little (2022a), including the belief updating process of decision-makers and the (motivated) belief utility function $\widetilde{B}(f)$, extending these from discrete to continuous domains.

Although our discussion naturally follows the whole logic of the paper—which is reflected in the formal details (above) and the discussion (below) of the model, it is still worth noticing to explicitly articulate that we do not deal with such subtopics of motivated reasoning literature as motivated memory (e.g., Saucet and Villeval, 2019; Chew et al., 2020; Hagenbach and Koessler, 2022; Amelio and Zimmermann, 2023; Sial et al., 2023; Fudenberg et al., 2024) or motivated social interactions (e.g., Marshall, 2019; Gonzalez-Jimenez, 2022; Momsen and Ohndorf, 2022; Oprea and Yuksel, 2022; Exley and Kessler, 2023; Ruzzier and Woo, 2023; Wang et al., 2023; Exley and Kessler, 2024; Stötzer and Zimmermann, 2024).

5.1.1 Advantages

First, the approach to modeling motivated reasoning used in this paper provides a simple closed-form solution applicable to any belief distribution and directional motive. This builds a simple and tractable formal model that accounts for several types of decision-makers and presents rigorous conditions for acquiring the signal.

Second, our approach, based on the framework developed by Little (2022a), encompasses several behavioral regularities documented in recent literature, both in political psychology and behavioral economics. While the comprehensive version of Little (2022a) (known as *Twice-Motivated reasoning*) addresses various empirical and theoretical findings (e.g., disputes regarding the quality of information linked to prior beliefs, variations in disagreement influenced by directional or accuracy primes, and asymmetric updating in response to "good" versus "bad" news), a simpler model (referred to as *Once-Motivated reasoning*) proves to be more powerful than the standard framework.

Third, our approach enables us to derive several testable predictions. Some of these predictions have already been examined in the literature and support the hypotheses outlined in the model presented in this paper, demonstrating the relevance of our theory in bridging behavioral regularities with a formal theoretical framework. Other predictions are novel contributions to the literature and offer potential avenues for future research. In subsection 5.2, we elaborate on how specific hypotheses are derived from the formal model and discuss them in the context of existing literature.

5.1.2 Assumptions

An important part of every formal modelling is assumptions. Here we discuss most crucial ones.

Unawareness vs. awareness of own biases. One key assumption is the possibility of distinguishing decision-makers who are self-aware of cognitive bias (in this case, motivated reasoning) from those who are unaware. This distinction is crucial for understanding how people process information, especially in contexts like political belief formation and decision-making. Self-aware motivated reasoners recognize that their reasoning may be influenced by their biases or desires. However, awareness does not guarantee the ability to overcome these biases; they may still engage in motivated reasoning but do so with an understanding of the factors influencing them. In contrast, unaware reasoners remain oblivious to the influence of their biases, genuinely believing they are evaluating evidence objectively, even when their reasoning is shaped by these biases.

Sensitivity to information. The case where individuals place greater weight on 'good news' signals is well-documented in both political psychology and behavioural economics literature. A notable example, increasingly studied in experimental economics, involves subjects who are asked to complete a series of cognitive tasks and subsequently rank themselves among other participants. This is often relevant to their identity, such as considering themselves intellectuals (e.g., Zimmermann, 2020; Drobner, 2022). In contrast, the situation where individuals assign more weight to 'bad news' signals is less represented in the economics literature. This scenario often arises when a subject's identity is tied to negative feedback. For instance, consider a peasant community where a strong social norm dictates staying connected to the land and discourages leaving it for any reason, such as pursuing education. In this context, receiving objectively negative news, such as failing a university entrance exam, is likely to be weighted more heavily by a peasant who does not want to leave his land, as it aligns with his desire to remain connected to his community. Thus, by incorporating δ, the absolute value of the difference between the values of 'good' and 'bad' news, we can avoid this context-dependency.

Objective vs. subjective value of information. A substantial body of literature explicitly documents that subjects care not only about the objective value of information but also about the potential psychological effects of encountering and incorporating it (e.g., Golman et al., 2017; Benjamin, 2019; Sharot and Sunstein, 2020; Stanovich, 2021; Kelly and Sharot, 2021; Molinaro et al., 2023; Vu et al., 2023). This holds true in the context of motivated reasoning as well (e.g., Chen et al., 2021; Dimant et al., 2024). Thus, distinguishing between b^{obj} and b^{psy} enhances the model's effectiveness.

5.2 Testable Hypotheses and Related Literature

In line with the insights provided by the formal model, we generate a range of hypotheses that can be tested empirically (perhaps, through experiment). These hypotheses are directly informed by the model's structure and allow for experimental validation of its theoretical implications.

First, following Theorem 2, the threshold for information acquisition increases with the precision of the signal. A more precise signal provides less flexibility for motivated belief formation, making it less desirable for the motivated reasoner to acquire the signal.

Hypothesis 1 (Signal precision)**.** *Individuals will be less likely to acquire information as the precision of the signal (λ) increases.*

Empirically, **H(1)** is consistent with recent experimental studies. For instance, Engelmann et al. (2024) test how subjects undertake pattern recognition tasks (where certain patterns could lead to an electric shock or financial loss). The tendency to engage with wishful thinking—i.e., as the authors define it, self-deception motivated by anticipatory utility concerns (e.g., by a desire to feel better about the future)[9]—becomes stronger when signals are less precise. More specifically, Engelmann et al. (2024) document that wishful thinking becomes more evident when participants face more ambiguous or chal-

[9]Namely, Engelmann et al. (2024) define wishful thinking as "self-deception that is driven by a desire to feel better about the future" (p. 927), or "self-deception motivated by anticipatory utility concerns" (p. 929).

lenging patterns: since it generates weaker signals which by construction are easier to interpret as more favorable outcomes to a subject[10].

Second, Corollary 2 shows that motivated decision-makers will acquire a signal s if and only if $\ell \geqslant \frac{1}{2}\alpha^2$. Since the threshold for information acquisition increases with the strength of the directional motives (α), the model predicts that more motivated individuals will be less inclined to seek information.

Hypothesis 2 (Directional motives)**.** *Individuals with higher directional motives (higher α) will be less likely to acquire information compared to those with lower directional motives.*

From an experimental economics literature perspective, **H(2)** also captures the recent behavioral patterns. For example, Drobner and Goerg (2024) introduce a simple but efficient mechanism that enables researchers to manipulate the extent of perceived ego-relevance, which, in terms, affects the subjects' ex ante degree of motivation or degree of directional motives. To do so Drobner and Goerg (2024) provide participants with different articles summarizing scientific evidence regarding the importance of IQ tests: half of the sample read about the strong correlation between IQ and ego-relevant outcomes like income and health (the High-Ego treatment) and the other half read about the evidence against IQ as a measure of intelligence (the Low-Ego treatment). Eventually, Drobner and Goerg (2024) document that subjects in the High-Ego treatment exhibited overconfident initial beliefs and ended with more optimistic final beliefs compared to those in the Low-Ego treatment. They also updated their beliefs more optimistically, showed a higher tendency to make incorrect or no updates when faced with negative signals, and rationalized negative feedback by diminishing the importance of the IQ test and downplaying their effort, whereas in the Low-Ego treatment, updates were consistent regardless of signal valence.

Another, though less subtle, solution to manipulate individual ego-relevance is presented in Burro and Castagnetti (2024). They designed the experiment such that, in the ego-relevant condition, subjects were ranked according to their actual performance on an IQ test, whereas in the non-ego-relevant condition, subjects were randomly assigned positions in the rank.

Third, Theorem 2 and Theorem 3 imply that self-aware motivated reasoners face a strictly positive threshold for acquiring information, whereas unaware motivated reasoners will acquire the signal as long as the cost-benefit ratio is non-negative. Therefore,

[10]The statistical analysis of the difficulty coefficient across different patterns indicates that subjects are less accurate with harder patterns. The variability of interaction coefficients—i.e., Shock/Loss pattern × HAB or Shock/Loss pattern × DP (see even-numbered columns in Table 2 (p. 940) in Engelmann et al., 2024)—demonstrates that the impact of loss or shock patterns intensifies as difficulty increases.

self-aware motivated reasoners will engage in more information avoidance compared to unaware motivated reasoners.

Hypothesis 3 (Self-aware vs. unaware motivated reasoner)**.** *Self-aware motivated reasoners are less likely to acquire information than unaware motivated reasoners.*

To the best of our knowledge, **H(3)** appears to be relatively new to the literature, and we provide a brief discussion on how this hypothesis can be tested in experimental settings in subsection 5.3.

Fourth, Corollary 4 shows that the threshold for acquiring information increases with the DM's sensitivity to the difference between favorable and unfavorable signals (δ). Therefore, the model predicts that decision-makers who react more strongly to the difference between good and bad news will be less likely to seek information.

