Over- and Underreaction to Information*

Cuimin Ba[†] J. Aislinn Bohren[‡] Alex Imas[§]

November 11, 2022

– Click here to see the most recent version –

Both over- and underreaction to information are well-documented empirically across a variety of domains. This paper explores how key features of the learning environment determine which bias emerges in a given setting. We first develop a two-stage model of belief formation. In the editing stage, limited attention leads the agent to use the *representativeness* heuristic to simplify the learning environment. In the evaluation stage, the agent forms subjective beliefs based on a noisy representation of the edited information structure. This model predicts underreaction when the state space is simple, signals are precise, and the prior is flat or diffuse; it predicts overreaction when the state space is complex, signals are noisy, and the prior is concentrated. A series of experiments provides direct support for these theoretical predictions. As a stark example, increasing the complexity of the state space from two to three states completely *reverses* the direction of the bias from underreaction to overreaction. The results highlight that both stages of belief updating are crucial, in that neither stage on its own can explain the observed patterns in the data. Our framework also rationalizes the disparate findings in prior work: the model predicts the prevalence of underreaction in laboratory studies—which typically use a binary state space, relatively informative signals, and flat priors—as well as the predominance of overreaction documented in financial markets—which feature a more complex state space and noisier signals.

Keywords: overreaction, underreaction, beliefs, noisy cognition, representativeness, behavioral economics, learning

*We thank Nick Barberis, Cary Frydman, David Hirshleifer, Ryan Oprea, Yuval Salant, Marciano Siniscalchi, Florian Zimmermann and seminar participants and conference participants at NBER Asset Pricing meeting, Northwestern, NYU Abu Dhabi, and USC for helpful comments. Josh Hascher, Christy Kang and Enar Leferink provided expert research assistance.

[†]Department of Economics, University of Pennslyvania: cuiminba@sas.upenn.edu
[‡]Department of Economics, University of Pennslyvania: abohren@sas.upenn.edu
[§]Booth School of Business, University of Chicago: alex.imas@chicagobooth.edu

1 Introduction

How do people interpret and react to new information? This question is fundamental across a variety of domains: investors adjust their beliefs about the quality of a stock based on past performance, managers learn from candidate interviews before making hiring decisions, and professional forecasters make economic predictions based on releases of new data. Standard rational models assume that people make statistically optimal use of new information through Bayesian updating with respect to an accurate representation of the signal process. However, a large literature in economics, finance, and psychology has documented systematic departures from this assumption.

Beginning with Edwards (1968), a large stream of research has used laboratory experiments to study learning and belief updating. In these experiments, participants are told about the information structure and report their posterior beliefs about the realization of a hidden state after seeing a signal of varying informativeness. The advantages of these settings are that the components of the information structure—the state space, prior, and signal diagnosticity/precision—are transparent and easy to exogenously manipulate, and it is straightforward to calculate the accurate Bayesian benchmark. As outlined in a recent in-depth review, the evidence from this literature typically documents *underreaction* to information (Benjamin 2019).[1] Another stream of research on financial markets has increasingly relied on surveys and explicit forecasts by households and industry professionals to study how people react to news.[2] In contrast, as noted in the review by Bordalo, Gennaioli, and Shleifer (2022), this literature mostly finds evidence for *overreaction*.[3]

This paper aims to better understand how features of the learning environment impact whether over- or underreaction emerges. We develop a two-stage model of belief-updating that incorporates two well-established psychological frictions—representativeness and noisy cognition—into the same framework. The model predicts underreaction when the state space is simple, signals are precise, and the prior is flat, whereas it predicts overreaction when the state space is more complex, signals are noisy, and the prior is more concentrated. Pre-registered experiments with

[1] Benjamin (2019) summarizes the literature thus: "The experimental evidence on inference taken as a whole suggests that even in small samples, people generally underinfer rather than overinfer."

[2] As discussed further in Section 4, it is important to distinguish between findings on under and overreaction in behavior (e.g. price movements) versus beliefs. Since behavior involves the interaction of preferences and beliefs, one cannot infer if under or overreaction in the former is due to belief-updating or biased preferences. For example, Frazzini (2006) shows that the famous case of underreaction to market news—the post earnings announcement drift (PEAD)—may be due to traders displaying the disposition effect. Given our aim of understanding how people react to information, our paper focuses on work that can isolate belief-updating.

[3] Bordalo, Gennaioli, and Shleifer (2022) write: "The expectations of professional forecasters, corporate managers, consumers, and investors appear to be systematically biased in the direction of overreaction to news."

$N = 1,649$ participants provide direct support for these predictions and highlight the necessity of our unified approach: neither psychological force on its own can explain the scope of our results. Our model and empirical findings can rationalize the discrepancy in prior work, predicting underreaction in laboratory studies—which typically use a simple binary state space, relatively precise signals, and flat priors—and overreaction in financial markets—which feature a richer, more complex state space and noisier signals.

We begin our investigation by developing a two-stage model of belief formation that incorporates an *editing* phase and an *evaluation* phase—where elements of bounded rationality can potentially enter each stage—and derive its implications for belief formation as a function of the information structure. The editing stage determines what elements of the information structure are attended to and to what extent. Research has found that working memory has a fixed capacity, allowing the consideration of a limited number of objects at a time when making judgments (Oberauer, Farrell, Jarrold, and Lewandowsky 2016; Luck and Vogel 1997).[4] In the case of belief formation, this implies that people making likelihood judgments would be limited in the number of potential states they can simultaneously consider. We propose that in the editing stage, the agent uses the *representativeness* heuristic to reduce complexity and simplify the informational environment by channelling attention towards states that are 'representative' of the signal, i.e. that generate the signal with the highest probability (Kahneman and Tversky 1972; Bordalo, Gennaioli, Porta, and Shleifer 2019; Bordalo, Coffman, Gennaioli, and Shleifer 2016). This heuristic distorts the objective Bayesian posteriors by putting the greatest weight on states whose likelihood increases the most relative to other states. In the case of good news or bad news signals (Milgrom 1981), this leads to an overweighting of extreme states.

As an example, take an investor who forms beliefs about a tech company that recently entered the public market. The state space includes the possibility that the firm is a zombie (non-viable and set to crash), a unicorn (e.g. Google, Facebook), as well as a slew of less extreme realizations. Upon seeing a signal, a boundedly-rational investor does not have the capacity to consider all of the states when forming beliefs; she focuses on the states whose likelihood increases the most conditional on the signal. Because unicorns are 'representative' of a price increase, i.e. they are most likely to generate a positive signal, the investor overweights this possibility when forming beliefs following a price increase. Similarly, zombies are 'representative' of a price decrease, and the investor overweights this possibility following a price decrease.

The evaluation stage determines how the (potentially distorted) information structure is processed to form subjective beliefs. A large literature in cognitive psychology has shown that an agent's response can be modeled as optimal subject to a noisy rep-

[4] For example, in the case of visual stimuli, participants can attend to only three to four item's worth of information at any given time (Bays, Gorgoraptis, Wee, Marshall, and Husain 2011)

resentation of the parameters (Green, Swets, et al. 1966; Thurstone 1927; Woodford 2020). A more recent literature in economics has applied the principles of such noisy cognition to explain anomalies in choice under uncertainty (Khaw, Li, and Woodford 2022; Frydman and Jin 2022; Enke and Graeber 2019) and forecasting (Azeredo da Silveira and Woodford 2019). In our framework, the agent processes the edited information structure as if she was facing a signal extraction problem, treating the parameters as unbiased noisy signals of the true underlying values. Greater cognitive noise decreases the agent's sensitivity to changes in the parameters and biases her subjective belief towards a *cognitive default*.

Given this belief-updating framework, we next formalize our definitions of under- and overreaction. We say an agent overreacts to the signal when her subjective posterior expectation of the state moves further from the prior expected state than the objective posterior expectation; an agent underreacts when her subjective expectation moves less than the objective expectation.[5] In order to compare the magnitude of over- and underreaction, we define a 'reaction' measure as the difference between subjective and objective belief movement, normalized by the objective belief movement. The sign of this measure determines whether the agent over- or underreacts and the normalization makes the magnitude of the measure comparable across information structures, allowing us to determine the *extent* of overreaction.

We then derive a series of comparative static predictions on how the direction and magnitude of reaction varies with different properties of the information structure. First, we show that underreaction is expected for a simple state space with relatively informative signals. However, as the complexity of the state space increases—measured as the number of states within a fixed range—the magnitude of underreaction decreases and overreaction emerges as the predominant bias. This prediction is stark: simulations show that even going from two to three potential states leads an agent to switch from underreaction to overreaction for a broad range of parameter values. To see the intuition, consider the case of two symmetric states of equal likelihood and a relatively informative signal. Since the objective posterior is already fairly extreme, the representativeness heuristic does not have much bite and cognitive noise pushes the agent towards underreaction. Critically, adding an interior state decreases the objective weight on the extreme states but representativeness leads the agent to continue to overweight the extreme state that is 'representative' of the realized signal and underweight the other two states; this pushes the agent towards overreact overall. Further increasing complexity increases the magnitude of such overreaction.

A similar logic implies that the magnitude of the reaction will have an inverse U-shaped relationship with respect to signal informativeness: agents will overreact most

[5] We discuss in Section 2 how this definition relates to other common definitions used in the literature.

(underreact least) to relatively uninformative signals and overreact least (underreact most) to relatively informative signals.

The prior distribution over the state space also influences the magnitude of the reaction. Our model predicts that overreaction will increase as the prior becomes more concentrated, in that it places more weight on interior states. This follows from the fact that representativeness overweights more extreme states; the smaller the objective likelihood of those states ex-ante, the greater the potential distortion from the heuristic. On the other hand, if the prior distribution already places substantial weight on the more extreme states, then representativeness plays a smaller role in determining posterior beliefs relative to cognitive noise, pushing the agent towards underreaction.

Finally, in the case of asymmetric prior distributions, the model predicts that agents underreact to confirmatory signals—those that increase the likelihood of states that are more likely ex-ante—and overreact to disconfirmatory, surprising signals. This prediction stems from the base rate neglect that arises due to cognitive noise. Since the prior and the signal are aligned when the latter is confirmatory, base rate neglect leads the agent to underreact; in contrast, the prior and the signal are misaligned when the latter is disconfirmatory, so base rate neglect causes the agent to overreact. Notably, in the case of confirmatory signals, cognitive noise can lead agents to update in the *opposite* direction of the objective posterior. Such wrong-direction updating is not predicted for disconfirmatory signals or for symmetric priors.

We test each of these predictions in a series of pre-registered experiments. We adopted the classic 'bookbag-and-poker-chip' design originally considered in Edwards (1968) and used extensively in the learning literature. In the standard paradigm, two urns are filled with different colored balls with known proportions. For example, Urn A is filled with 70 red balls and 30 blue balls while Urn B is filled with 30 red balls and 70 blue balls. One urn is chosen at random with a known probability, e.g., 0.5. Next, a ball is drawn from it and shown to the participant. The participant then reports her beliefs about the likelihood that Urn A and Urn B were selected, respectively.

Parameters in the design have a straightforward correspondence to informational environment we consider in the model. The urns represent the states, the proportion of balls in each urn represents the signal diagnosticity, while the probability that each urn is selected corresponds to the prior distribution. In the example outlined in the preceding paragraph, the information structure involves two states, Urn A and Ur B, the signal diagnosticity from observing a red ball in state A (B) is 0.7 (0.3), and the prior distribution corresponds to an equal chance that each state is realized.

We adopted this design for several reasons. First, the paradigm allows for a transparent calculation of the objective Bayesian benchmark (e.g. an agent seeing a red ball should update her prior belief that Urn A was selected from 0.5 to a posterior of 0.7.) Second, the parameters of the information structure are straightforward to manipulate in testing the predictions of our model. For example, testing the comparative static on the complexity of the state space can be done by increasing the number of urns. Third, the bookbag-and-poker-chip design has been used extensively in the literature reviewed by Benjamin (2019). Critically, the vast majority of papers used a simple state space (two urns) and, consistent with our framework, have predominantly found underreaction. It would therefore be particularly noteworthy to show substantial evidence for overreaction in the setting where underreaction has thus far been documented as the norm.

Our experiments have three main sources of treatment variation. First, we manipulate the complexity of the state space by expanding from the standard two-state setting to include 3, 5, and 11 potential states. Second, conditional on a level of complexity, we vary the signal diagnosticity from very informative to almost uninformative. Finally, we vary the concentration and symmetry of the prior distribution.