Hypothesis 4 (Good vs. Bad news asymmetry)**.** *Individuals who are more sensitive to asymmetric information (higher δ) will be less likely to acquire information compared to those who are less sensitive.*

From an empirical point of view, the major part of the literature on motivated reasoning shows that the good vs. bad news asymmetry is present, which gives rather strong evidence for **H(4)**. That is subjects update (significantly) much more under favorable signals than under unfavorable ones as long as these signals are not orthogonal to the decision-maker's identity. The asymmetry of updating is a rather robust finding which is captured by many designs and specifications (e.g., Coutts, 2019; Zimmermann, 2020; Thaler, 2021; Drobner, 2022; Drobner and Goerg, 2024; Melnikoff and Strohminger, 2024; Thaler, 2024a). Since, on average, the identity of intellectuals is considered important for Western students, one of the widespread contexts in which the asymmetry of belief updating is tested is the IQ test (also known as Raven matrices).
Overall, the presence of identity relevance seems to be the major factor of motivated reasoning. While it is not yet perfectly clear[11], the valence (positive vs. negative) of news in the context of financial decision-making—regardless of the subject's identity—does not appear to cause motivated reasoning (e.g., Barron, 2021). Similarly, the valence of news—whether independent of or relevant to their functional values—in the context of message trust does not lead to motivated reasoning (e.g., Thaler, 2024b).

Fifth, according to eq. (18) and Corollary 2, the threshold for information acquisition increases with the prior variance σ_0^2. The belief utility with the posterior decreases as the precision of prior beliefs improves. Therefore, the DM with greater dispersion of his prior

[11]While the financial domain does not inherently evoke the relevance to subjects that triggers motivated reasoning, an interesting direction for research—both fundamentally significant and methodologically sound—would be to investigate the specific mechanisms through which a financial context may induce motivated reasoning.

beliefs (i.e., higher variance σ_0^2) will have a lower threshold for acquiring information, as the potential benefit of reducing uncertainty is greater.

Hypothesis 5 (Prior Confidence)**.** *Individuals with stronger prior beliefs (lower σ_0^2) are less likely to seek new information compared to those with more uncertain prior beliefs (higher σ_0^2).*

While dealing with a novel approach to identifying motivated reasoning, Thaler (2024a) addresses the issue of the (over-)precision of prior beliefs. First, he shows that the results on overprecision indicate that subjects often overestimate the accuracy of their beliefs, particularly on politicized topics, with confidence intervals containing the correct answer less than 50 percent of the time. This overprecision is more pronounced in partisans than in moderates, supporting the idea that individuals with strong prior beliefs (lower σ_0^2) might be overconfident in their assessments. Consequently, overprecision may reduce their likelihood of seeking new information, as they perceive their current beliefs to be sufficiently accurate. Furthermore, Thaler (2024a) notes that overprecise individuals are more likely to trust error-reinforcing Fake News and are less likely to trust True News compared to underprecise individuals, suggesting that motivated beliefs can shape how individuals process and assess information. This behavior aligns with **H(5)**.

Sixth, following Theorems 2, 3, and 4, the threshold for information acquisition increases as the exogenous benefit (b^{obj}) rises.

Hypothesis 6 (Value of higher incentives)**.** *Individuals under higher incentives will (higher b^{obj}) will be more likely to acquire information.*

From an empirical point of view, there is some evidence supporting **H(6)**, although the results remain somewhat mixed. For instance, Zimmermann (2020) investigates the effect of low vs. high stakes specifically in the context of motivated reasoning (i.e., €2 in the *Recall* treatment vs. €50 in the *RecallHigh* treatment). The construction of the feedback in Zimmermann (2020) is organized in such a way that, after undertaking the IQ test, subjects are randomly assigned to a group of ten and ranked according to their performance on the IQ test, although the rank is not disclosed to the subjects. For each subject in the group, three (out of the remaining nine) participants are randomly chosen, and the feedback provided specifies only the relative position of subject i compared to these three randomly chosen subjects. (For instance, if subject i is ranked fifth in the group of ten, and the positions of the three randomly picked subjects are 2, 8, and 9, then the feedback provided will convey: "1 subject is ranked higher than you, and 2 subjects are ranked lower than you," where "you" refers to subject i.) Eventually, Zimmermann (2020) finds that higher incentives improve recall accuracy only among those subjects who received negative (unfavorable) feedback; higher incentives have no statistically significant effect on subjects who received positive feedback.

At the same time, Engelmann et al. (2024) aims to find the effect of stakes on accuracy. In the first two experiments (experiments 1 and 2), Engelmann et al. (2024) uses patterns that are either tilted to the right or left (commonly referred to as Gabor patterns in the paper)[12]. Since these patterns may be too specific to capture a general behavioral phenomenon, the authors introduce two additional experiments. In these experiments, they replace the initial visual task of identifying a single right-or-left tilted pattern in a picture (a single Gabor flash) with a series of eight such pictures (eight Gabor flashes in experiment 3) and a square consisting of red and blue dots that exhibits another type of pattern (colored dots in experiment 4)[13]. Based on these latter types of visual tasks, Engelmann et al. (2024) explicitly examines the effect of stake sizes (namely: €1 vs. €20 in experiment 1, £0.1 vs. £10 in experiment 2, £0.05 vs. £10 in experiment 3, and £0.05 vs. £10 in experiment 4) on the accuracy of pattern recognition. In all four experiments, for both OLS specifications—with and without interaction terms—the effect of the high accuracy bonus is not significant, except in one case: only the regression specification without the interaction term shows the significant effect of higher stakes on accuracy[14].

Seventh, following Theorem 2 and Corollary 2, the threshold for information acquisition is lower when decision-makers are more focused on objective decision-making. When γ is approaching its maximum (fully objective reasoners), information is acquired more readily because psychological concerns do not affect the DM. Therefore, as the weight γ assigned to the objective value of information increases, the likelihood of information acquisition increases.

Hypothesis 7 (Trade-off between objective and subjective value of information)**.** *Individuals who assign greater weight to the objective value of information (higher γ) will be more likely to acquire information.*

From an empirical perspective, the discussion of this hypothesis is closely related to one we had on **H(6)**. However, additional insights can be drawn in connection with **H(7)**. Specifically, we find an explicit investigation of willingness to pay very useful in this context. One rather popular method is based on an incentive-compatible mechanism introduced by Becker et al. (1964). For instance, Engelmann et al. (2024) use it to estimate the individual value of having distorted beliefs. Or, as another somewhat relevant

[12]Although Gabor (1946) pioneered the concept, it was Granlund (1978) who generalized it to two dimensions.

[13]Examples for all three types of visual tasks can be found on p. 933 in Engelmann et al. (2024).

[14]Namely, specification 7 regresses accuracy on the shock/loss pattern (SLP), high accuracy bonus (HAB), and a difficult pattern (DP). The effect of HAB is significant (coefficient: 1.732, standard error: 0.628), indicating that providing a high accuracy bonus tends to increase accuracy by 1.732 percentage points. However, when the interaction terms are added in the specification (namely, SLP × HAB and SLP × DP), the effect for high accuracy bonus becomes non-significant (coefficient: 1.050, standard error: 0.856), and neither is the interaction term significant (coefficient: 1.363, standard error: 1.325). Constants are included in both regression specifications.

example, Guan et al. (2023), interested in measuring preferences for informativeness, derive strict cardinal preferences by assessing subjects' willingness to pay for access to the information structure that impacts their forthcoming decisions.

5.3 Further Notes on the Experiment Design

Prior and posterior belief distributions. Using the experimental framework for belief-updating experiments, we can easily collect the whole distribution of both prior and posterior beliefs. For example, specifically, in the context of motivated reasoning Zimmermann (2020) collects the whole prior and posterior distribution of beliefs about subjects' position in the ranking based on their performance in the IQ test. Using similar procedures, this will give a researcher both μ_0^2 and σ_0^2.

Eliciting subjects' (un)awareness of own directional motives. In general, in psychological literature, awareness can be tested in various ways: designing the experiment based on inattentional blindness, where, after the primary task, participants are asked whether they noticed anything unusual (as in the famous test with the gorilla from Simons and Chabris, 1999), or basing it on change blindness, where participants are asked if they notice any differences between the images (see, e.g., Simons and Levin, 1997; Simons and Rensink, 2005). While serving as potentially interesting controls for the general notion of awareness, these measures cannot be properly applied to experimental settings in the context of motivated reasoning.

One straightforward way to manipulate awareness is to vary the level of information provided about the issue of motivated reasoning specifically, or about different (belief) biases in general. In this context, using a conventional design capable of eliciting motivated reasoning and incorporating two treatments can serve to test **H(3)**. While minor, yet crucial, details of the experimental design need refinement, the general idea is relatively simple: in one treatment, subjects are provided with a brief introduction on cognitive biases that humans are generally susceptible to, the consequences of relying on these biases in decision-making, and general strategies to mitigate them[15]. Naturally, embedding initial (and subsequent) measures before (and after) introducing information on biases may be essential. Additionally, beyond directly structuring the treatments, measures of reaction or response time—which are increasingly gaining popularity—can be employed as a useful metric.

[15]Depending on the specific research question, researchers may find it more useful and precise to separate the testing of each element. For example, disentangling the introduction of biases and explanation of their consequences from explanations of optimal solutions and ways to avoid these biases within individual treatments could comprise two separate treatments.

Eliciting subjects' psychological value of information. While operationalizing the benefit b of information acquisition within a Bayesian belief updating framework can be challenging—since it must capture a clear and measurable advantage for the decision-maker—this can be accomplished in the conventional experimental economics framework through monetary rewards. To assess the effect of the objective value of information (b^{obj}), a straightforward approach is to use two treatments with high and low payoffs. In a recent study, Enke et al. (2023b) investigate the impact of incentives on four well-documented biases—anchoring, base-rate neglect, failure of contingent thinking, and intuitive reasoning—and find that while very high incentives result in only minor improvements in performance, they do increase response times by 40%. To capture the subjective value of information (b^{psy}), an experimenter can design the incentives to reflect subjects' willingness to pay. For instance, Guan et al. (2023) examine subjects' demand for information—shaped by the *informativeness* of information structures—in an abstract laboratory experiment context.

6 Conclusion

In this paper, we build on the theoretical framework introduced by Little (2022a) to develop a formal minimalistic framework of motivated information acquisition. We contrast this with standard belief updating, focusing on how a decision maker (DM) with directional motives approaches the decision to observe or ignore feedback (signals) about the state of the world.

Our findings reveal several key insights. First, for a DM with objective beliefs and directional motives, there is a unique solution to the belief updating problem. Second, for a self-aware motivated reasoner, the decision to acquire information becomes costlier with increased sensitivity to directional motives and decreases with higher signal precision. Conversely, unaware motivated reasoners acquire information as long as there is any positive benefit. Finally, for information-sensitive motivated reasoners, the threshold for acquiring information is influenced by both the DM's inclination toward directional motives and sensitivity to information asymmetry, with more precise signals making information acquisition less appealing.

For future research, we believe there are several fruitful directions. In a narrow sense, one could investigate specific cases of prior belief distributions. For instance, flat priors, which assume that all values are equally likely, represent no prior information, while vague priors, which set a very large but finite variance, model the DM's absence or lack of prior knowledge in such a way that the distribution is so spread out it provides minimal information. Finally, given explicit predictions of the model, it would be natural to test them in experimental settings.

In a broad(er) sense, a crucial question is the development of an axiomatization of the DM's costs and benefits in making decisions. While the standard paradigm has captured researchers' attention for at least three-quarters of a century (e.g., Bohnenblust et al., 1949; Blackwell et al., 1951), behavioral refinements are still in their infancy. Furthermore, one could explore how to adapt our minimalistic framework to choices involving different signals of known characteristics.

Overall, our results underscore the complexities of information acquisition when directional motives are at play and highlight the nuanced role of self-awareness in decision-making under uncertainty.

References

Amelio, Andrea and Florian Zimmermann, "Motivated Memory in Economics—A Review," *Games*, 2023, *14* (1), 15.

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani, "Strategic information acquisition and transmission," *American Economic Journal: Microeconomics*, 2016, *8* (3), 119–155.

Banovetz, James and Ryan Oprea, "Complexity and procedural choice," *American Economic Journal: Microeconomics*, 2023, *15* (2), 384–413.

Barlevy, Gadi and Pietro Veronesi, "Information acquisition in financial markets," *The Review of Economic Studies*, 2000, *67* (1), 79–90.

Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya, "Facts, alternative facts, and fact checking in times of post-truth politics," *Journal of Public Economics*, 2020, *182*, 104123.

Barron, Kai, "Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains?," *Experimental Economics*, 2021, *24* (1), 31–58.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9* (3), 226–232.

Bénabou, Roland, "The economics of motivated beliefs," *Revue d'économie politique*, 2015, *125* (5), 665–685.

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

_ **and** _ , "Mindful economics: The production, consumption, and value of beliefs," *Journal of Economic Perspectives*, 2016, *30* (3), 141–164.

Benjamin, Daniel J, "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2019, *2*, 69–186.

Bergemann, Dirk and Juuso Välimäki, "Information acquisition and efficient mechanism design," *Econometrica*, 2002, *70* (3), 1007–1033.

Blackwell, David et al., "Comparison of experiments," *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, 1951, *1* (93-102), 26.

Bohnenblust, Henri Frédéric, Lloyd S Shapley, and Seymour Sherman, "Reconnaissance in game theory," *Research Memorandum*, 1949.

Bracha, Anat and Donald J Brown, "Affective decision making: A theory of optimism bias," *Games and Economic Behavior*, 2012, *75* (1), 67–80.

Bronnikov, Egor, "Convenience Over Conviction: How Noise Aversion Shapes Information Choices," *Working paper*, 2024.

— , "Motivated Efforts," *Working paper*, 2024.

Bulmer, MG, *Principles of statistics*, Dover, 1979.

Burro, Giovanni and Alessandro Castagnetti, "Will I tell you that you are smart (dumb)? Deceiving Others about their IQ or about a Random Draw," *Journal of Behavioral and Experimental Economics*, 2022, *100*, 101920.

— **and** — , "The ego is no fool: Absence of motivated belief formation in strategic interactions," *Journal of Behavioral and Experimental Economics*, 2024, *109*, 102167.

Carter, Erin Baggott and Brett L. Carter, *Propaganda in autocracies: institutions, information, and the politics of belief*, Cambridge University Press, 2023.

Charness, Gary, Ryan Oprea, and Sevgi Yuksel, "How do people choose between biased information sources? Evidence from a laboratory experiment," *Journal of the European Economic Association*, 2021, *19* (3), 1656–1691.

Chen, Si, Carl Heese et al., "Fishing for good news: Motivated information acquisition," *Workign paper*, 2021.

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

Colombo, Luca, Gianluca Femminis, and Alessandro Pavan, "Information acquisition and welfare," *The Review of Economic Studies*, 2014, *81* (4), 1438–1483.

Coutts, Alexander, "Good news and bad news are still news: Experimental evidence on belief updating," *Experimental Economics*, 2019, *22* (2), 369–395.

Cover, Thomas M, *Elements of information theory*, John Wiley & Sons, 1999.

DeGroot, Morris H, *Optimal statistical decisions*, John Wiley & Sons, 2005/1970.

Dimant, Eugen, Fabio Galeotti, and Marie Claire Villeval, "Motivated information acquisition and social norm formation," *European Economic Review*, 2024, p. 104778.

Dominiak, Adam, Matthew Kovach, and Gerelt Tserenjigmid, "Minimum distance belief updating with general information," *Working paper*, 2021.

Drobner, Christoph, “Motivated beliefs and anticipation of uncertainty resolution,” *American Economic Review: Insights*, 2022, *4* (1), 89–105.

__ **and Sebastian J Goerg**, “Motivated belief updating and rationalization of information,” *Management Science*, 2024.

Druckman, James N and Mary C McGrath, “The evidence for motivated reasoning in climate change preference formation,” *Nature Climate Change*, 2019, *9* (2), 111–119.

Egorov, Georgy and Konstantin Sonin, “The Political Economics of Non-democracy,” *Journal of Economic Literature*, 2024, *University of Chicago, Becker Friedman Institute for Economics Working Paper 2023-67.*

Engelmann, Jan B, Maël Lebreton, Nahuel A Salem-Garcia, Peter Schwardmann, and Joël J van der Weele, “Anticipatory anxiety and wishful thinking,” *American Economic Review*, 2024, *114* (4), 926–960.

Enke, Benjamin, “The Cognitive Turn in Behavioral Economics,” *Working paper*, 2024.

__ **, Thomas Graeber, and Ryan Oprea**, “Complexity and hyperbolic discounting,” Technical Report, National Bureau of Economic Research 2023.

__ **, Uri Gneezy, Brian Hall, David Martin, Vadim Nelidov, Theo Offerman, and Jeroen Van De Ven**, “Cognitive biases: Mistakes or missing stakes?,” *Review of Economics and Statistics*, 2023, *105* (4), 818–832.

Exley, Christine L and Judd B Kessler, “Information avoidance and image concerns,” *The Economic Journal*, 2023, *133* (656), 3153–3168.

__ **and** __ , “Motivated errors,” *American Economic Review*, 2024, *114* (4), 961–987.

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack, “Pathwise concentration bounds for Bayesian beliefs,” *Theoretical Economics*, 2023, *18* (4), 1585–1622.

__ , __ **, and** __ , “Selective memory equilibrium,” *Journal of Political Economy*, 2024.

Gabor, Dennis, “Theory of communication. Part 1: The analysis of information,” *Journal of the Institution of Electrical Engineers-part III: radio and communication engineering*, 1946, *93* (26), 429–441.

Gehlbach, Scott and Konstantin Sonin, “Government control of the media,” *Journal of Public Economics*, 2014, *118*, 163–171.

__ , __ **, and Milan W. Svolik**, “Formal models of nondemocratic politics,” *Annual Review of Political Science*, 2016, *19* (3), 565–584.

Gerardi, Dino and Leeat Yariv, "Information acquisition in committees," *Games and Economic Behavior*, 2008, *62* (2), 436–459.

Gerber, Alan and Donald Green, "Misperceptions about perceptual bias," *Annual review of political science*, 1999, *2* (1), 189–210.

Gino, Francesca, Michael I. Norton, and Roberto A. Weber, "Motivated Bayesians: Feeling moral while acting egoistically," *Journal of Economic Perspectives*, 2016, *30* (3), 189–212.

Golman, Russell, David Hagmann, and George Loewenstein, "Information avoidance," *Journal of economic literature*, 2017, *55* (1), 96–135.

__ **, George Loewenstein, Moene Karl, and Luca Zarri**, "The Preference for Belief Consonance," *Journal of Economic Perspectives*, 2016, *30* (3), 165–188.

Gonzalez-Jimenez, VÃctor, "Bayes rule as a descriptive model: The representativeness heuristic," *Journal of Public Economics*, 2022, *211*, 104662.

Granlund, Goesta H, "In search of a general picture processing operator," *Computer Graphics and Image Processing*, 1978, *8* (2), 155–173.

Guan, Menglong, Ryan Oprea, and Sevgi Yuksel, "Too Much Information," *Working paper*, 2023. [HTML].