We find that increasing the complexity of the state space has a striking effect on belief-updating. Focusing on uniform priors, we first replicate the standard finding that people generally underreact to information in the simple two-state case. This result flips when the state space expands by even a single urn: the majority of participants *overreact* in the 3-state case, and do so across all levels of signal diagnosticity. The extent of overreaction increases monotonically with the complexity of the state space such that the largest fraction of people overreacting is observed in the 11 state case. Note that this result would not be predicted if people were simply insensitive to the size of the state space; 'representativeness' is key because it highlights exactly *which* states receive more weight.

Moreover, we also document the predicted relationship between signal diagnosticity and belief-updating. The inverse U-shape is observed across all levels of complexity; for example, in the 3 state case, we observe the most overreaction for relatively uninformative signals and the least for relatively informative signals (though overreaction is observed across all diagnosticities).

Next, we examine how the shape of the prior distribution affects people's reaction to information. Consistent with our predictions, the extent of overreaction increases as the prior distribution becomes more concentrated; the most (least) overreaction is observed when interior states have greater (smaller) prior likelihoods than more extreme states. Looking at asymmetric priors in the simple two-state case, we find the hypothesized underreaction to confirmatory signals and overreaction to disconfirmatory signals. That is, even in the setting where people generally underreact, we still find evidence for significant overreaction when the prior distribution is asymmetric. Notably, the fraction of wrong-direction updates varies predictably with the signal type: consistent with the framework, we observe nearly four times as many wrong-direction updates for confirmatory signals than disconfirmatory signals. As we outline formally in Section 2, neither of the psychological mechanisms we consider can explain the full set of results; both stages of the model are required to generate the comparative statics predictions.

Our findings contribute to the literature on belief-updating and learning from information. Section 4 provides an in-depth review of the prior work on over versus underreaction and discusses how our findings can help rationalize some of the core results. Notably, our framework predicts underreaction as the predominant phenomenon in simple settings such as the two-state experiments reviewed in Benjamin (2019). At the same time, it predicts overreaction in more complex environments that involve forming expectations about future returns of a stock or forecasting macroeconomic variables—settings that feature a large number of potential states—which is consistent with the findings reviewed in Bordalo, Gennaioli, and Shleifer (2022). Finally, we discuss how our findings relate to the evidence on how investor behavior (prices) responds to news in financial markets, e.g. short-lag autocorrelations (momentum) (Daniel, Hirshleifer, and Subrahmanyam 1998), price reactions to earnings announcements (Barberis, Shleifer, and Vishny 1998), and macroeconomic news (Klibanoff, Lamont, and Wizman 1998).

The paper also contributes to the literature on cognitive foundations for economic decision-making. One line of work explores the role of complexity in judgment and decision-making. This research argues that people are averse to complexity (Oprea 2020). As a result, individuals adopt simpler mental models (Kendall and Oprea 2021; Molavi 2022) and use heuristics to reduce the mental costs of judgments and decisions (Banovetz and Oprea 2020). Another strand of research models an agent as optimally responding to a stimulus given a noisy representation of the decision problem. These models of noisy cognition have been used to explain phenomena such as small stakes risk aversion (Khaw, Li, and Woodford 2021), state-dependent risk attitudes (Khaw, Li, and Woodford 2022), and myopia in time preferences (Gabaix and Laibson 2017). Awareness of this noise is correlated with the extent of people's insensitivity to the parameters of the decision problem (Enke and Graeber 2019). Our theoretical framework is linked to both lines of work, where the proposed two-stage model of belief-updating incorporates complexity aversion into the editing stage and noisy cognition into the evaluation stage.

Finally, our unified approach of considering multiple forms of bounded rationality at different stages of the judgment process has significant precedent in the psychology and economics literature. Arguably the most prominent model in behavioral economics—prospect theory—combines an editing and evaluation phase (Kahneman and Tversky 1979; Thaler and Johnson 1990). Two-stage models of editing and evaluation have also been applied in finance (Barber and Odean 2008; Akepanidtaworn, Di Mascio, Imas, and Schmidt 2022). We contribute to this literature by showing how the interaction of multiple psychological frictions can add explanatory power in explaining how people learn from information.

The rest of the paper proceeds as follows. Section 2 outlines the theoretical framework and derives the predictions. Section 3 presents the experimental paradigm and the empirical results. Section 4 discusses the prior literature and discusses the findings in the context of our framework. Section 5 concludes.

2 Theoretical Framework

In this section, we formalize a two-stage model of belief formation that combines an 'editing' and 'evaluation' stage. The 'editing' stage guides what factors of the information structure are salient to the agent and to what extent. The 'evaluation' stage determines how this edited input is processed to form a subjective posterior belief after observing the signal. We then define a general measure of over- and underreaction and use this framework to derive a series of comparative static predictions on how the extent of under- or overreaction varies with the complexity of the state space, the precision of the signal, and the shape of the objective prior.

2.1 Model

2.1.1 Information Structure

Consider a finite state space $\Omega \equiv \{\omega_1, ..., \omega_N\} \subset [0, 1]$ with N distinct states in ascending order, i.e. $\omega_1 < ... < \omega_N$, distributed according to prior $p \in \Delta(\Omega)$. A binary signal $s \in \{r, b\}$ provides information about the state. Conditional on state ω, the signal is distributed according to $p(r|\omega) = \omega$ and $p(b|\omega) = 1 - \omega$. Note that in this environment, changing the values assigned to each state changes the signal structure. For example, signal realization r is more likely in state ω_2 for $\Omega = \{0.3, 0.7\}$ relative to $\Omega' = \{0.4, 0.6\}$. Since Ω fully pins down the signal distribution, we refer to the pair (Ω, p) as an *information structure.*

We focus on *symmetric* state spaces, which means that if $\omega \in \Omega$, then $1-\omega \in \Omega$. We say a prior is *symmetric* if for any $\omega \in \Omega$, $p(\omega) = p(1 - \omega)$. Note that prior symmetry implies state space symmetry but not vice versa. We define two structural properties of state spaces. We say state space Ω' is more *extreme* than Ω if the maximum and the minimum states in Ω' are weakly more extreme, $\omega'_1 \leq \omega_1$ and $\omega'_N \geq \omega_N$. We say state space Ω' is more *complex* than Ω if Ω' contains weakly more states than Ω, i.e. $|\Omega'| \geq |\Omega|$. Finally, related to individual states, we say state ω' is more *interior* than state ω if $|\omega' - \frac{1}{2}| \leq |\omega - \frac{1}{2}|$.

Given an information structure, by Bayes rule, the (objective) posterior probability of any state ω following signal realization $s \in \{r, b\}$ is

$$p(\omega|s) = \frac{p(s|\omega)p(\omega)}{\sum_{\omega' \in \Omega} p(s|\omega')p(\omega')}, \tag{1}$$

and the (objective) posterior expected state is

$$E(\omega|s) = \sum_{\omega \in \Omega} p(\omega|s)\omega. \tag{2}$$

This information structure mirrors the experimental environment in Section 3. The two-stage updating process, definitions of under- and overreaction, and analysis is straightforward to extend to richer information structures with more general signals.

2.1.2 Two-Stage Updating Process

An agent forms her subjective belief $\hat{p}(\omega|s)$ using a two-step process. First, she edits the objective posterior belief, where constraints on working memory and attention lead her to use the *representativeness* heuristic to simplify the state space. Second, following a given signal realization, she updates her belief using Bayes rule subject to a noisy representation of the information structure.[6]

Editing. We define the representativeness of state ω given signal realization s by $R(\omega|s) \equiv \frac{p(\omega|s)}{p(\omega)}$. The most representative state is the state whose likelihood increases the most after observing the signal, relative to the prior for that state. In the information structures we consider, the most representative state is one of the extreme states, ω_1 or ω_N. Given signal realization s, the agent's subjective posterior after the first updating step is equal to the objective posterior times a representativeness weight for each state,

$$p_R(\omega|s) \equiv p(\omega|s)\frac{R(\omega|s)^\theta}{Z(s)}, \tag{3}$$

where $\theta \geq 0$ is a parameter capturing the severity of the representativeness distortion and $Z(s)$ is a normalization factor such that $\sum_{\omega \in \Omega} p_R(\omega|s) = 1$. This expression follows Bordalo, Gennaioli, Porta, and Shleifer (2019) and Bordalo, Gennaioli, Ma, and Shleifer (2020) who formalize the represenativeness heuristic in the context of financial markets. When $\theta = 0$, the agent's subjective posterior corresponds to the objective posterior, $p_R(\omega|s) = p(\omega|s)$. Cases where $\theta > 0$ correspond to the agent simplifying the informational environment by channeling greater attention—and therefore inflating the weight on—representative states.[7] Given signal realization

[6]The predictions of our model do not change qualitatively if we switch the order of the stages.

[7]An alternative interpretation of this heuristic is that the agent simplifies the informational environment by "counting" a signal $\theta + 1$ times—a common parameterization of updating with overreaction. To see that these two heuristics are equivalent, note that the odds ratio of any two states ω and ω' is given by

$$\frac{p_R(\omega|s)}{p_R(\omega'|s)} = \left(\frac{p(s|\omega)}{p(s|\omega')}\right)^{\theta+1} \frac{p(\omega)}{p(\omega')}, \tag{4}$$

and the distorted posterior is given by

$$p_R(\omega|s) = \frac{p(s|\omega)^{\theta+1} p(\omega)}{\sum_{\omega' \in \Omega} p(s|\omega')^{\theta+1} p(\omega')}. \tag{5}$$

s, the posterior expectation of the state is given by

$$E_R(\omega|s) = \sum_{\omega \in \Omega} p_R(\omega|s)\omega \tag{6}$$

Evaluation. In the second stage, the agent uses a noisy representation of the simplified information structure from the first stage to form a subjective posterior belief. We assume that her noisy representation of $p_R(\omega|s)$ is normally distributed in each state and following each signal realization,

$$p_R(\omega|s) \sim \mathcal{N}(\overline{p}(\omega), \sigma^2), \tag{7}$$

where $\overline{p}$ is her *cognitive default* prior and $\sigma > 0$.[8] As in the literature on noisy cognition (Enke and Graeber 2019; Khaw, Li, and Woodford 2022), the cognitive default corresponds to an agent's prior about parameters of the decision environment *before* she is given information about those parameters. Here we assume the default prior is the 'ignorance prior' that does not place greater weight on any given state, i.e. the default prior is uniform.[9]

Following the noisy cognition literature (Woodford 2020), given this noisy representation, the agent forms her posterior following signal realization s as if she observed a noisy signal $y(\omega|s)$ about $p_R(\omega|s)$,

$$y(\omega|s) \sim \mathcal{N}\left(p_R(\omega|s), v^2\right), \tag{8}$$

for each state ω, where $v > 0$. From Bayes' rule, the agent's expected subjective posterior conditional on signal $y(\omega|s)$ is

$$\lambda y(\omega|s) + (1-\lambda)\overline{p}(\omega), \tag{9}$$

where $\lambda \equiv \sigma^2/(v^2 + \sigma^2)$. When $\lambda < 1$, the agent biases her posterior towards the cognitive default.[10] As cognition becomes noisier (higher v^2), the agent becomes more reliant on her cognitive default (lower λ).

For our predictions, we focus on the expectation of the subjective posterior with

It is straightforward to see that Eq. (3) and Eq. (5) are equivalent.

[8] As in this literature, we acknowledge the problem of assuming normality, which is that probabilities may be out of the feasible range $[0, 1]$ and the sum of the probabilities may not be exactly 1.

[9] We provide direct evidence this assumption in the context of our experiment in Section Section 3. There, we incentivize elicited beliefs *before* participants are provided with the parameters of the information structure. The 'ignorance prior' is the modal answer and the average is not significantly different from uniform assumption.

[10] This process is akin to the anchoring-and-adjustment heuristic in the judgment and decision-making literature (Tversky and Kahneman 1974), where the agent enters a decision environment with an 'anchor' of $\overline{p}$ and insufficiently adjusts to new information.

respect to $y(\omega|s)$,

$$\hat{p}(\omega|s) \equiv E_y[\lambda y(\omega|s) + (1-\lambda)\overline{p}(\omega)] = \lambda p_R(\omega|s) + (1-\lambda)\overline{p}(\omega). \tag{10}$$

For simplicity, we will refer to this as the subjective posterior. The agent's subjective posterior expected state can also be expressed as a linear combination of the posterior expected state from the first stage and the expected state from the cognitive default,

$$\hat{E}(\omega|s) = \lambda E_R(\omega|s) + (1-\lambda)\overline{E}(\omega), \tag{11}$$

where $\overline{E}(\omega) = \sum_{\omega\in\Omega}\overline{p}(\omega)\omega = 1/2$. Note that when $\theta = 0$ and $\lambda = 1$, the subjective posterior is equal to the objective posterior.