Guriev, Sergei and Daniel Treisman, "A theory of informational autocracy," *Journal of Public Economics*, 2020, *186*, 104158.

Hagenbach, Jeanne and Frédéric Koessler, "Selective memory of a psychological agent," *European Economic Review*, 2022, *142*, 104012.

Horz, Carlo M., "Propaganda and skepticism," *American Journal of Political Science*, 2021, *65* (3), 717–732.

__ , "Identity propaganda," *American Journal of Political Science*, 2023, pp. 1–26.

Kelly, Christopher A and Tali Sharot, "Individual differences in information-seeking," *Nature communications*, 2021, *12* (1), 7062.

Kimbrough, Erik O. and Alexander Vostroknutov, "Affective decision-making and moral sentiments," *Working paper*, 2022.

Kobayashi, Hisashi, Brian L Mark, and William Turin, *Probability, random processes, and statistical analysis: applications to communications, signal processing, queueing theory and mathematical finance*, Cambridge University Press, 2011.

Little, Andrew T., "Propaganda and credulity," *Games and Economic Behavior*, 2017, *102*, 224–232.

—, "Detecting motivated reasoning," *OSF Working paper*, 2022. [HTML].

Little, Andrew T, "Information Theory and Biased Beliefs," *OSF Working paper*, 2022. [HTML].

Little, Andrew T., Keith E. Schnakenberg, and Ian R. Turner, "Motivated reasoning and democratic accountability," *American Political Science Review*, 2022, *116* (2), 751–767.

Liu, Qingmin, "Information acquisition and reputation dynamics," *The Review of Economic Studies*, 2011, *78* (4), 1400–1425.

Loewenstein, George and Andras Molnar, "The renaissance of belief-based utility in economics," *Nature Human Behaviour*, 2018, *2* (3), 166–167.

Marshall, John, "Signaling sophistication: How social expectations can increase political information acquisition," *The Journal of Politics*, 2019, *81* (1), 167–186.

Mayraz, Guy, "Priors and Desires: A Bayesian Model of Wishful Thinking and Cognitive Dissonance," *Working paper*, 2019. [HTML].

Melnikoff, David E and Nina Strohminger, "Bayesianism and wishful thinking are compatible," *Nature Human Behaviour*, 2024, *8* (4), 692–701.

Molinaro, Gaia, Irene Cogliati Dezza, Sarah Katharina Bühler, Christina Moutsiana, and Tali Sharot, "Multifaceted information-seeking motives in children," *Nature Communications*, 2023, *14* (1), 5505.

Molnar, Andras and George Loewenstein, "Thoughts and players: an introduction to old and new economic perspectives on beliefs," *The cognitive science of belief: a multidisciplinary approach*, 2022, pp. 321–350.

Momsen, Katharina and Markus Ohndorf, "Information avoidance, selective exposure, and fake (?) news: Theory and experimental evidence on green consumption," *Journal of Economic Psychology*, 2022, *88*, 102457.

Nieuwerburgh, Stijn Van and Laura Veldkamp, "Information acquisition and under-diversification," *The Review of Economic Studies*, 2010, *77* (2), 779–805.

Oprea, Ryan and Sevgi Yuksel, "Social exchange of motivated beliefs," *Journal of the European Economic Association*, 2022, *20* (2), 667–699.

Persico, Nicola, “Information acquisition in auctions,” *Econometrica*, 2000, *68* (1), 135–148.

Peterson, Erik and Shanto Iyengar, “Partisan gaps in political information and information-seeking behavior: Motivated reasoning or cheerleading?,” *American Journal of Political Science*, 2021, *65* (1), 133–147.

Pitman, Jim, *Probability*, Springer Science & Business Media, 2012.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz, “The cost of information: The case of constant marginal costs,” *American Economic Review*, 2023, *113* (5), 1360–1393.

Rabin, Matthew and Joel L Schrag, “First impressions matter: A model of confirmatory bias,” *The Quarterly Journal of Economics*, 1999, *114* (1), 37–82.

Raiffa, Howard and Robert Schlaifer, *Applied statistical decision theory*, Vol. 78, John Wiley & Sons, 2000.

Reza, Fazlollah M, *An introduction to information theory*, Courier Corporation, 1994.

Ruzzier, Christian A and Marcelo D Woo, “Discrimination with inaccurate beliefs and confirmation bias,” *Journal of Economic Behavior & Organization*, 2023, *210*, 379–390.

Ryan, Timothy J and Amanda R Aziz, “Is the political right more credulous? Experimental evidence against asymmetric motivations to believe false political information,” *The Journal of Politics*, 2021, *83* (3), 1168–1172.

Saucet, Charlotte and Marie Claire Villeval, “Motivated Memory in Dictator Games,” *Games and Economic Behavior*, 2019, *117*, 250–275.

Savage, Leonard J, *The Foundations of Statistics*, Courier Corporation, 1972.

Sharot, Tali and Cass R Sunstein, “How people decide what they want to know,” *Nature Human Behaviour*, 2020, *4* (1), 14–19.

Sial, Afras Y, Justin R Sydnor, and Dmitry Taubinsky, “Biased Memory and Perceptions of Self-Control,” Technical Report, National Bureau of Economic Research 2023.

Simons, Daniel J and Christopher F Chabris, “Gorillas in our midst: Sustained inattentional blindness for dynamic events,” *Perception*, 1999, *28* (9), 1059–1074.

_ **and Daniel T Levin**, “Change blindness,” *Trends in cognitive sciences*, 1997, *1* (7), 261–267.

— and **Ronald A Rensink**, "Change blindness: Past, present, and future," *Trends in cognitive sciences*, 2005, *9* (1), 16–20.

Soch, Joram, Thomas J Faulkenberry, Kenneth Petrykowski, and Carsten Allefeld, *The Book of Statistical Proofs*, Zenodo, 2024. [HTML].

Sprengholz, Philipp, Luca Henkel, Robert Böhm, and Cornelia Betsch, "Historical narratives about the COVID-19 pandemic are motivationally biased," *Nature*, 2023, *623* (7987), 588–593.

Stanovich, Keith E, *The bias that divides us: The science and politics of myside thinking*, MIT Press, 2021.

Stoetzer, Lasse S and Florian Zimmermann, "A representative survey experiment of motivated climate change denial," *Nature Climate Change*, 2024, *14* (2), 198–204.

Stötzer, Lasse S and Florian Zimmermann, "A note on motivated cognition and discriminatory beliefs," *Games and Economic Behavior*, 2024, *147*, 554–562.

Thaler, Michael, "Gender differences in motivated reasoning," *Journal of Economic Behavior & Organization*, 2021, *191*, 501–518.

— , "The fake news effect: Experimentally identifying motivated reasoning using trust in news," *American Economic Journal: Microeconomics*, 2024.

— , "Good news is not a sufficient condition for motivated reasoning," *Working paper*, 2024.

Vu, Linh, Ivan Soraperra, Margarita Leib, Joël van der Weele, and Shaul Shalvi, "Ignorance by choice: A meta-analytic review of the underlying motives of willful ignorance and its consequences.," *Psychological Bulletin*, 2023, *149* (9-10), 611.

Wang, Ao, Shaoda Wang, and Xiaoyang Ye, "When information conflicts with obligations: the role of motivated cognition," *The Economic Journal*, 2023, *133* (654), 2533–2552.

Yang, Ming, "Coordination with flexible information acquisition," *Journal of Economic Theory*, 2015, *158*, 721–738.

Zhao, Chen, "Pseudo-Bayesian updating," *Theoretical Economics*, 2022, *17* (1), 253–289.

Zimmermann, Florian, "The dynamics of motivated beliefs," *American Economic Review*, 2020, *110* (2), 337–363.

A Proofs

Here we present the proofs of propositions, theorems, and corollaries in the order in which they appear in the paper.

Proof of Remark 1. To derive the posterior belief, we proceed in four steps.

Prior distribution. The prior belief about the random variable w is normal: $w \sim \mathcal{N}(\mu_0, \sigma_0^2)$. That is

$$\mathbb{P}(w) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(w-\mu_0)^2}{2\sigma_0^2}\right) \tag{41}$$

Likelihood function. The signal s is observed as $s = w + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Given w, the conditional distribution of s is $s \mid w \sim \mathcal{N}(w, \sigma_\epsilon^2)$. That is

$$\mathbb{P}(s \mid w) = \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left(-\frac{(s-w)^2}{2\sigma_\epsilon^2}\right) \tag{42}$$