2.1.3 Defining Over- and Underreaction

Given signal realization s, define

$$r(s) \equiv \frac{|\hat{E}(\omega|s) - E(\omega)| - |E(\omega|s) - E(\omega)|}{|E(\omega|s) - E(\omega)|}, \tag{12}$$

where the numerator captures the difference between how far the agent's subjective posterior expected state moves relative to the prior compared to how far the objective posterior expected state moves, and the denominator captures the movement of the objective expected state. We say an agent *overreacts* to signal realization s if her subjective posterior expected state is further from the prior expected state as compared to the objective posterior expected state, and *underreacts* if it is closer.

Definition 1 (Over- and underreaction)**.** *The agent exhibits* overreaction *to* s *if* $r(s) > 0$ *and* underreaction *to* s *if* $r(s) < 0$.

In the case where $\theta = 0$ and $\lambda = 1$, $r(s)$ is equal to 0 for both signal realizations, and the agent neither overreacts or underreacts. Otherwise, $r(s)$ may be non-zero. Its sign potentially varies across signal realizations; the agents can overreact following one signal realization and underreact following the other.

The denominator of $r(s)$ is positive, so this definition is equivalent to saying that the numerator of $r(s)$ is positive or negative. However, the numerator on its own does not provide a valid measure of the magnitude of over and underreaction that is comparable across information structures. Dividing by the expected movement of the objective posterior standardizes the measure so that it is comparable across different information structures.[11]

Definition 1 defines over- and underreaction with respect to the posterior expected

[11] To see why this is necessary, note that when all possible states are close to each other, the numerator of $r(s)$ is naturally small, and the opposite is true if the states are very far apart. In addition, if we double the value of all states, the numerator is automatically doubled. Therefore, the numerator is not a sensible measure of magnitude.

state. This is consistent with both the finance and experimental literatures. The former typically studies asset prices and average forecasts instead of the entire belief distribution. The latter typically compares the movement of posterior beliefs in binary state spaces, which is equivalent to a comparison of posterior expectations.[12]

Finally, we distinguish between updates that move in the same direction versus the opposite direction as the objective update.

Definition 2 (Wrong Direction Updates). *A subjective posterior belief is a* same direction update *at s if $\hat{E}(\omega|s) \leq E(\omega)$ when $E(\omega|s) \leq E(\omega)$ and $\hat{E}(\omega|s) \geq E(\omega)$ when $E(\omega|s) \geq E(\omega)$. Otherwise it is a* wrong direction update. *The subjective posterior belief features* same direction updates *if this holds for both $s \in \{r, b\}$.*

2.2 Predictions

We next derive predictions for whether an agent over- or underreacts to a signal realization based on (i) the complexity of the state space, (ii) the informativeness of the signal, and (iii) the shape of the objective prior. All proofs for this section are in Appendix A.

2.2.1 Benchmark

As a benchmark, we consider the cases where the agent exhibits bias in only one of the belief formation stages, i.e. either $\theta = 0$ or $\lambda = 1$. If the agent uses the representativeness heuristic to process the signal in the editing stage, but does not exhibit noisy cognition in the evaluation stage ($\lambda = 1$), then she overreacts to both signal realizations.

Prediction 1 (Representativeness Only). *If $\theta > 0$ and $\lambda = 1$, the agent* overreacts *to both signal realizations.*

If the agent does not exhibit representativeness in the editing stage ($\theta = 0$), but has noisy cognition in the evaluation stage, then when the prior is symmetric, she underreacts to both signal realizations.

Prediction 2 (Noisy Cognition Only). *If $\theta = 0$ and $\lambda < 1$, then under a symmetric prior the agent* underreacts *to both signal realizations.*

Therefore, observing underreaction is evidence against the agent only using the representativeness heuristic to form beliefs (i.e. $\lambda = 1$) and observing overreaction in an environment with a symmetric prior is evidence against the agent only using noisy cognition to form beliefs (i.e. $\theta = 0$).

[12] When Ω is binary, it can be shown that $r(s) > 0$ if and only if the subjective posterior for any state moves further away from the prior than the objective posterior, i.e. $|\hat{p}(\omega|s) - p(\omega)| - |p(\omega|s) - p(\omega)| > 0$ for all $\omega \in \Omega$.

2.2.2 Bounded Rationality and Belief-Updating

We next explore the interaction between representativeness and noisy cognition and show that it gives rise to a predictable pattern of overreaction and underreaction. We first focus on the case of a symmetric objective prior. In this case, the prior expected state is equal to the cognitive default expected state, which in turn is equal to 1/2, i.e. $E(\omega) = \overline{E}(\omega) = 1/2$. Therefore, it is possible to simplify $r(s)$. From Eqs. (11) and (12),

$$r(s) = \lambda r_R(s) - (1 - \lambda), \tag{13}$$

where

$$r_R(s) \equiv \frac{|E_R(\omega|s) - E(\omega)| - |E(\omega|s) - E(\omega)|}{|E(\omega|s) - E(\omega)|}. \tag{14}$$

Note that $r(s)$ is linear and increasing in both λ and $r_R(s)$. Intuitively, the agent reacts more to s if cognitive noise is smaller (higher λ) and representativeness is stronger (higher $r_R(s)$). More importantly, the agent exhibits overreaction if $r_R(s) > (1 - \lambda)/\lambda$ and underreaction if $r_R(s) < (1 - \lambda)/\lambda$. While $(1 - \lambda)/\lambda$ is a positive constant, $r_R(s)$ can range from 0 to a potentially large number, which depends on both the size of the representativeness parameter θ and the information structure.

We derive three comparative statics of $r(s)$ with respect to the agent's information structure. First, we compare the agent's reaction as a function of *state space complexity* and *signal precision*. Next, we derive comparative statics with respect to *concentration* of a symmetric prior over a fixed state space.

Complexity of the State Space. Fix any two distinct information structures (Ω, p) and (Ω', p') with symmetric state spaces. Let $r(s)$ denote the overreaction ratio to signal realization s for (Ω, p) and $r'(s)$ analogously for (Ω', p'). Our next prediction fixes the extremeness of the state space and derives how overreaction changes as the state space becomes more complex. We show that when the influence of representativeness is sufficiently large, overreaction becomes larger as the complexity of the state space increases.

Prediction 3 (Complexity)**.** *Suppose $\theta > 0$ and $\lambda \leq 1$. Consider two distinct information structures (Ω, p) and (Ω', p') with symmetric state spaces, the same range, and uniform priors. If Ω' is more complex than Ω, and every state in $\Omega' \setminus \Omega$ is more interior than every state in Ω, then for sufficiently large θ, the agent overreacts more in (Ω', p') than (Ω, p) following both signal realizations, $r'(s) > r(s)$ for $s \in \{r, b\}$.*

For example, for any $x \in (0, 0.5)$, the agent reacts more to both signal realizations from information structure (Ω', p') than (Ω, p) when $\Omega = \{x, 1 - x\}$ contains two states and $\Omega' = \{x, 0.5, 1 - x\}$ contains three states or $\Omega' = \{x, y, 1 - y, 1 - x\}$ contains four states with $y \in (x, 0.5)$.

The intuition behind this result is as follows. When the state space becomes more complex and the states are more interior, extreme states become less likely. Therefore, the objective expected state moves less following any signal realization. However, the agent does not fully internalize this change because she puts more weight on the representative states. When θ is very high, the agent's subjective expected state is almost entirely driven by the most representative state, which remains unchanged as the state spaces are equally extreme. Overall, this leads to an increase in movement of the subjective expected state relative to the objective expected state.

Note that Prediction 3 also holds in a representativness-only model (i.e. $\theta > 0$ and $\lambda = 1$), but such a model would predict the agent overreacts in both (Ω, p) and (Ω', p'). In contrast, when $\lambda < 1$, it is possible to have underreaction in (Ω, p) and overreaction in (Ω', p'), or underreaction in both but less underreaction in (Ω', p').

Extremeness of the State Space. We now vary the extreme states, holding fixed the complexity and the set of non-extreme states. The impact on the agent's reaction is more subtle in this case, because moving the extreme states closer together leads to less movement in both the objective expected state and the subjective expected state following both signal realizations. This tension resolves towards a higher $r(s)$ if the former effect dominates. We show that this is the case if

$$W(\Omega) \equiv \sum_{m=1,N} (\omega_m - 1/2)^2 - \sum_{n \neq 1,N} (\omega_n - 1/2)^2 > 0, \tag{15}$$

and $W(\Omega') > 0$. Intuitively, when the extreme states are close to the boundary and the interior states are close to 1/2, the objective posterior belief attaches high probability to one of the extreme states and low probabilities to the others. In this case, the objective expected state is sensitive to the value of ω_1 and ω_N. On the contrary, if $W(\Omega) < 0$ and $W(\Omega') < 0$, the second effect dominates, in which case moving the extreme states inwards results in less overreaction.

Prediction 4 (Extremeness)**.** *Suppose $\theta > 0$, $\lambda \leq 1$. Consider two distinct information structures (Ω, p) and (Ω', p') with symmetric and equally complex state spaces, the same set of non-extreme states, $\omega_n = \omega'_n$ for $n \notin \{1, N\}$, and uniform priors. Suppose Ω' is less extreme than Ω, then for sufficiently large θ:*

(i) If $W(\Omega) > 0$ and $W(\Omega') > 0$, then the agent overreacts more *to both signal realizations in (Ω', p') than (Ω, p), $r'(s) > r(s)$ for all $s \in S$.*

(ii) If $W(\Omega) < 0$ and $W(\Omega') < 0$, then the agent overreacts less *to both realizations in (Ω', p') than (Ω, p), $r'(s) < r(s)$ for all $s \in S$.*

Note that the quantity $W(\Omega)$ tends to be positive when the state space is small because the extreme states naturally have higher weights. For example, we can show that $W(\Omega)$ must be positive if Ω contains no more than 5 states. However, if Ω contains more states, increasing its extremeness may lead to more overreaction.

Signal Structure. Under the conditions of Prediction 4, the prediction implies that $r(s)$ increases as the states become closer together. We show that this comparative statics holds for all values of θ. Moreover, when both $\theta > 0$ and $\lambda < 1$, the agent must underreact to precise signals, and if θ is sufficiently high, the agent overreacts to imprecise signals.

Corollary 1. *Suppose $\theta > 0$, $\lambda < 1$. Consider a binary symmetric state space $\Omega = \{\omega_1, \omega_2\}$ with a uniform prior, then $r(s)$ decreases in ω_2. Moreover, there exists a cutoff $c \in [1/2, 1)$ such that the agent overreacts to all signals if $\omega_2 \in (1/2, c)$ and underreacts to all signals if $\omega_2 \in (c, 1)$. When θ is large enough, c is strictly larger than $1/2$.*

When the state space becomes more complex, signal precision can no longer be summarized by a scalar (i.e. ω_2). We combine Prediction 3 and Prediction 4 to show that if $W(\Omega)$ is positive, then the agent reacts more as the signal has lower diagnosticity—in the sense that *all* states become more interior—when θ is sufficiently large. For example, consider state space $(x, y, 1-y, 1-x)$ with $x \in (0, 0.5)$ and $y \in (x, 0.5)$ and a uniform prior. Then the agent reacts more to signals as x and y move closer to 0.5.

Corollary 2. *Suppose $\theta > 0, \lambda \leq 1$. Consider two distinct information structures (Ω, p) and (Ω', p') with uniform priors and equally complex state spaces such that $W(\Omega') > 0$ and $W(\Omega) > 0$. If ω'_n is more interior than Ω_n for all $n = 1, ..., N$, then for sufficiently large θ, the agent reacts more in (Ω', p') than (Ω, p) following both signal realizations, $r'(s) > r(s)$ for $s \in \{r, b\}$.*

2.2.3 Shape of the Prior

Concentration. Next, consider two symmetric information structures (Ω, p) and (Ω', p') that share the same state space but differ in the priors. We say that p' is more *concentrated* than p if there exists a cutoff $c \in (1/2, 1)$ such that $p'(\omega) \geq p(\omega)$ for all $\omega \in [1-c, c]$ and $p'(\omega) \geq p(\omega)$ for all $\omega \in [0, 1-c] \cup [c, 1]$, and strictly more concentrated if at least one of the inequalities is strict. In words, a more concentrated prior assigns higher probability to the interior states. We show in Prediction 5 that the agent exhibits more overreaction as the prior becomes more concentrated.

Prediction 5 (Prior concentration)**.** *Suppose $\theta > 0, \lambda \leq 1$, and $\Omega = \Omega'$. Then if p' is strictly more concentrated than p, for large enough θ, the agent overreacts more in (Ω', p') than in (Ω, p) to all signal realizations, $r'(s) > r(s)$ for all $s \in S$.*

The intuition behind Prediction 5 is similar to that of Prediction 3. With a more concentrated prior, the objective expected state moves less following a signal realization, but the representativeness heuristic is not sensitive to the shape of the prior. Therefore, the agent exhibits more overreaction. Again, the model with only representativeness predicts overreaction under both priors, but the two-stage model with cognitive noise allows for underreaction under the less concentrated prior or for both priors.

Asymmetry. Finally, we consider information structures with asymmetric priors. For this analysis, we restrict to a binary symmetric state space $\Omega = \{\omega_1, \omega_2\}$ where it is straightforward to manipulate the direction of the prior and the signal. We define whether a signal realization is *confirmatory* or *disconfirmatory* based on its alignment with the prior. For example, if the prior assigns higher probability to ω_1, then a signal realization is confirmatory if it is more likely under ω_1 and disconfirmatory if more likely under ω_2. The formal definition follows:

Definition 3. *Under a binary symmetric state space, a signal realization s is* confirmatory *if either (1) $p(\omega_1) > p(\omega_2)$ and $p(s|\omega_1) > p(s|\omega_2)$, or (2) $p(\omega_1) < p(\omega_2)$ and $p(s|\omega_1) < p(s|\omega_2)$. A signal realization s is* disconfirmatory *if either (3) $p(\omega_1) > p(\omega_2)$ and $p(s|\omega_1) < p(s|\omega_2)$, or (4) $p(\omega_1) < p(\omega_2)$ and $p(s|\omega_1) > p(s|\omega_2)$.*

The model generates a rich set of predictions with under asymmetric priors. With $\lambda < 0$, the agent can overreact to disconfirmatory signals and underreact to confirmatory signals. Moreover, the agent is predicted to update in the wrong direction for relatively uninformative confirmatory signals. This relative pattern does not change when the agent also exhibits representativeness. Moreover, when representativeness is sufficiently strong, the agent may also overreact to a confirmatory signal with intermediate signal diagnosticity.

Prediction 6 (Asymmetric priors)**.** *Suppose $\theta \geq 0$ and $\lambda < 1$. Consider an information structure with a binary state space $\Omega = \{\omega_1, \omega_2\}$ and a fixed prior.*

(i) Suppose s is confirmatory. There exist cutoffs $1/2 < c_1 \leq c_2 \leq c_3 \leq 1$ such that the agent has wrong direction updates *if $\omega_2 \in (1/2, c_1)$,* overreacts *if $\omega_2 \in (c_2, c_3)$, and* underreacts *if $\omega_2 \in (c_1, c_2) \cup (c_3, 1)$. Moreover, (c_2, c_3) is nonempty for sufficiently large θ.*

(ii) Suppose s is disconfirmatory. There exist a cutoff $c_4 \in (1/2, 1)$ such that the agent overreacts *if $\omega_2 \in (1/2, c_4)$ and* underreacts *if $\omega_2 \in (c_4, 1)$.*

To illustrate, let us first consider the case where the agent is not subject to representativeness ($\theta = 0$). Suppose ω_2 is more likely according to the prior and so the prior mean $E(\omega)$ is strictly larger than $1/2$. A relatively uninformative confirmatory signal s increases the objective expected state, $E(\omega|s) > E(\omega)$, but this increase is small. However, noisy cognition pulls the subjective expected state towards the cognitive default $1/2$ with a non-trivial weight of $1 - \lambda$, resulting in a wrong direction update, $\hat{E}(\omega|s) < E(\omega)$. By contrast, following a relatively uninformative disconfirmatory signal s, the objective expected state decreases slightly,

$1/2 < E(\omega|s) < E(\omega)$. By moving towards the cognitive default, the subjective expected state decreases even more, implying overreaction. As signal diagnosticity increases, the signal outweighs the prior in determining the posterior. Since noisy cognition implies not only base-rate neglect but also *signal-diagnosticity neglect*, it now leads to underreaction.

On the other hand, compared to the objective benchmark, representativeness induces the agent to react more to all signals, and more so for relatively uninformative signals. When this force is strong enough, the agent may overreact to a confirmatory signal with an intermediate diagnosticity. As signal diagnosticity approaches 1, similar to our observation in Corollary 1, the distortion induced by representativeness diminishes and the agent underreacts.

3 Empirical investigation

In this section, we directly test the predictions of our framework in a controlled experimental setting. Our experiments test how people's reactions to information depend on the complexity of the state space, informativeness of the signal, and shape of the prior.

3.1 Method

Participants were recruited from the Prolific crowdsourcing platform and a total of 1,649 (48.6% female, 38.9 average age) took part in our experiment.[13] Participants first had to pass an attention check before reading any experimental instructions. Those who did not pass the first attention check did not proceed to the rest of the study; we do not collect data from these participants and they are not included in the participant totals.

After passing an initial screen, participants were told that in addition to the base payment of \$2, they could earn two additional payments as a bonus. First, they would earn \$1 for answering another set of comprehension checks that followed the instructions. Second, they would earn an additional \$10 if their response to a randomly-chosen belief elicitation question was within 3% of the corresponding objective posterior.[14] We used this incentivization procedure as opposed to more complex mechanisms such as quadratic or binarized scoring rules because recent evidence shows that these procedures can systematically bias truthful reporting towards conservatism and underreaction. Danz, Vesterlund, and Wilson (2022) show that the binarized scoring rule leads to conservatism in elicited beliefs and greater error rates compared to simpler mechanisms. The paper argues that incentives based on belief quantiles—like the one we use here—will result in more truthful reporting and lower cognitive burdens.

[13] Preregistration materials can be found here: https://aspredicted.org/LTJ_CS7 and https://aspredicted.org/Q77_3LG.

[14] See Enke and Graeber (2019) for similar use of objective posterior as the incentivized benchmark.

3.2 Design

Participants who passed the first attention check were given the following information about the information structure. Each was told that there is a deck of 100 cards. Every card has the name of a bag written on it, e.g., 'Bag A,' and every bag has 100 red and blue balls. Participants completed a series of trials where they reported their beliefs on the likelihood that each type of card was drawn from the deck.

In each trial, they knew how many cards have the name of each bag written on them and how many red and blue balls were contained in each bag. The computer would randomly draw one card from the deck and show it to the participant. The participant's task was to report how likely they thought that each type of card was drawn, e.g. Bag A, Bag B, etc., by reporting a percentage from 0 to 100, and these percentages must add up to 100.

It is straightforward to see how each parameter in the experiment corresponds to a parameter in the information structure. As illustrated in Fig. 1, which depicts the 3-state case, the number of bags corresponds to the size of the state space. The number of cards for each bag corresponds to the objective prior. The 3-state example depicts a concentrated prior with more probability mass on the interior states (Bag B) and less on the more extreme ones (Bag A and Bag C). Finally, the number of red and blue balls in each bag corresponds to the signal diagnosticity. In the example, seeing a red ball has a .6 diagnosticity of coming from Bag A, a .5 diagnosticity of coming from Bag B, and has a .4 diagnosticity of coming from Bag C.[15]

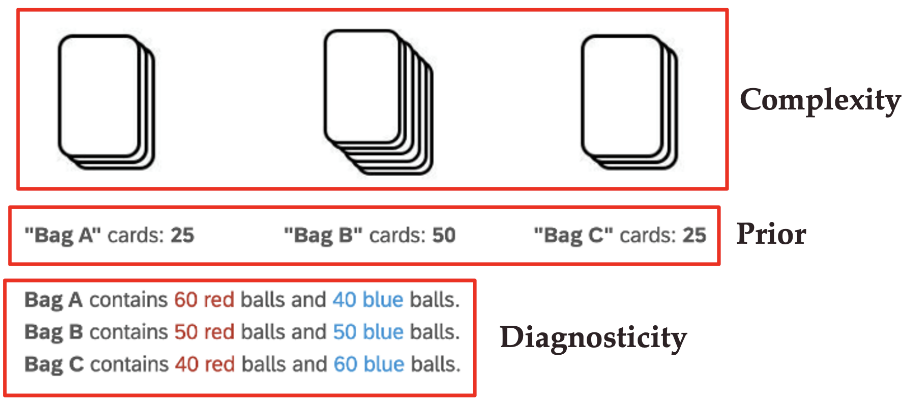


FIGURE 1. Experimental Setup: 3-State Case

The experiments manipulated three factors in the information structure:

- **Complexity:** The number of states in $\Omega \equiv \{\omega_1, ..., \omega_M\}$.
- **Diagnosticity:** For signal $s \in \{r, b\}$, the probability that signal s is drawn given state ω_m, $p(s|\omega_m)$. To simplify notation, we define $d_m \equiv p(r|\omega_m)$ as the probability of a red signal in state ω_m.

[15]The specific instructions can be found in Appendix C.

- **Concentration:** The objective prior probability $p(\omega)$ placed on states more or less informative signals,
- **Symmetry:** The relative probability placed on states that are more likely to generate one signal $s\{r, b\}$ versus the other by the objective prior $p(\omega)$.

Table 1 outlines all of the parameter combinations used in our experiments. After reading the instructions, participants completed a set of comprehension questions. They were then randomized into one of the state space conditions and completed a set number of trials in random order. Each condition had at least 200 participants. The total number of possible trials for each condition was equal to the product of the unique prior distributions, the unique diagnosticities, and the unique signals (which was always set at 2). The maximum number of trials was capped at 15. After completing all the trials, participants answered a set of basic demographic questions and exited the study.

Defining Over and Underreaction: Our main dependent variables compare participants' responses to the objective posteriors. In every trial, we calculate a) the participant's expectation of drawing a red ball given their reported beliefs, b) the prior expectation of drawing a red ball, and c) the objective posteriors for each state. We then compute three metrics. Using these measures, we compute the ratio r that corresponds to Eq. (12) in Section 2. A positive (negative) r represents over (under) reaction in that trial. We use r as our primary measure of over and underreaction in the main text. Per our pre-registration, we exclude trials in which participants update in the wrong direction. Appendix B replicates the analyses with wrong direction updates included; the results do not meaningfully change.

TABLE 1. Experiment Parameters

STATE SPACE	PRIORS	DIAGNOSTICITIES
2	$p(\omega_1) \in \{0.3, 0.5, 0.7\}$ $p(\omega_2) = 1 - p(\omega_1)$	$d_1 \in \{0, 0.1, 0.2, 0.3, 0.4,$ $0.49, 0.51, 0.6, 0.7, 0.8, 0.9, 1\}$ $d_2 = (1 - d_1)$
3	$p(\omega_1) = 0.33$ $p(\omega_2) = 0.34$ $p(\omega_3) = 0.33$	$d_1 \in \{0.6, 0.7, 0.8, 0.9\}$ $d_2 = 0.5$ $d_3 = (1 - d_1)$
3	$p(\omega_1) = 0.25$ $p(\omega_2) = 0.50$ $p(\omega_3) = 0.25$	$d_1 \in \{0.6, 0.7, 0.8, 0.9\}$ $d_2 = 0.5$ $d_3 = (1 - d_1)$
3	$p(\omega_1) = 0.4$ $p(\omega_2) = 0.2$ $p(\omega_3) = 0.4$	$d_1 \in \{0.6, 0.7, 0.8, 0.9\}$ $d_2 = 0.5$ $d_3 = (1 - d_1)$

Table 1. Experiment Parameters

State Space	Priors	Diagnosticities
5	$p(\omega_m) = 0.2$ $\forall m \in M$	$(d_1, d_2 \in \{(60, 55), (70, 55), (80, 55), (90, 55), (70, 60), (80, 70), (80, 60), (90, 80), (90, 70), (90, 60)\}$ $d_3 = 0.5$ $d_4 = (1 - d_2)$ $d_5 = (1 - d_1)$
11	$p(\omega_m) = (1/11)$ $\forall m \in M$	$d_m = \frac{m-1}{10}$ $\forall m \in M$

3.3 Results

We begin by examining how complexity of the state space and signal diagnosticity impacts people's reactions to information. We then proceed to test our predictions on the shape of the prior, both with regard to its concentration and symmetry.