Posterior distribution. Bayes' theorem states that the posterior distribution of w given the signal s is proportional to the product of the prior and the likelihood:

$$\mathbb{P}(w \mid s) \propto \mathbb{P}(s \mid w)\mathbb{P}(w) \tag{43}$$

Since $\frac{1}{\sqrt{2\pi\sigma_\epsilon^2}}$ and $\frac{1}{\sqrt{2\pi\sigma_0^2}}$ terms are constants with respect to w, we substitute the expressions for the likelihood and the prior and get:

$$\begin{aligned}
\mathbb{P}(w \mid s) &\propto \exp\left(-\frac{(s-w)^2}{2\sigma_\epsilon^2}\right)\exp\left(-\frac{(w-\mu_0)^2}{2\sigma_0^2}\right) && (44)\\
&\propto \exp\left(-\frac{1}{2}\left[\frac{(s-w)^2}{\sigma_\epsilon^2} + \frac{(w-\mu_0)^2}{\sigma_0^2}\right]\right) && (45)\\
&\propto \exp\left(-\frac{1}{2}\left[\frac{s^2}{\sigma_\epsilon^2} - \frac{2sw}{\sigma_\epsilon^2} + \frac{w^2}{\sigma_\epsilon^2} + \frac{w^2}{\sigma_0^2} - \frac{2w\mu_0}{\sigma_0^2} + \frac{\mu_0^2}{\sigma_0^2}\right]\right) && (46)\\
&\propto \exp\left(-\frac{1}{2}\left[w^2\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_0^2}\right) - 2w\left(\frac{s}{\sigma_\epsilon^2} + \frac{\mu_0}{\sigma_0^2}\right) + \left(\frac{s^2}{\sigma_\epsilon^2} + \frac{\mu_0^2}{\sigma_0^2}\right)\right]\right) && (47)\\
&\propto \exp\left(-\frac{1}{2}\left[w^2\left(\frac{\sigma_\epsilon^2 + \sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}\right) - 2w\left(\frac{s\sigma_0^2 + \mu_0\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}\right) + \left(\frac{s^2\sigma_0^2 + \mu_0^2\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}\right)\right]\right) && (48)
\end{aligned}$$

Since $\left(\frac{s^2\sigma_0^2+\mu_0^2\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}\right)$ is constant and given the transformation to complete the (full) square

$$w^2\left(\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}\right)-2w\left(\frac{s\sigma_0^2+\mu_0\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}\right)=\left[\sqrt{\left(\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}\right)}\left(w-\frac{\frac{s\sigma_0^2+\mu_0\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}}{\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}}\right)\right]^2 \tag{49}$$

we, finally, get

$$\mathbb{P}(w\mid s)\propto\exp\left(-\frac{1}{2}\left[\sqrt{\left(\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}\right)}\left(w-\frac{\frac{s\sigma_0^2+\mu_0\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}}{\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}}\right)\right]^2\right) \tag{50}$$

$$\propto\exp\left(-\frac{1}{2}\left(\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}\right)\left(w-\frac{\frac{s\sigma_0^2+\mu_0\sigma_\epsilon^2}{\sigma_\epsilon^2\sigma_0^2}}{\frac{\sigma_\epsilon^2+\sigma_0^2}{\sigma_\epsilon^2\sigma_0^2}}\right)^2\right) \tag{51}$$

$$\propto\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_0^2}\right)\left(w-\frac{\left(\frac{s}{\sigma_\epsilon^2}+\frac{\mu_0}{\sigma_0^2}\right)}{\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_0^2}\right)}\right)^2\right) \tag{52}$$

This expression shows that the posterior distribution of w given s is normal with mean μ_s:

$$\mu_s=\frac{\frac{s}{\sigma_\epsilon^2}+\frac{\mu_0}{\sigma_0^2}}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_0^2}} \tag{53}$$

$$=\frac{s\sigma_\epsilon^{-2}+\mu_0\sigma_0^{-2}}{\sigma_\epsilon^{-2}+\sigma_0^{-2}} \tag{54}$$

$$=\left(\frac{\sigma_\epsilon^{-2}}{\sigma_\epsilon^{-2}+\sigma_0^{-2}}\right)s+\left(\frac{\sigma_0^{-2}}{\sigma_\epsilon^{-2}+\sigma_0^{-2}}\right)\mu_0 \tag{55}$$

and variance σ_s^2:

$$\sigma_s^2=\frac{1}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_0^2}} \tag{56}$$

$$=\frac{1}{\sigma_\epsilon^{-2}+\sigma_0^{-2}} \tag{57}$$

Final simplifications. To simplify the expressions of μ_s and σ_s^2, we can introduce λ such that

$$\lambda=\frac{\sigma_\epsilon^{-2}}{\sigma_\epsilon^{-2}+\sigma_0^{-2}} \tag{58}$$

and, hence,

$$1-\lambda=\frac{\sigma_0^{-2}}{\sigma_\epsilon^{-2}+\sigma_0^{-2}} \tag{59}$$

Then,

$$\mu_s=\lambda s+(1-\lambda)\mu_0 \tag{60}$$

$$\sigma_s^2=(1-\lambda)\sigma_0^2 \tag{61}$$

Thus, given the objective belief follows a normal distribution $w\sim\mathcal{N}(\mu_0,\sigma_0^2)$ and a signal $s=w+\epsilon$, where $\epsilon\sim\mathcal{N}(0,\sigma_\epsilon^2)$, the posterior belief w_s is normal with mean $\mu_s=\lambda s+(1-\lambda)\mu_0$ and variance $\sigma_s^2=(1-\lambda)\sigma_0^2$, where $\lambda=\sigma_\epsilon^{-2}/(\sigma_\epsilon^{-2}+\sigma_0^{-2})$. ***Q.E.D.***

Proof of Theorem 1. It is conceptually similar to Theorem 1 in Little (2022a); however, while in Theorem 1 in Little (2022a) the optimization is performed over a finite-dimensional probability vector with summations over discrete indices, in 1 (here) the optimization is performed over a space of functions with integrals over a continuous domain. The proof proceeds in three steps.

The Problem. Firstly, let us restate the DM's optimization problem.

$$\widetilde{f}=\arg\max_{f'\in\mathcal{F}}\left\{B(f,f')\right\} \tag{62}$$

$$=\arg\max_{f'\in\mathcal{F}}\left\{-D_{KL}(f'||f)+v(f')\right\} \tag{63}$$

$$=\arg\max_{f'\in\mathcal{F}}\left\{-\int_{-\infty}^{\infty}f'(w)\log\left(\frac{f'(w)}{f(w)}\right)dw+\int_{-\infty}^{\infty}v(w)f'(w)dw\right\} \tag{64}$$

$$=\arg\max_{f'\in\mathcal{F}}\left\{\int_{-\infty}^{\infty}f'(w)\left[-\log\left(\frac{f'(w)}{f(w)}\right)+v(w)\right]dw\right\} \tag{65}$$

$$=\arg\max_{f'\in\mathcal{F}}\left\{\int_{-\infty}^{\infty}f'(w)\left[v(w)+\log(f(w))-\log\left(f'(w)\right)\right]dw\right\} \tag{66}$$

$$=\arg\max_{f'\in\mathcal{F}}\left\{\int_{-\infty}^{\infty}f'(w)[v(w)+\log(f(w))]dw-\int_{-\infty}^{\infty}f'(w)\left(\log f'(w)\right)dw\right\} \tag{67}$$

Optimization. Next, we introduce a Lagrange multiplier λ to account for the constraint $\int_{-\infty}^{\infty}f'(w)dw=1$, leading to the following Lagrangian:

$$\mathscr{L}(f',\lambda)=\int_{-\infty}^{\infty} f'(w)\left[v(w)+\log\left(f(w)\right)\right]dw - \int_{-\infty}^{\infty} f'(w)\log\left(f'(w)\right)dw+\lambda\left(1-\int_{-\infty}^{\infty} f'(w)dw\right) \quad (68)$$

Taking the functional derivative with respect to $f'(w)$ and setting it equal to zero gives:

$$\frac{\delta\mathscr{L}}{\delta f'(w)}=v(w)+\log f(w)-\log f'(w)-1-\lambda=0 \quad (69)$$

Simplifying this equation yields:

$$\log f'(w)=v(w)+\log f(w)-\lambda-1 \quad (70)$$

Expanantioation gives us:

$$f'(w)=\exp(v(w)+\log f(w)-\lambda-1) \quad (71)$$
$$=e^{v(w)}\cdot e^{-\lambda-1}\cdot f(w) \quad (72)$$

That is:

$$f'(w)=f(w)\frac{e^{v(w)}}{e^{1+\lambda}} \quad (73)$$

Normalization. To ensure that $f'(w)$ is a valid probability density function, we normalize it by requiring that $\int_{-\infty}^{\infty} f'(w)dw=1$. Also denote $e^{1+\lambda}=c$ This gives:

$$1=\int_{-\infty}^{\infty} f'(w)dw \quad (74)$$
$$=\int_{-\infty}^{\infty} c\cdot f(w)\cdot e^{v(w)}dw \quad (75)$$
$$=c\int_{-\infty}^{\infty} f(w)\cdot e^{v(w)}dw \quad (76)$$

That is

$$C=\frac{1}{\int_{-\infty}^{\infty} f(w)e^{v(w)}dw} \quad (77)$$

Therefore, the optimal density function $\widetilde{f}(w)$ is:

$$\widetilde{f}(w) = \frac{f(w)\exp(v(w))}{\int_{-\infty}^{\infty} f(w')\exp(v(w'))\,dw'} \tag{78}$$

Q.E.D.