3.3.1 Complexity and Diagnosticity

Table 2 and Fig. 2 present the results on people's belief updating as a function of complexity and diagnosticity. As shown in Fig. 2, we replicate the underreaction result in the simple 2 state case: people update their beliefs significantly less than the objective benchmark across all signal diagnosticities (one-sample t-test against 0, $p < .001$). This is consistent with the evidence outlined in Benjamin (2019) that shows underinference from signals in a host of experiments using a similar paradigm as our own.

TABLE 2. Belief-Updating by Complexity and Diagnosticity

	(1) Overreaction Ratio	(2) Overreaction Ratio
3 States	0.342***	0.174***
	(0.0358)	(0.0356)
5 States	0.331***	0.235***
	(0.0381)	(0.0389)
Uninformative		0.682***
		(0.0588)
3 States * Uninformative		-0.339***
		(0.0471)
5 States * Uninformative		-0.351***
		(0.0511)
Constant	-0.0822***	-0.0796***
	(0.0253)	(0.0304)
N	5293	5293
adj. R^2	0.040	0.075

Standard errors clustered at the individual level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

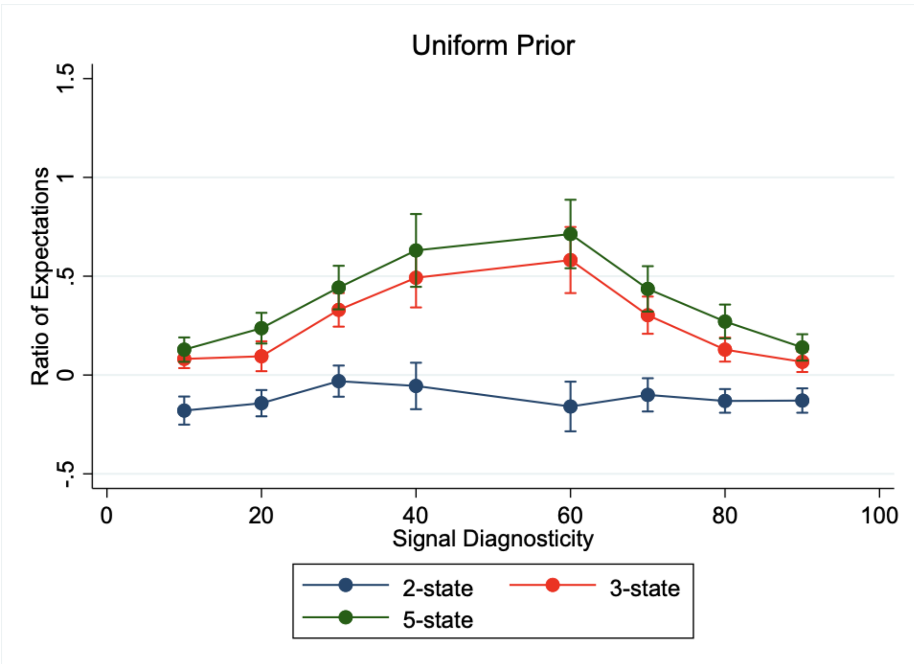


FIGURE 2. Belief-updating by complexity with uniform priors

Increasing the complexity of the state space completely reverses this result. In the 3, 5, and 11-state case we observe significant significant deviations from the objective updating benchmark, $r \neq 0$ (one-sample t-test against 0, $p < .001$). But the effect goes in the opposite direction as the 2-state case: people display overreaction across all signal diagnosticities. Column 1 in Table 2 compares the 3 and 5-state case to the 2-state case, showing that increasing complexity generates substantially more overreaction. Fig. 3 presents results using a discrete overreaction measure that captures the frequency of deviations from the objective benchmark in the direction of either over or underreaction, trial by trial. Specifically, this metric takes the fraction of trials with overreaction and the fraction of trails with underreaction, and then computes the difference. Positive values indicate the prevalence of overreaction and negatve values indicate the prevalence of underreaction. We again see that people tend to underreact in the 2-state but switch to overreacting when the complexity increases to more states. These results provide strong support for Prediction 3.

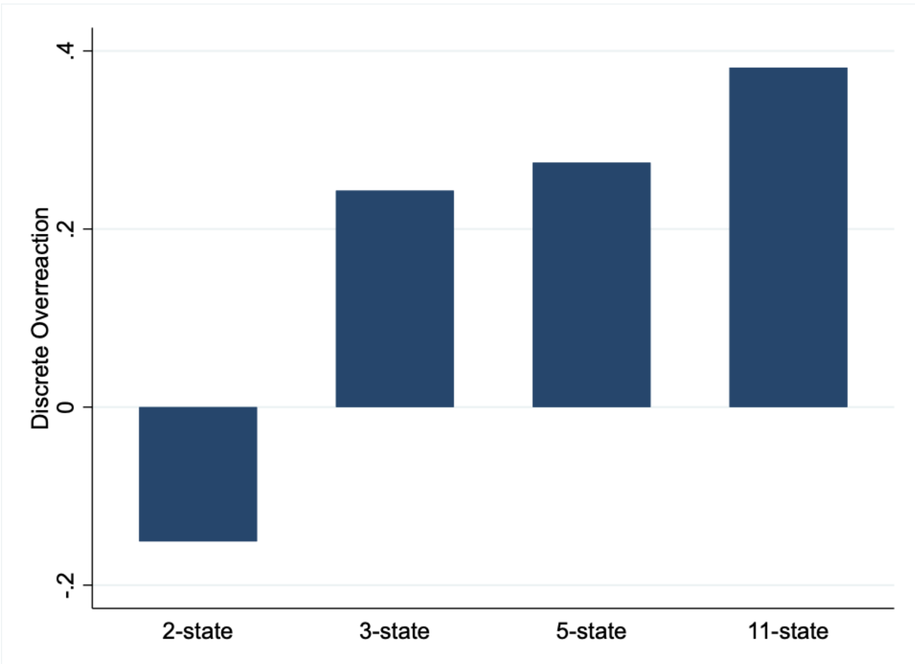


FIGURE 3. Discrete Overreaction by M

Table 2 tests the related prediction in Corollary 1 on the informativeness of signals. We create a dummy variable that correspond to Uninformative signals if the diagnosticity is between 0.3 and 0.7. Consistent with the prediction, coefficients in Column 2 show that there is less underreaction in the 2-state case for noisier signals than for more precise signals. The coefficients on the interaction terms between the number of states and signal informativeness show that noisier signals also generate significantly more overreaction in the 3 and 5-state cases, though the relative impact of noisier signals is smaller than in the 2-state case.

3.3.2 Concentration of Prior

Next, we examine how the shape of the prior affects belief-updating. Our experiments focus on the 3-state for simplicity, though the predictions hold for any number of states greater than 2 (where the prior concentration cannot shift by definition). Table 3 and Fig. 4 present the results. Consistent with Prediction 5, we observe substantially more overreaction as the prior becomes more concentrated. Table 3 Column 1 compares belief-updating for concentrated and diffuse priors relative to the uniform case. People overreact significantly more when the prior is concentrated and significantly less when the prior is diffuse.

TABLE 3. Belief-Updating by Concentration and Diagnosticity

	(1) Overreaction Ratio	(2) Overreaction Ratio
Concentrated	0.213***	0.127***
	(0.0547)	(0.0414)
Diffuse	-0.215***	-0.216***
	(0.0321)	(0.0232)
Uninformative		0.505***
		(0.0651)
Concentrated * Uninformative		-0.171***
		(0.0594)
Diffuse * Uninformative		-0.00325
		(0.0395)
Constant	0.260***	0.0941***
	(0.0253)	(0.0187)
N	4026	4026
adj. R^2	0.048	0.108

Standard errors clustered at the individual level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

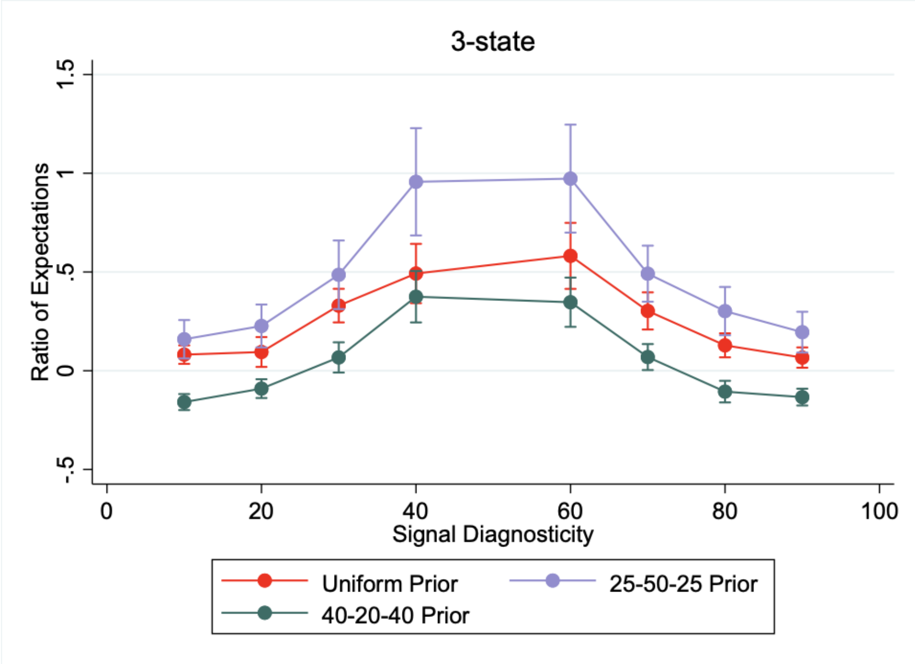


FIGURE 4. Belief-updating by concentration (3-states)

Note that while the ratio $r > 0$ in all three cases, it is no longer significant when the prior is diffuse. As shown in Fig. 4 and Column 2 of Table 3, We observe significantly less overreaction for more precise signals across all three shapes of the prior. Looking at the diffuse case, we actually see significant *underreaction* when the signals are precise enough. Together, these findings provide strong support for our Prediction 5.

3.3.3 Confirmatory versus Disconfirmatory Information

We now proceed to examine belief-updating when the signal is confirmatory versus disconfirmatory with respect to the objective prior. Fig. 5 presents the ratio of expectations r in the 2-state case when aggregating across all three sets of priors (i.e. symmetric and asymmetric). We again see significantly more underreaction for more precise signals, but also some overreaction for noisier signals. However, aggregating across priors masks significant heterogeneity with respect to whether the signal was congruent with the objective prior (confirmatory) or not (disconfirmatory).

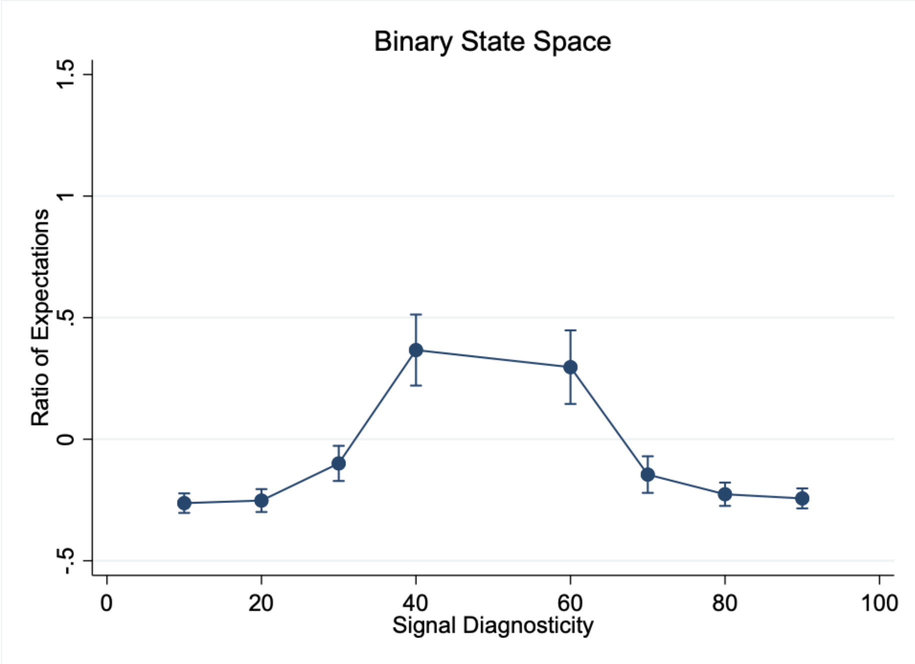


FIGURE 5. Belief-Updating with 2-states Across all Priors

Table 4 Column 1 presents regressions for whether a signal was confirmatory or disconfirmatory in the case of asymmetric priors relative to the uniform case. Consistent with Prediction 6 we see more underreaction for confirmatory signals and significant *overreaction*—even in the 2-state case—when signals are disconfirmatory.

TABLE 4. Belief-Updating by Signal Type and Diagnosticity

	(1) Overreaction Ratio	(2) Overreaction Ratio
Confirm	-0.0645	0.0510
	(0.0587)	(0.0776)
Disconfirm	3.831***	5.715***
	(0.305)	(0.482)
Uninformative		-5.379***
		(0.486)
Confirm * Uninformative		0.444***
		(0.0877)
Disconfirm * Uninformative		4.927***
		(0.479)
Constant	-0.0822***	-0.0796***
	(0.0254)	(0.0304)
N	3661	3661
adj. R^2	0.092	0.143

Standard errors clustered at the individual level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Fig. 6 presents the same results graphically: we see substantially more overreaction for disconfirmatory signal (right side of red curve, left side of green curve) than for confirmatory signals (left side of red curve, right side of green curve). People indeed appear to overreact more to surprising news compared to news that is expected.

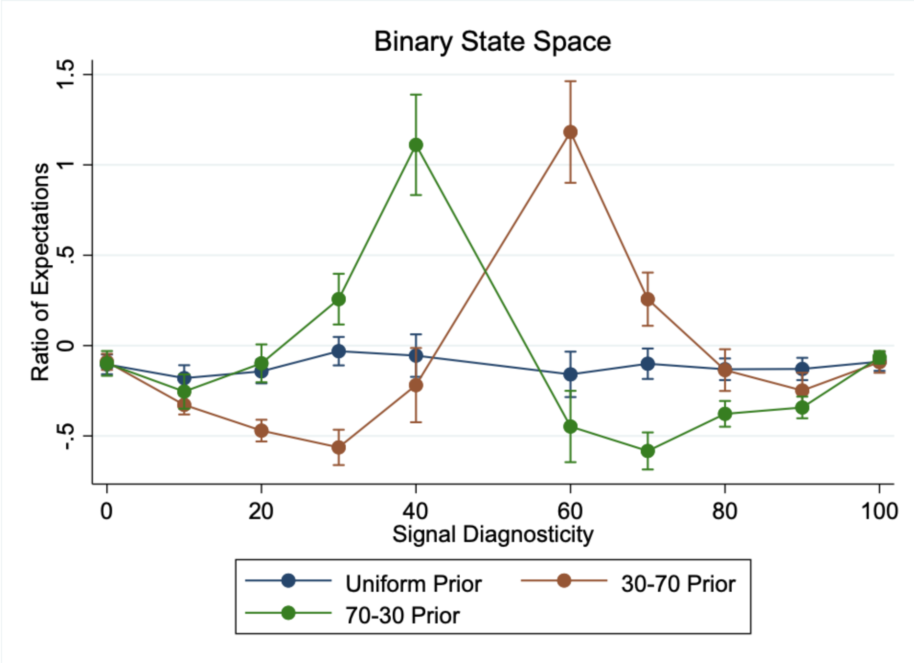


FIGURE 6. Belief-Updating with 2-States by Prior

Finally, we explore the prediction that people are more likely to update in the direction opposite of the objective benchmark, i.e. wrong direction updates, for confirmatory signals compared to discofirmatory ones. Fig. 7 presents the fraction of wrong direction updates by whether the signal was confirmatory or disconfirmatory given the shape of the prior. Consistent with Prediction 6, we see a substantial discrepancy in frequencies. While wrong direction updates occur relatively infrequently in the case of the uniform prior (approximately 10%), they occur directionally less for disconfirmatory, surprising signals than for confirmatory signals. In the latter case, nearly 30% of updates are in the direction opposite the objective benchmark. Importantly, this high incidence of wrong direction updates is not arbitrary noise (e.g. inattentive subjects), but is actually predicted by our model as a function of the decision environment.

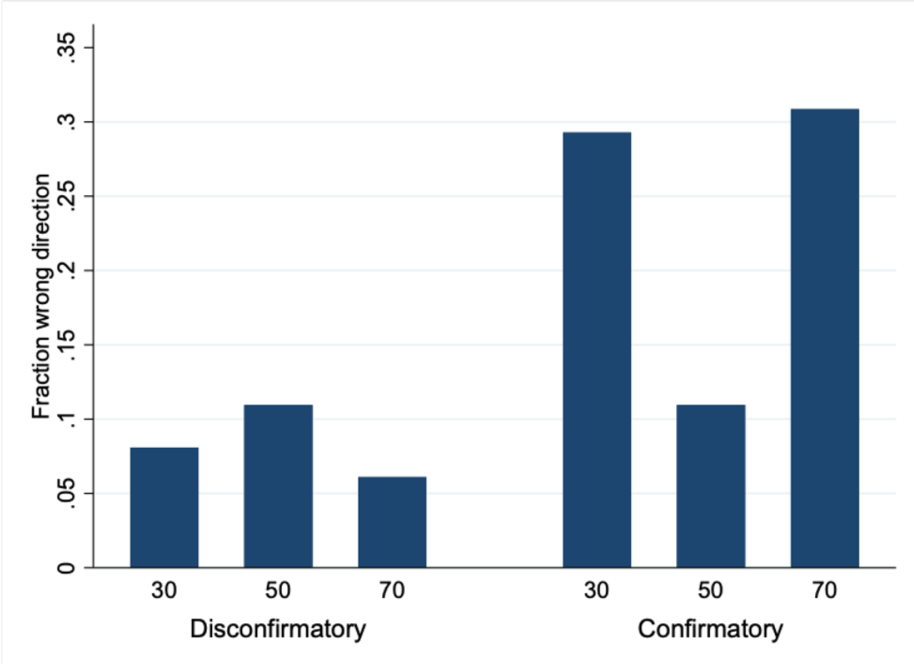


FIGURE 7. Wrong Direction Updates by Prior

4 Over and underreaction in prior work

In this section, we synthesize the empirical findings of over and underreaction in laboratory studies and the research in macroeconomics and financial markets. Importantly, we restrict attention to settings where agents observe one signal from the information structure and updates her beliefs accordingly.[16]

Laboratory studies. Biases in belief updating have been extensively studied in both psychology and economics. Edward's pioneering study jump-started a paradigm of bookbag-and-poker-chip laboratory experiments to examine such biases (Edwards 1968). To categorize the relevant experimental studies, we use the analyzing framework of Benjamin (2019) and talk about signals $s \in \{a, b\}$ which are drawn i.i.d. with diagnosticities $p(a|A) > 0.5$ in state A and $p(b|B) > 0.5$ in state B. Then, the posterior odds ratio of state A compared to state B could be represented as

$$\frac{\hat{p}\,(A|s)}{\hat{p}\,(B|s)} = \left[\frac{p\,(s|A)}{p\,(s|B)}\right]^c \left[\frac{p\,(A)}{p\,(B)}\right]^d, \tag{16}$$

[16]A number of papers explore belief-updating after seeing a sequence of signals from the same information structure. For example, in Griffin and Tversky (1992), participants see multiple signal draws and report their beliefs after seeing the entire sequence. The authors argue that people focus too much on the strength of evidence (e.g. sample proportions) and not enough to the weight (e.g. sample size). In Massey and Wu (2005), participants observe signal draws from an *unstable* information structure, where the signal generating process can change with a certain probability. People tend tend to neglect the possibility of a regime shift, which leads to either over- or underreaction depending on the stability of the system and precision of the signals.

with $c = d = 1$ if the decision maker makes accurate Bayesian inferences. This model has been widely used to analyze deviations from the accurate Bayesian benchmark.[17] Taking the logarithm of the posterior odds in (16) yields the canonical *Grether regressions* (Grether 1980).

The vast majority of experimental studies that use this basic framework find that people on average underreact to new information.[18] Benjamin (2019) presents a meta-analysis incorporating data from a multitude of experiments with a binary state space, symmetric signal diagnosticities ($p(a|A) = p(b|B)$), and a uniform prior. When restricted to experiments where subjects observe only one signal, which we do in the current work, Benjamin estimates that $\hat{c} = 0.70$; this implies that subjects underinfer from signals. Only a small share of this experimental literature finds that individuals overreact on average, and when they do, the paradigm features more than two states (Hartzmark, Hirshman, and Imas 2021; Fan, Liang, and Peng 2021).

A great number of studies have made effort to identify the factors that affect how much people underinfer, among which the most well-studied is the signal diagnosticity. Consistent with our predictions and empirical results, several studies have found that people exhibit greater underreaction when receiving signals with higher diagnosticities. When signal diagnosticities are symmetric, the more informative the signal becomes (i.e. $p(a|A)$ and $p(b|B)$ increase) the more extreme underinference becomes. For example, Edwards (1968) experiments with a uniform prior and signal diagnosticities of $\{0.55, 0.7, 0.85\}$. When the signal was less diagnostic ($p(a|A) = p(b|B) = 0.55$) subjects exhibited overinference; as the diagnosticity increased, they exhibited underinference.[19] Kieren and Weber (2020) find underreaction to informative signals and overreaction to totally uninformative signals, and that the direction of the update depends on the valence of the signal. Benjamin (2019) notes that nearly correct inference or overinference is found in Peterson, DuCharme, and Edwards (1968); DuCharme (1970); Gustafson, Shukla, Delbecq, and Walster (1973), where the signal is generated from a normal distribution and is close to its expected value. Fixing the signal diagnosticities, a larger sample size is found to lead to more underinference.[20]

Our results also reconcile the findings of many papers that experiment with vary-

[17]This model implicitly assumes that the distorted use of priors is independent from the use of signal likelihood ratios. Note that the representativeness heuristic corresponds to $c > 1, d = 0$, but the noisy cognition heuristic is not nested in this model.

[18]See Benjamin (2019) for a comprehensive review.

[19]Similar monotone patterns are documented in Phillips and Edwards (1966); Peterson, Schneider, and Miller (1965); Kahneman and Tversky (1972); Grether (1992); Holt and Smith (2009); Benjamin (2019). When signal proportions are asymmetric, a similar pattern holds: agents tend to overinfer when $p(a|A)$ and $p(b|B)$ are close together (and thus close to 0.5) and underinfer when they are further apart. See Peterson, Schneider, and Miller (1965); Ambuehl and Li (2018).

[20]See Peterson, Schneider, and Miller (1965); Green, Halbert, and Robinson (1965); Peterson and Swensson (1968); Peterson, DuCharme, and Edwards (1968); Sanders (1968); Kahneman and Tversky (1972); Griffin and Tversky (1992).

ing levels of priors. The majority of relevant studies find evidence for base-rate neglect.[21] In a meta-analysis, Benjamin (2019) estimates that $\hat{d} = 0.60$, which is smaller than 1. On the other hand, whether base-rate neglect leads to under or over-reaction is less clear-cut. Holt and Smith (2009) vary the prior odds for a two-urn case and show that there is underreaction for single draws. In line with our experimental results, they show that when priors were more extreme, and a disconfirmatory signal was drawn, subjects overreacted; in other cases, such as a confirmatory signal, or as priors became less extreme, they underreacted. Kieren, Müller-Dethard, and Weber (2022) find that in both experiments and financial market data, investors systematically overreact to new disconfirming information. Among other studies that find an insensitivity to objective priors, Green, Halbert, and Robinson (1965) and Grether (1992, Study 3) find underreaction, and Griffin and Tversky (1992, Study 2) find underreaction in the case of one-shot draws.[22]

The key contribution of our paper to the experimental literature is to explicitly consider how the complexity of the state space affects individuals' reaction to news. The literature has primarily focused on experiments with a binary number of states (urns), and research that varies the size of state space is scant. For example, Phillips and Edwards (1966) conduct two additional experiments to the one mentioned above by Edwards (1968) in which the state space consisted of 10 urns. However, all r urns had p red chips and q blue chips, while the other $(10 - r)$ bags share the same inverse proportions. Thus, this experiment is equivalent to varying the prior, and it is likely subjects did not see it as a meaningful expansion in state-space such as in our experiment. Recently, Fan, Liang, and Peng (2021) find that people underinfer when making inferences and overinfer when forming forecasts. It is straightforward to rationalize their experimental results in our framework. In their main treatment, subjects were shown a normally distributed signal (the stock price growth of a firm this month). They find that over half of the subjects underreacted when they were asked to report their posterior about a binary state (whether the firm is in good or bad condition); and over half overreacted when they were asked to report their expectation about the next signal (the stock price growth next month). Our framework provides a simple explanation—the state to be forecasted is binary in the inference task but substantially more complex in the forecast task since the next signal can be a range of potential numbers. In addition, the signal is more informative about the state than about the next-period signal, which also pushes towards underreaction in the first task and overreaction in the second.[23]

[21] Two exceptions that find over-use of priors include (Peterson, Schneider, and Miller 1965; Grether 1992, Study 2). Consistent with our framework, these studies find both under and overreaction depending on the diagnosticities of asymmetric signals.