Proof of Corollary 1. If $v(\cdot)$ is linear in w, such that $v(w) = \alpha w$, then we can can express $v(w) = \alpha_0 + \alpha_1 w + \alpha_2 w^2$, where $\alpha_2 = 0$. Following the logic of Corollary 6 in Little (2022a), we prove that the density of the motivated belief must be proportional to:

$$\exp\left(\frac{-(w-\mu)^2}{2\sigma^2}\right) \cdot \exp\left(\alpha_2 w^2 + \alpha_1 w + \alpha_0\right) = \exp\left(\frac{-(w-\widetilde{\mu})^2}{2\widetilde{\sigma}^2}\right) \cdot \exp(k) \tag{79}$$

Taking logs of both sides gives the following:

$$\frac{-(w-\mu)^2}{2\sigma^2} + \alpha_2 w^2 + \alpha_1 w + \alpha_0 = \frac{-(w-\widetilde{\mu})^2}{2\widetilde{\sigma}^2} + k \tag{80}$$

Expanding (80) and solving for w^2, we get:

$$-\frac{1}{2\sigma^2} + \alpha_2 = -\frac{1}{2\widetilde{\sigma}^2} \tag{81}$$

$$\widetilde{\sigma}^2 = \frac{\sigma^2}{1 - 2\alpha_2\sigma^2} \tag{82}$$

Since $\alpha_2 = 0$, $\widetilde{\sigma}^2\big|_{\alpha_2=0} = \sigma^2$.

Expanding (80) and solving for w, we get:

$$\frac{\mu}{\sigma^2} + \alpha_1 = \frac{\widetilde{\mu}}{\widetilde{\sigma}^2} \tag{83}$$

$$\widetilde{\mu} = \frac{\mu + \alpha_1\sigma^2}{1 - 2\alpha_2\sigma^2} \tag{84}$$

Since $\alpha_2 = 0$, $\widetilde{\mu}|_{\alpha_2=0} = \mu + \alpha_1\sigma^2$ where we equate $\alpha_1 = \frac{1}{2}\alpha$, $\alpha \geqslant 0$. ***Q.E.D.***

Proof of Proposition 1. This is a proof of Proposition 1. It is conceptually similar to Corollary 3 in Little (2022a); however, here we extend it to a continuous case.

To do so, we need to have the objective function being evaluated at the optimal belief:

$$
\begin{aligned}
B\left(f,\widetilde{f}\right) &= -D_{KL}\left(\widetilde{f}||f\right) + v\left(\widetilde{f}\right) && (85)\\
&= -\int_{-\infty}^{\infty} \widetilde{f}(w)\log\left(\frac{\widetilde{f}(w)}{f(w)}\right)dw + \int_{-\infty}^{\infty} v(w)\widetilde{f}(w)dw && (86)\\
&= -\int_{-\infty}^{\infty} \widetilde{f}(w)\log\left(\frac{\exp(v(w))}{\int_{-\infty}^{\infty} f(w')\exp(v(w'))\,dw'}\right)dw + \int_{-\infty}^{\infty} v(w)\widetilde{f}(w)dw && (87)\\
&= -\int_{-\infty}^{\infty} \widetilde{f}(w)\left[\log\left(e^{v(w)}\right) - \log\left(\int_{-\infty}^{\infty} f(w')e^{v(w')}dw'\right)\right]dw \\
&\quad + \int_{-\infty}^{\infty} v(w)\widetilde{f}(w)dw && (88)\\
&= -\int_{-\infty}^{\infty} \widetilde{f}(w)v(w)dw + \int_{-\infty}^{\infty} \widetilde{f}(w)\log\left(\int_{-\infty}^{\infty} f(w')e^{v(w')}dw'\right)dw \\
&\quad + \int_{-\infty}^{\infty} v(w)\widetilde{f}(w)dw && (89)\\
&= \log\left(\int_{-\infty}^{\infty} f(w')e^{v(w')}dw'\right)\int_{-\infty}^{\infty} \widetilde{f}(w)dw && (90)\\
&= \log\left(\int_{-\infty}^{\infty} f(w')e^{v(w')}dw'\right) && (91)
\end{aligned}
$$

That is the objective function being evaluated at the optimal belief $B\left(f,\widetilde{f}\right)$ is equal to $\log\left(\int_{-\infty}^{\infty} f(w')e^{v(w')}dw'\right)$. ***Q.E.D.***

Proof of Theorem 2. Given that the DM is aware of their directional motive and expects the posterior mean μ_s to be equal to the prior mean μ_0, the expected belief utility after acquiring the signal can be written as:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] = \alpha\mu_0 + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{92}$$

The utility from not acquiring the signal is simply the belief utility with the prior:

$$\widetilde{B}(f_0) = a\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 \tag{93}$$

The DM will choose to acquire the signal if the utility with the signal plus the overall benefit b exceeds the utility without the signal:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] + b \geqslant \widetilde{B}(f_0) \tag{94}$$

Substituting the respective expressions for $\mathbb{E}_s\left[\widetilde{B}(f_s)\right]$ and $\widetilde{B}(f_0)$, we get:

$$\alpha\mu_0 + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 + b \geqslant \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 \tag{95}$$

Simplifying, this reduces to:

$$b \geqslant \frac{1}{2}\alpha^2\lambda\sigma_0^2 \tag{96}$$

Since $\frac{1}{2}\frac{\alpha^2}{\beta^2}\sigma_0^2\lambda > 0$, the DM will only acquire the signal if the overall benefit b is greater than or equal to this term. ***Q.E.D.***

Proof of Theorem 3. The expected belief utility after acquiring the signal can be written as:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] = \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 \tag{97}$$

The utility from not acquiring the signal is simply the belief utility with the prior:

$$\widetilde{B}(f_0) = \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 \tag{98}$$

The DM will choose to acquire the signal if the utility with the signal plus the overall benefit b exceeds the utility without the signal:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] + b \geqslant \widetilde{B}(f_0) \tag{99}$$

Substituting the respective expressions for $\mathbb{E}_s\left[\widetilde{B}(f_s)\right]$ and $\widetilde{B}(f_0)$, we get:

$$\alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 + b \geqslant \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 \tag{100}$$

$$b \geqslant 0 \tag{101}$$

The DM will only acquire the signal if the overall benefit b is greater than or equal to zero. ***Q.E.D.***

Proof of Theorem 4. This is a proof of Theorem 4. Let us start with the self-aware (SA) information-sensitive motivated reasoner. His expected belief utility, $\mathbb{E}_s^{SA}\left[\widetilde{B}(f_s)\right]$, after acquiring the signal can be written as:

$$\mathbb{E}_s^{SA}\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda(\mu_0-\delta) + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2\sigma_0^2 \tag{102}$$

$$= \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 - \alpha\lambda\delta \tag{103}$$

Next, let us start with the unaware (UA) information-sensitive motivated reasoner. His expected belief utility, $\mathbb{E}_s^{UA}\left[\widetilde{B}(f_s)\right]$, after acquiring the signal can be written as:

$$\mathbb{E}_s^{UA}\left[\widetilde{B}(f_s)\right] = \alpha\left(\lambda(\widetilde{\mu}_0 - \delta) + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{104}$$

$$= \alpha\left(\lambda\left(\mu_0 + \frac{1}{2}\alpha\sigma_0^2 - \delta\right) + (1-\lambda)\mu_0\right) + \frac{1}{2}\alpha^2(1-\lambda)\sigma_0^2 \tag{105}$$

$$= \alpha\mu_0 + \frac{1}{2}\alpha^2\sigma_0^2 - \alpha\lambda\delta \tag{106}$$

Thus, both self-aware (SA) and unaware (UA) motivated reasoners have equal expected belief utility after acquiring the signal. This utility is identical to the belief utility in the prior, minus the weighted (i.e., by $\alpha\lambda$) absolute difference between the value of the favorable signal and the unfavorable signal (i.e., δ):

$$\mathbb{E}_s^{SA}\left[\widetilde{B}(f_s)\right] = \mathbb{E}_s^{UA}\left[\widetilde{B}(f_s)\right] = \widetilde{B}(f_0) - \alpha\lambda\delta \tag{107}$$

Both self-aware (SA) and unaware (UA) information-sensitive motivated reasoners will choose to acquire the signal if the utility with the signal plus the overall benefit b exceeds the utility without the signal:

$$\mathbb{E}_s\left[\widetilde{B}(f_s)\right] + b \geqslant \widetilde{B}(f_0) \tag{108}$$

$$\widetilde{B}(f_0) - \alpha\lambda\delta + b \geqslant \widetilde{B}(f_0) \tag{109}$$

$$b \geqslant \alpha\lambda\delta \tag{110}$$

The DM—both self-aware (SA) and unaware (UA) information-sensitive motivated reasoners—will only acquire the signal if the overall benefit b is greater than or equal to $\alpha\lambda\delta$. ***Q.E.D.***

Proof of Corollary 2. Since $b = \ell\lambda\sigma_0^2$, the statement of Theorem 2 turns to the fact that the DM, who is self-aware of his own directional motive, will acquire information with an overall benefit b if and only if $\ell\lambda\sigma_0^2 \geqslant \frac{1}{2}\alpha^2\lambda\sigma_0^2$, which simplifies to $\ell \geqslant \frac{1}{2}\alpha^2$. ***Q.E.D.***

Proof of Corollary 3. Since $b = \ell\lambda\sigma_0^2$, the statement of Theorem 3 turns to the fact that the DM who is unaware of his directional motives when making an information acquisition choice, will acquire information if and only if $\ell\lambda\sigma_0^2 \geqslant 0$, which is always the case. ***Q.E.D.***

Proof of Corollary 4. Since $b = \ell\lambda\sigma_0^2$, the statement of Theorem 4 turns to the fact that the DM who is information-sensitive, disregarding whether he is self-aware or unaware of his directional motives, when making an information acquisition choice, will acquire information if and only if $\ell\lambda\sigma_0^2 \geqslant \alpha\lambda\delta$, which simplifies to $\ell \geqslant \alpha\delta\sigma_0^{-2}$. ***Q.E.D.***

B Examples

Here we consider several examples for the main theorems in Section 3.