[22] Robalo and Sayag (2018) find that with 60/40 priors, subjects' posteriors were, depending on the degree of base-rate neglect, close to the correct benchmark or exhibited overinference.

[23] Our model can explain their findings in the other treatments as well. In the *Cross-variable Forecast* treatment, the authors replaced the forecast variable with a binary variable (whether

Prat-Carrabin and Woodford (2022) find underreaction in an information structure that has a continuous state space $[0, 1]$. While the state space is complex, the finding of underreaction is consistent with our model because the state space contains extreme states such as 0 and 1, the signal is relatively precise, and the prior is flat. As illustrated in Section 2, the comparative statics on the complexity of the state space rely on the assumption that the additional states are interior. While $[0, 1]$ is more complex, the flat prior means that the objective posterior still places substantial weight on the extreme states, leaving less room for representativeness to generate overreaction. Consistent with this, we find the least amount of overreaction in the 3-state case when the state space contains extreme values, $\{0.1, 0.5, 0.9\}$. It is thus consistent with our findings that the addition of even more extreme states, combined with a flat prior, in Prat-Carrabin and Woodford (2022) generated underreaction.[24]

Our paper contributes to the literature on theories that have attempted to explain underreaction in the laboratory studies reviewed here. Phillips and Edwards (1966) propose that people suffer from the *conservatism* bias: they underweight the likelihood ratio of the signal and underinfer from its informational content. This corresponds to belief updating as in (16) with $c < 1$ and $d = 1$. A related concept is *extreme-belief aversion*, which refers to the aversion to holding beliefs close to certainty (Benjamin, Rabin, and Raymond 2016). As pointed out by DuCharme (1970), both conservatism and extreme-belief aversion can lead to under-reaction when individuals receive signals that are strongly indicative of one state, but with the conservatism bias under-reaction should also occur when the objective posterior is not as extreme. DuCharme (1970) devise two experiments with normal signals and show that under-reaction is small or non-existent when the objective posterior is close to the uniform prior, which is consistent with extreme-belief aversion but not conservatism. Underreaction could also be a consequence of base-rate neglect (Kahneman and Tversky 1972), captured by (16) with $d < 1$.[25] The literature on noisy

revenue growth is good or bad next month) that takes the value of 100 if the state is good and 0 if the state is bad. The subjects still exhibited overreaction when they were instructed to report their expectation about the new variable. Note that although this variable takes only two values, its expectation could still be a range of potential numbers; the complexity of the state space thus explains the finding of overreaction. In the *Obvious Connection* treatment and the *Binary Signal* treatment, the forecast variable was replaced by two other binary-valued variables, but the subjects were asked to report their posterior about one of the two possible value (rather than expectation), and thus the state spaces here were binary. Consistent with our model, the authors have found in these two treatments that the fractions of under and overreacting subjects were similar across the inference and forecast tasks—they found primarily underreaction in both.

[24]Our model predicts that either removing the more extreme states in Prat-Carrabin and Woodford (2022) or making the prior more concentrated would lead to overreaction in their setting.

[25]Relatedly, Massey and Wu (2005) study situations where individuals are asked to determine if the regime has shifted based on the observed signals and find that people suffer from a bias they call *system-neglect*. In particular, they find that individuals pay more attention to signals but neglect diagnosticity and transition probability. It leads to a predictable pattern of under and overreaction: individuals underreact in unstable environments with precise signals and overreact in stable environments with noisy signals.

cognition develops a model of decision making in which an agent's response is optimal subject to a noisy representation of the parameters (Woodford 2020).[26] Noisy cognition immediately generates both underreaction and base-rate neglect because subjects only respond partially to variation in the objective posterior. While these theories can explain the underreaction documented in previous laboratory studies with two states and relatively precise signals, none of them can reconcile our results on overreaction with complex information structures and concentrated priors.

Belief updating in financial markets. In contrast to the pervasive experimental findings of underreaction, an extensive literature in finance and macroeconomics suggests that people tend to overreact to information.

A growing literature use surveys and explicit forecasts by professionals and households to study the departures from rational expectations. A common approach is to examine the predictability of forecast errors from forecast revisions (Coibion and Gorodnichenko 2015).[27] Bordalo, Gennaioli, Ma, and Shleifer (2020) analyze the time series data on a large group of financial and macro variables and individual forecasts from professionals. Their analysis finds that except for short-term interest rates (fed funds and 3-month treasury-bill rate) and unemployment rate, a number of financial and macro variables display overreaction, including five-year and ten-year treasury-bill rates, several corporate bond rates, nominal GDP, real GDP, industrial production, CPI, real consumption, real federal government expenditures, and real state and local government expenditures. These findings suggest that forecasts with longer term structures tends to exhibit overreaction. Many other studies produce supporting evidence. For example, d'Arienzo (2020); Wang (2021) find that individual analysts' forecasts of long-term interest rates exhibit overreaction, and Bordalo, Gennaioli, Porta, and Shleifer (2019) also find overreaction in the expectations of long-term corporate earnings growth. This literature is reviewed in Bordalo, Gennaioli, and Shleifer (2022) who summarize it with "the expectations of professional forecasters, corporate managers, consumers, and investors appear to be systematically biased in the direction of overreaction to news."

Our framework and empirical results can potentially reconcile the seemingly contradictory findings in the lab versus financial markets. The prominent feature of real-world settings is that decision makers tend to face a much larger state space (e.g. the positive segment of the real line) and much noisier signals than they would in laboratory settings; our framework predicts that we should observe overreaction, which is consistent with the data. On the other hand, as noted above, laboratory studies have tended to use simple, binary state spaces, flat priors, and relatively

[26] A similar formulation could be found in Epstein, Noor, Sandroni, et al. (2010). They consider the implication of such under-reaction on a decision maker's asymptotic learning outcomes.

[27] See Augenblick and Rabin (2021) for a new statistical test of behavior that deviates from the accurate Bayesian benchmark.

informative signals; our framework would predict underreaction, which is consistent with the findings from that literature.

One thing to note is that we focus on studies that collect data on beliefs, whether by eliciting them directly or through forecasts and surveys; we do not attempt to address prior work on over and underreaction using behavioral data such as prices. A rich literature starting with Ball and Brown (1968) and De Bondt and Thaler (1985) has examined over and underreaction to information by looking at price movements. Prices have been found to adjust slowly to firm-specific (Ball and Brown 1968) and macro (Klibanoff, Lamont, and Wizman 1998) announcements, and display short-term autocorrelation (i.e. momentum); these effects have been interpreted as underreaction (Hirshleifer, Lim, and Teoh 2009; Daniel, Hirshleifer, and Subrahmanyam 1998). Prices also display long-term reversal, i.e. negative autocorrelation, which has been interpreted as overreaction (Barberis, Shleifer, and Vishny 1998). However, it is not clear whether these phenomena are driven by preferences or beliefs. For example, Frazzini (2006) shows that the slow price adjustment to earnings announcements—the famous post earnings announcement drift (PEAD)—is consistent with the disposition effect, which has been explained through prospect theory preferences (Barberis 2012; Heimer, Iliewa, Imas, and Weber 2021). Recent work by Charles, Frydman, and Kilic (2021) shows that noisy cognition can weaken the transmission from beliefs to behavior, such that overreaction in the former can still lead to underreaction in the latter. Since our paper is focused on belief-updating, we do not attempt to apply our framework to findings on over or underreaction in behavior since the effects on beliefs cannot be identified.

5 Conclusion

In this paper we examine the incidence and underlying drivers of over and underreaction. We develop a two-stage model of belief-updating where the first stage determines what elements of the information structure are attended to and the second evaluates this 'edited' information structure with some cognitive noise. This simple model predicts that as the information structure becomes more complex, attention will be channeled towards 'representative' states and people will be more likely to overreact. Overreaction will also be more prevalent for concentrated priors and 'surprising', disconfirmatory signals. We test these predictions in a stylized setting where the relevant features of the information structure can be manipulated exogenously. We first replicate the well-known finding of underreaction to information in the simple 2-state setting used in prior work. However, consistent with our predictions, overreaction is the predominant response when the information structure becomes more complex, the prior is more concentrated, and signals are surprising given the shape of the prior.

Given the complexity of information structures found in real-world settings, our results suggest that overreaction in beliefs may be the predominant phenomenon across a myriad of settings. This is consistent with the findings from the finance literature that primarily documents overreaction to information (Bordalo, Gennaioli, and Shleifer 2022).

Our findings also point to the benefits of studying judgment and decision making as an interaction between multiple psychological processes. While the majority of papers in psychology and behavioral economics have focused on identifying the implications of a single psychological mechanism, it is likely the case that observed judgments and choice are the product of multiple mechanisms interacting. We believe modeling and testing more 'unified' frameworks across economically-important domains to be a fruitful research agenda.

References

Akepanidtaworn, K., R. Di Mascio, A. Imas, and L. Schmidt (2022): "Selling Fast and Buying Slow: Heuristics and Trading Performance of Institutional Investors," *Journal of Finance*.

Ambuehl, S., and S. Li (2018): "Belief updating and the demand for information," *Games and Economic Behavior*, 109, 21–39.

Augenblick, N., and M. Rabin (2021): "Belief movement, uncertainty reduction, and rational updating," *The Quarterly Journal of Economics*, 136(2), 933–985.

Azeredo da Silveira, R., and M. Woodford (2019): "Noisy memory and over-reaction to news," in *AEA Papers and Proceedings*, vol. 109, pp. 557–61.

Ball, R., and P. Brown (1968): "An empirical evaluation of accounting income numbers," *Journal of accounting research*, pp. 159–178.

Banovetz, J., and R. Oprea (2020): "Complexity and procedural choice," *American Economic Journal: Microeconomics*.

Barber, B. M., and T. Odean (2008): "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *The review of financial studies*, 21(2), 785–818.

Barberis, N. (2012): "A model of casino gambling," *Management Science*, 58(1), 35–51.

Barberis, N., A. Shleifer, and R. Vishny (1998): "A model of investor sentiment," *Journal of financial economics*, 49(3), 307–343.

Bays, P. M., N. Gorgoraptis, N. Wee, L. Marshall, and M. Husain (2011): "Temporal dynamics of encoding, storage, and reallocation of visual working memory," *Journal of vision*, 11(10), 6–6.

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

Benjamin, D. J., M. Rabin, and C. Raymond (2016): "A model of nonbelief in the law of large numbers," *Journal of the European Economic Association*, 14(2), 515–544.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016): "Stereotypes," *The Quarterly Journal of Economics*, 131(4), 1753–1794.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in macroeconomic expectations," *American Economic Review*, 110(9), 2748–82.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019): "Diagnostic expectations and stock returns," *The Journal of Finance*, 74(6), 2839–2874.

Bordalo, P., N. Gennaioli, and A. Shleifer (2022): "Overreaction and Diagnostic Expectations in Macroeconomics," *Journal of Economic Perspectives*, 36(3), 223–44.

Charles, C., C. Frydman, and M. Kilic (2021): "Insensitive Investors," *Available at SSRN.*

Coibion, O., and Y. Gorodnichenko (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105(8), 2644–78.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998): "Investor psychology and security market under-and overreactions," *the Journal of Finance*, 53(6), 1839–1885.

Danz, D., L. Vesterlund, and A. J. Wilson (2022): "Belief elicitation and behavioral incentive compatibility," *American Economic Review.*

d'Arienzo, D. (2020): "Maturity increasing overreaction and bond market puzzles," *Available at SSRN 3733056.*

De Bondt, W. F., and R. Thaler (1985): "Does the stock market overreact?," *The Journal of finance*, 40(3), 793–805.

DuCharme, W. M. (1970): "Response bias explanation of conservative human inference.," *Journal of Experimental Psychology*, 85(1), 66.

Edwards, W. (1968): "Conservatism in human information processing," *Formal representation of human judgment.*

Enke, B., and T. Graeber (2019): "Cognitive uncertainty," Discussion paper, National Bureau of Economic Research.