Self-aware Motivated Reasoner. Now, let us consider several examples, to illustrate Theorem 2. First, imagine a self-aware citizen facing a choice of consuming media sources during the 2022 Russian full-scale invasion of Ukraine.

Example 1 (Self-Aware Citizen Choosing Media). *Consider a self-aware citizen living in a country where both state-controlled propaganda and independent media are available during the 2022 Russian full-scale invasion of Ukraine. This citizen has a preference for the right of the strong and does not condemn the aggression ($\alpha > 0$). However, they are aware of this directional motive and understand that engaging with independent media might present a more nuanced and potentially unsettling perspective on the invasion compared to state-controlled narratives. When deciding whether to seek information from independent media, the citizen evaluates both the objective and psychological benefits ($\gamma \in (0,1)$). The objective benefit includes acquiring accurate information about the conflict, which may empower them to make informed decisions or take appropriate actions (e.g., publicly disagreeing with the invasion or, conversely, conforming). The psychological cost, however, stems from the potential cognitive dissonance caused by confronting challenging truths that undermine their previously held beliefs. This citizen will choose to read independent media only if the psychological benefits of aligning their worldview with the truth or increasing their informational agency outweigh the discomfort associated with revising their prior beliefs ($b \geq \frac{1}{2}\alpha^2\lambda\sigma_0^2$). Conversely, if the cost of discomfort outweighs the benefit, they may continue consuming state-controlled propaganda that reinforces their preferred narrative.*

Second, consider a self-aware political activist evaluating gender studies report.

Example 2 (Self-Aware Political Activist). *A political activist advocating for gender equality is deeply committed to promoting policies that address gender disparities. They hold strong prior beliefs about the severity of these issues, rooted in both personal experience and ideological commitment ($\alpha > 0$). Recognizing their bias, the activist decides whether to read a comprehensive research report on gender studies that could potentially refine or challenge their understanding of systemic inequities. The activist is aware that engaging with the report might introduce conflicting evidence, such as data suggesting areas where disparities are less pronounced than initially believed. They evaluate the trade-off between the objective utility of acquiring robust, evidence-based knowledge that could improve their advocacy and the psychological cost of encountering data that might challenge their perspective. The activist will decide to read the report if the overall benefit of acquiring new knowledge (b^{obj}) and the satisfaction of becoming a more informed advocate (b^{psy}) outweighs* *the discomfort associated with processing conflicting evidence. If the potential reduction in cognitive consonance is too great, they might forego reading the report to maintain their current belief system.*

Third, have a envision a self-aware journalist investigating ethical dilemmas in reporting.

Example 3 (Self-Aware Journalist). *A journalist covering ethical dilemmas in war reporting is committed to objectivity and public interest but recognizes a personal bias toward sensationalist stories that garner attention and readership ($\alpha > 0$). Aware of this directional motive, they must decide whether to review guidelines from an independent journalistic board that critiques sensationalism and emphasizes balanced reporting. The journalist weighs the psychological cost of confronting their biases against the benefit of improving their professional integrity and public trust. If the anticipated discomfort of recognizing their tendency toward sensationalism is outweighed by the long-term benefits of adhering to ethical standards, they will engage with the guidelines. Otherwise, they might avoid the guidelines, justifying it as unnecessary or impractical, to maintain their current approach to reporting.*

Finally, have a look at a self-aware policy maker assessing a research on migration policy.

Example 4 (Self-Aware Policymaker). *A policymaker deeply concerned about the current level and type of migration ($\alpha > 0$) is tasked with formulating policies to address its socioeconomic impacts. His prior beliefs are shaped by both political commitments and personal views, which incline him to see migration as a straining public resources and culturally hostile phenomenon. Aware of these directional motives, the policymaker considers whether to review an independent research report that provides a nuanced analysis of migrants' positive contributions to economic growth, including evidence that might challenge his initial preconceptions. The policymaker evaluates the trade-off between the objective benefit of acquiring a comprehensive understanding of migration dynamics (b^{obj})—which could improve the effectiveness and fairness of their policies—and the psychological discomfort of confronting data that might contradict their existing beliefs or preferred narrative. The policymaker will decide to engage with the report if the anticipated benefits of crafting better-informed, more balanced policies outweigh the cognitive dissonance and potential political risks associated with revising their stance ($b \geq \frac{1}{2}\alpha^2\lambda\sigma_0^2$). If the psychological cost of integrating conflicting information is perceived as too high, they may avoid the report, favoring sources that align with their prior beliefs to maintain consistency and political alignment.*

Each example illustrates how self-awareness can shape decision-making in scenarios involving directional motives and the acquisition of potentially conflicting information.

Unaware Motivated Reasoner. Now, let us consider several examples, to illustrate Theorem 3. First, imagine a unaware citizen facing a choice of consuming media sources during the 2022 Russian full-scale invasion of Ukraine.

Example 5 (Unaware Citizen Choosing Media)**.** *Consider a citizen living in a country where both state-controlled propaganda and independent media are available during the 2022 Russian full-scale invasion of Ukraine. The citizen holds a preference for the right of the strong and does not condemn the aggression ($\alpha > 0$). However, they are unaware of this directional motive, which implicitly influences their evaluation of media sources. When deciding whether to seek information, the citizen perceives the act of consuming independent media as an opportunity to acquire more precise knowledge about the situation. Since their anticipated belief about the signal is biased toward their motivated belief ($\mathbb{E}_s[s] = \widetilde{\mu}_0$), they are motivated by the idea that new information will confirm their existing worldview or enhance their ability to defend it. The citizen will always choose to acquire information from either independent or state-controlled media as long as the overall benefit b is non-negative.*

Second, consider a unaware political activist evaluating gender studies report.

Example 6 (Unaware Political Activist)**.** *A political activist advocating for gender equality is deeply committed to promoting policies addressing gender disparities. They hold strong prior beliefs about the severity of these issues, rooted in both personal experience and ideological commitment ($\alpha > 0$). Unaware of their directional motive, they believe their evaluations are purely objective. The activist considers reading a comprehensive research report on gender studies, which might contain conflicting evidence. However, because their belief about the signal is influenced by their motivated prior ($\mathbb{E}_s[s] = \widetilde{\mu}_0$), they anticipate that the report will support their worldview or highlight actionable insights consistent with their goals. Given the absence of a threshold for acquiring information, the activist will always choose to engage with the report as long as the overall benefit b is not negative. Unaware of their bias, they are likely to interpret the information through the lens of their preexisting beliefs, reinforcing their commitment regardless of the report's actual content.*

Third, have a envision a unaware journalist investigating ethical dilemmas in reporting.

Example 7 (Unaware Journalist)**.** *A journalist covering ethical dilemmas in war reporting is committed to objectivity and public interest but unknowingly favors sensationalist stories that garner attention and readership ($\alpha > 0$). Unaware of this directional motive, they believe their approach to selecting stories is guided purely by professional standards. The journalist considers reviewing guidelines from an independent journalistic board critiquing sensationalism and emphasizing balanced reporting. Since their anticipated belief* *about the signal is biased by their motivated prior ($\mathbb{E}_s[s] = \tilde{\mu}_0$), they expect the guidelines will validate their current reporting style or provide constructive suggestions aligned with their practices. Because there is no threshold for acquiring information, the journalist will always choose to review the guidelines if the overall benefit b is non-negative. However, their lack of awareness about their bias may lead them to selectively interpret or dismiss recommendations that challenge their existing approach, reinforcing their tendency toward sensationalism.*

Finally, have a look at a unaware policy maker assessing a research on migration policy.

Example 8 (Unaware Policymaker). *A policymaker deeply concerned about the current level and type of migration ($\alpha > 0$) is tasked with formulating policies to address its socioeconomic impacts. Their prior beliefs are shaped by political commitments and personal views, inclining them to see migration as a strain on public resources and a cultural challenge. Unaware of their directional motives, they believe their policy decisions are guided by objective evaluations. The policymaker considers reviewing an independent research report analyzing migration's positive contributions to economic growth. Their anticipated belief about the signal is biased toward their motivated prior ($\mathbb{E}_s[s] = \tilde{\mu}_0$), leading them to expect the report will confirm their stance or offer practical insights that align with their policy agenda. Since there is no threshold for acquiring information, the policymaker will always choose to review the report if the overall benefit b is not negative. However, being unaware of their bias, they may unconsciously interpret the findings to support their existing narrative, thereby limiting the report's potential to challenge or reshape their policies.*

These examples illustrate how unaware motivated reasoners tend to acquire information as long as the benefits are non-negative but interpret the data in a manner consistent with their directional motives, perpetuating their biased beliefs.

Information-Sensitive Motivated Reasoner. Now, let us consider several examples, to illustrate Theorem 4. First, imagine a sensitivity to the asymmetry of feedback citizen facing a choice of consuming media sources during the 2022 Russian full-scale invasion of Ukraine.

Example 9 (Information Sensitive Citizen Choosing Media). *Consider an unaware citizen living in a country where both state-controlled propaganda and independent media are available during the 2022 Russian full-scale invasion of Ukraine. The citizen has a preference for the right of the strong and does not condemn the aggression ($\alpha > 0$). However, they are unaware of this directional motive, believing instead that their views are grounded solely in objective reasoning. When deciding whether to consume independent media, the* *citizen evaluates the potential costs and benefits of obtaining information. Their sensitivity to the asymmetry of information (δ) plays a significant role, as they unconsciously anticipate that "bad" signals (contradicting their beliefs) will decrease their perceived utility, while "good" signals (confirming their beliefs) will increase it. The citizen will engage with independent media if the net benefit (b) of acquiring more precise information outweighs their sensitivity to unfavorable signals, adjusted for the precision of the source ($b \geq \alpha\lambda\delta$). If the psychological cost of encountering unsettling or challenging truths is too high, the citizen may choose to consume state-controlled propaganda that aligns with their directional motives, even if they do not consciously recognize this bias.*

Second, consider a sensitivity to the asymmetry of feedback political activist evaluating gender studies report.

Example 10 (Information Sensitive Political Activist)**.** *A political activist advocating for gender equality holds strong beliefs about the severity of systemic inequities, rooted in personal and ideological commitment ($\alpha > 0$). However, they are unaware of their directional motives, believing instead that their views are entirely evidence-based. When deciding whether to read a comprehensive gender studies report, the activist is unconsciously influenced by their sensitivity to the asymmetry of information (δ). They perceive "good" signals (evidence supporting their belief in significant disparities) as more valuable than "bad" signals (evidence suggesting disparities may be less severe than previously thought). This asymmetry in perceived value affects their decision-making. The activist will review the report if the expected benefit of acquiring information exceeds the anticipated psychological cost of encountering unfavorable signals, adjusted for the precision of the report's findings ($b \geq \alpha\lambda\delta$). Otherwise, they may avoid the report, justifying their decision as a matter of priority or relevance, while unconsciously shielding themselves from potential cognitive dissonance.*

Third, have a envision a sensitivity to the asymmetry of feedback journalist investigating ethical dilemmas in reporting.

Example 11 (Information Sensitive Journalist)**.** *A journalist covering ethical dilemmas in war reporting values objectivity and public interest but unconsciously favors sensationalist stories that attract attention and readership ($\alpha > 0$). Believing their decisions are guided solely by professional standards, they are unaware of their directional motives. When deciding whether to review guidelines from an independent journalistic board critiquing sensationalism, the journalist's sensitivity to the asymmetry of information (δ) influences their choice. They perceive "good" signals (validating their approach) as more valuable than "bad" signals (challenging their practices), leading to a skewed assessment of the guidelines' utility. The journalist will engage with the guidelines if the benefit of improving their professional* *standards exceeds the cost of confronting critiques of their approach, adjusted for the guidelines' precision and their sensitivity to unfavorable signals ($b \geq \alpha\lambda\delta$). If the anticipated cost is too high, they may dismiss the guidelines as irrelevant or overly prescriptive, unconsciously preserving their current practices.*

Finally, have a look at a sensitivity to the asymmetry of feedback policy maker assessing a research on migration policy.

Example 12 (Information Sensitive Policymaker)**.** *A policymaker concerned about migration's socioeconomic impacts holds prior beliefs that migration strains public resources and poses cultural challenges ($\alpha > 0$). Unaware of their directional motives, they perceive their stance as objective and evidence-driven. When deciding whether to review an independent report highlighting migrants' positive contributions to economic growth, the policymaker's sensitivity to the asymmetry of information (δ) affects their choice. They unconsciously view "good" signals (aligning with their belief in migration's challenges) as more valuable than "bad" signals (contradicting their belief). The policymaker will review the report if the net benefit of acquiring detailed insights into migration dynamics exceeds the psychological cost of confronting evidence that challenges their views, adjusted for the report's precision ($b \geq \alpha\lambda\delta$). If the cost is perceived as too great, they may disregard the report, citing time constraints or questioning its relevance, while unconsciously maintaining consistency with their prior beliefs.*

C Choosing (Not) to Acquire the Signal Based on Beliefs: Comparative Statics

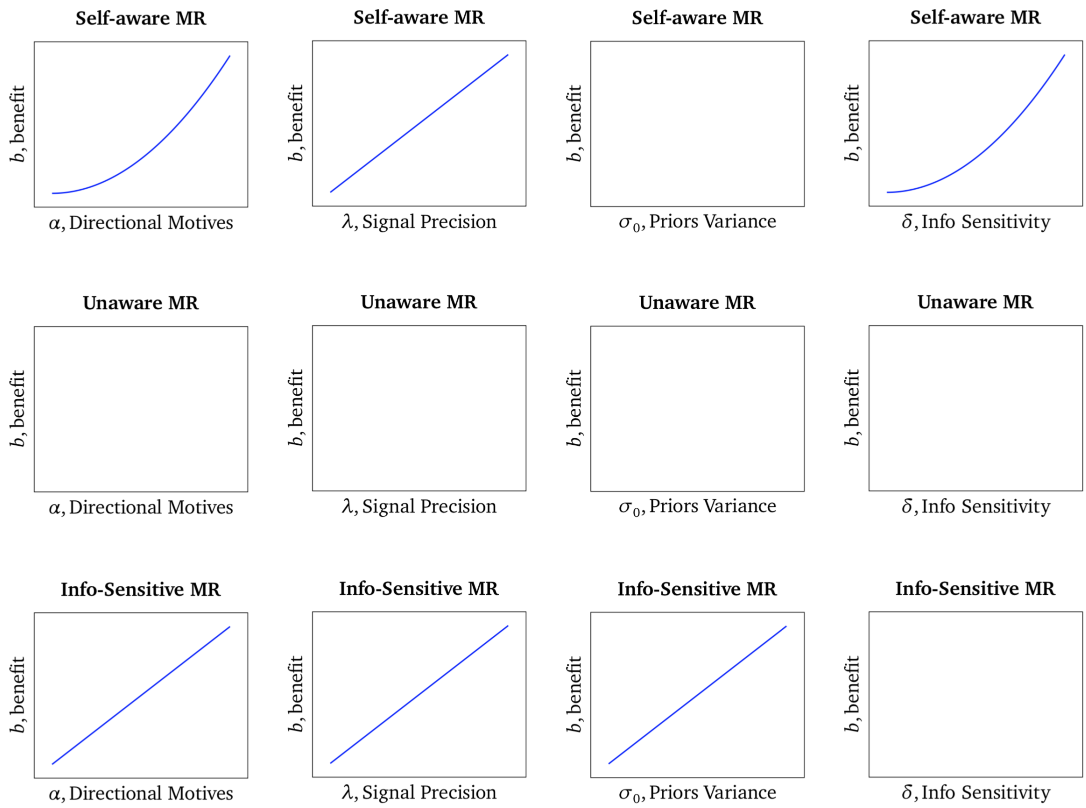

Figure 1: Comparative statics graphs

Note: These graphs depict the conditions presented in eq. (23), eq. (27), and eq. (32) for Self-Aware Motivated Reasoners (MR), Unaware MR, and Information-Sensitive MR, respectively. The empty graphs indicate that the respective variables are not present in the aforementioned equations.

D Choosing (Not) to Acquire the Signal Based on Actions: Comparative Statics

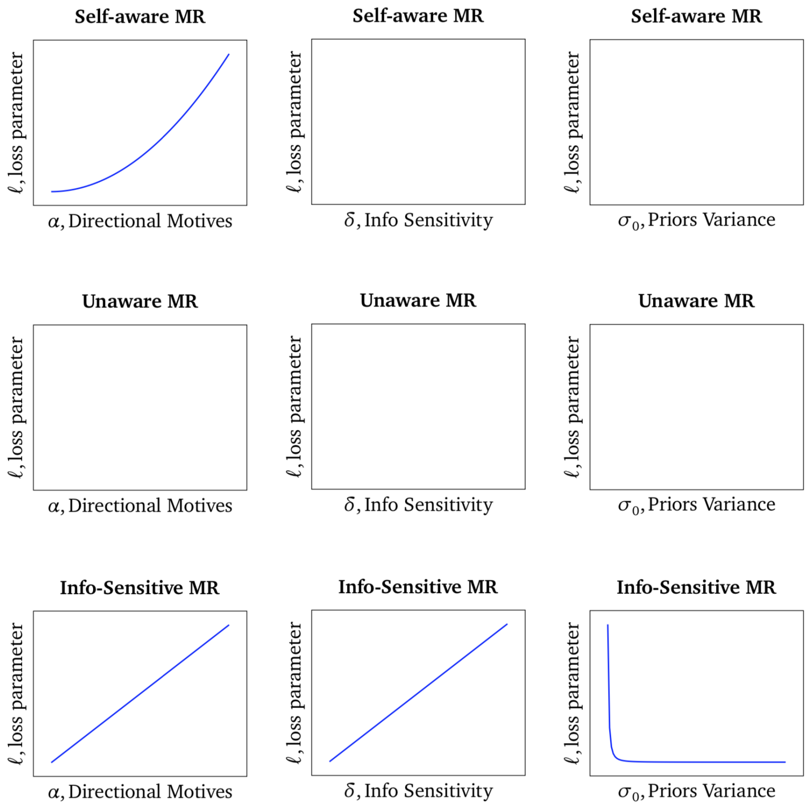

Figure 2: Comparative statics graphs

Note: These graphs depict the conditions presented in eq. (38), eq. (39), and eq. (40) for Self-Aware Motivated Reasoners (MR), Unaware MR, and Information-Sensitive MR, respectively. The empty graphs indicate that the respective variables are not present in the aforementioned equations.