Epstein, L. G., J. Noor, A. Sandroni, et al. (2010): "Non-bayesian learning," *The BE Journal of Theoretical Economics*, 10(1), 1–20.

Fan, T. Q., Y. Liang, and C. Peng (2021): "The Inference-Forecast Gap in Belief Updating," .

Frazzini, A. (2006): "The disposition effect and underreaction to news," *The Journal of Finance*, 61(4), 2017–2046.

Frydman, C., and L. J. Jin (2022): "Efficient coding and risky choice," *The Quarterly Journal of Economics*, 137(1), 161–213.

Gabaix, X., and D. Laibson (2017): "Myopia and discounting," .

Green, D. M., J. A. Swets, et al. (1966): *Signal detection theory and psychophysics*, vol. 1. Wiley New York.

Green, P. E., M. H. Halbert, and P. J. Robinson (1965): "An experiment in probability estimation," *Journal of Marketing Research*, 2(3), 266–273.

GRETHER, D. M. (1980): "Bayes rule as a descriptive model: The representativeness heuristic," *The Quarterly journal of economics*, 95(3), 537–557.

——— (1992): "Testing Bayes rule and the representativeness heuristic: Some experimental evidence," *Journal of Economic Behavior & Organization*, 17(1), 31–57.

GRIFFIN, D., AND A. TVERSKY (1992): "The weighing of evidence and the determinants of confidence," *Cognitive psychology*, 24(3), 411–435.

GUSTAFSON, D. H., R. K. SHUKLA, A. DELBECQ, AND G. W. WALSTER (1973): "A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, Delphi groups, and nominal groups," *Organizational behavior and human performance*, 9(2), 280–291.

HARTZMARK, S. M., S. D. HIRSHMAN, AND A. IMAS (2021): "Ownership, learning, and beliefs," *The Quarterly Journal of Economics*, 136(3), 1665–1717.

HEIMER, R., Z. ILIEWA, A. IMAS, AND M. WEBER (2021): "Dynamic inconsistency in risky choice: Evidence from the lab and field," *Available at SSRN 3600583*.

HIRSHLEIFER, D., S. S. LIM, AND S. H. TEOH (2009): "Driven to distraction: Extraneous events and underreaction to earnings news," *The Journal of Finance*, 64(5), 2289–2325.

HOLT, C. A., AND A. M. SMITH (2009): "An update on Bayesian updating," *Journal of Economic Behavior & Organization*, 69(2), 125–134.

KAHNEMAN, D., AND A. TVERSKY (1972): "Subjective probability: A judgment of representativeness," *Cognitive psychology*, 3(3), 430–454.

——— (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263–292.

KENDALL, C., AND R. OPREA (2021): "On the Complexity of Forming Mental Models," .

KHAW, M. W., Z. LI, AND M. WOODFORD (2021): "Cognitive imprecision and small-stakes risk aversion," *The review of economic studies*, 88(4), 1979–2013.

——— (2022): "Cognitive Imprecision and Stake-Dependent Risk Attitudes," .

KIEREN, P., J. MÜLLER-DETHARD, AND M. WEBER (2022): "Can Agents Add and Subtract when Forming Beliefs? Evidence from the Lab and Field," *Proceedings of Paris December 2021 Finance Meeting EUROFIDAI - ESSEC*.

KIEREN, P., AND M. WEBER (2020): "Expectation Formation under Uninformative Signals," *Available at SSRN 3971733*.

KLIBANOFF, P., O. LAMONT, AND T. A. WIZMAN (1998): "Investor reaction to salient news in closed-end country funds," *The Journal of Finance*, 53(2), 673–699.

LUCK, S. J., AND E. K. VOGEL (1997): "The capacity of visual working memory for features and conjunctions," *Nature*, 390(6657), 279–281.

MASSEY, C., AND G. WU (2005): "Detecting regime shifts: The causes of under-and overreaction," *Management Science*, 51(6), 932–947.

Milgrom, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, pp. 380–391.

Molavi, P. (2022): "Simple Models and Biased Forecasts," *arXiv preprint arXiv:2202.06921*.

Oberauer, K., S. Farrell, C. Jarrold, and S. Lewandowsky (2016): "What limits working memory capacity?," *Psychological bulletin*, 142(7), 758.

Oprea, R. (2020): "What makes a rule complex?," *American economic review*, 110(12), 3913–51.

Peterson, C. R., W. M. DuCharme, and W. Edwards (1968): "Sampling distributions and probability revisions.," *Journal of Experimental Psychology*, 76(2p1), 236.

Peterson, C. R., R. J. Schneider, and A. J. Miller (1965): "Sample size and the revision of subjective probabilities.," *Journal of Experimental Psychology*, 69(5), 522.

Peterson, C. R., and R. G. Swensson (1968): "Intuitive statistical inferences about diffuse hypotheses," *Organizational Behavior and Human Performance*, 3(1), 1–11.

Phillips, L. D., and W. Edwards (1966): "Conservatism in a simple probability inference task.," *Journal of experimental psychology*, 72(3), 346.

Prat-Carrabin, A., and M. Woodford (2022): "Imprecise Probabilistic Inference from Sequential Data," .

Robalo, P., and R. Sayag (2018): "Paying is believing: The effect of costly information on Bayesian updating," *Journal of Economic Behavior & Organization*, 156, 114–125.

Sanders, A. (1968): "Choice among bets and revision of opinion," *Acta Psychologica*, 28, 76–83.

Thaler, R. H., and E. Johnson (1990): "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice," *Management science*, 36(6), 643–660.

Thurstone, L. L. (1927): "A law of comparative judgment.," *Psychological review*, 101(2), 266.

Tversky, A., and D. Kahneman (1974): "Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty.," *science*, 185(4157), 1124–1131.

Wang, C. (2021): "Under-and Overreaction in Yield Curve Expectations," *Available at SSRN 3487602*.

Woodford, M. (2020): "Modeling imprecision in perception, valuation, and choice," *Annual Review of Economics*, 12, 579–601.

A Proofs

Proof of Prediction 1. Suppose the signal realization s is r. Note that

$$p(\omega|s) = \frac{\omega}{\sum_{\omega' \in \Omega} \omega' p(\omega')} p(\omega),$$

which is a re-scaling of beliefs using increasing weights $\frac{\omega}{\sum_{\omega' \in \Omega} \omega' p(\omega')}$. Therefore, the posterior $p(\omega|s)$ first-order stochastically dominates $p(\omega)$. It follows that $E(\omega|s) - E(\omega|s) > 0$. Analogously, we have $p_R(\omega|s)$ first-order stochastically dominates $p(\omega|s)$, because

$$p_R(\omega|s) = \frac{\omega^{\theta+1} p(\omega)}{\sum_{\omega' \in \Omega} \omega'^{\theta+1} p(s|\omega')} = \frac{\omega^{\theta} \sum_{\omega'' \in \Omega} \omega'' p(s|\omega'')}{\sum_{\omega' \in \Omega} \omega'^{\theta+1} p(s|\omega')} p(\omega|s),$$

and the weight increases in ω. Therefore, $E_R(\omega|s) - E(\omega) > E(\omega|s) - E(\omega) > 0$. Analogously, when the signal s is b, we can show that $E_R(\omega|s) - E(\omega) < E(\omega|s) - E(\omega) < 0$. When $\lambda = 1$, $\hat{E}(\omega|s) \equiv E_R(\omega|s)$. Therefore, $r(s) > 0$ for all $s \in S$. $\square$

Proof of Prediction 2. When $\theta = 0$, $\hat{E}(\omega|s) = \lambda E(\omega|s) + (1-\lambda)\overline{E}(\omega) = \lambda E(\omega|s) + (1-\lambda)E(\omega)$, where $\overline{E}(\omega) = E(\omega)$ follows from prior symmetry. So $r(s) = \lambda - 1 < 0$. $\square$

Proof of Prediction 3. Suppose the signal realization s is r. The objective posterior of any state ω in Ω is

$$p(\omega|s) = \frac{\overline{p}(\omega)\omega}{\sum_{\tilde{\omega} \in \Omega} \overline{p}(\tilde{\omega})\tilde{\omega}} = \frac{2}{N}\omega$$

We can write the Bayesian expected state as

$$E(\omega|s) = \sum_{\omega \in \Omega} p(\omega|s)\omega = \frac{2}{N} \sum_{\omega \in \Omega} \omega^2$$

Suppose Ω contains an even number of states and $N = 2K$, then

$$\begin{aligned}
E(\omega|s) - E(\omega) &= \frac{2}{N} \sum_{\omega \in \Omega} \omega^2 - \frac{1}{2} \\
&= \frac{2}{N} \left[(1 - \omega_N)^2 + ... + (1 - \omega_{K+1})^2 + \omega_{K+1}^2 + ... + \omega_N^2 - \frac{K}{2} \right] \\
&= \frac{2}{N} \left[\left(\omega_{K+1} - \frac{1}{2} \right)^2 + ... + \left(\omega_N - \frac{1}{2} \right)^2 \right].
\end{aligned}$$

When Ω contains an odd number of states and $N = 2K - 1$, symmetry implies that the K-th state must be $\frac{1}{2}$. We therefore obtain the same expression for $E(\omega|s) - E(\omega)$.

On the other hand,

$$E_R(\omega|s) = \sum_{\omega\in\Omega} p_R(\omega|s)\omega = \sum_{\omega\in\Omega} \frac{\overline{p}(\omega)\omega^{\theta+2}}{\sum_{\tilde{\omega}\in\Omega}\overline{p}(\tilde{\omega})\tilde{\omega}^{\theta+1}} = \frac{\sum_{\omega\in\Omega}\omega^{\theta+2}}{\sum_{\omega\in\Omega}\omega^{\theta+1}}.$$

Note that $E_R(\omega|s)$ converges to the most representative state as θ grows to infinity. That is, $\lim_{\theta\to\infty} E_R(\omega|s) = \omega_N$. It follows that

$$\begin{aligned}\lim_{\theta\to\infty} r_R(s) + 1 &= \lim_{\theta\to\infty} \frac{|E_R(\omega|s) - E(\omega)|}{|E(\omega|s) - E(\omega)|} \\ &= \frac{\omega_N - \frac{1}{2}}{\frac{4}{N}\left[\left(\omega_{K+1} - \frac{1}{2}\right)^2 + ... + \left(\omega_N - \frac{1}{2}\right)^2\right]}.\end{aligned} \tag{17}$$

Analogously for $r'(s)$. Since Ω' is equally extreme as Ω, $\omega'_N = \omega_N$. Since Ω' is more complex than Ω and every state in $\Omega' \setminus \Omega$ is more interior than every state in Ω,

$$\frac{4}{N'}\left[\left(\omega'_{K+1} - \frac{1}{2}\right)^2 + ... + \left(\omega'_{N'} - \frac{1}{2}\right)^2\right] < \frac{4}{N}\left[\left(\omega_{K+1} - \frac{1}{2}\right)^2 + ... + \left(\omega_N - \frac{1}{2}\right)^2\right].$$

Therefore, when θ is sufficiently large, it follows from Eq. (13) that $r'(s) > r(s)$. The proof is analogous for signal b. □

Proof of Prediction 4. As in the proof of Prediction 3, we have

$$\begin{aligned}\lim_{\theta\to\infty} r_R(s) + 1 &= \lim_{\theta\to\infty} \frac{|E_R(\omega|s) - E(\omega)|}{|E(\omega|s) - E(\omega)|} \\ &= \frac{\omega_N - \frac{1}{2}}{\frac{2}{N}\left[\left(\omega_{K+1} - \frac{1}{2}\right)^2 + ... + \left(\omega_N - \frac{1}{2}\right)^2\right]},\end{aligned}$$

and analogously for $r'(s)$. Fixing $\omega_{K+1}, ..., \omega_{N-1}$, the above expression is increasing in ω_N if $\left(\omega_N - \frac{1}{2}\right)^2 < \left(\omega_{K+1} - \frac{1}{2}\right)^2 + ... + \left(\omega_{N-1} - \frac{1}{2}\right)^2$, i.e. $W(\Omega) < 0$, and decreasing in ω_N if $\left(\omega_N - \frac{1}{2}\right)^2 > \left(\omega_{K+1} - \frac{1}{2}\right)^2 + ... + \left(\omega_{N-1} - \frac{1}{2}\right)^2$, i.e. $W(\Omega) > 0$. □

Proof of Corollary 1. When the state space is binary,

$$\begin{aligned}r_R(s) + 1 &= \frac{|E_R(\omega|s) - E(\omega)|}{|E(\omega|s) - E(\omega)|} \\ &= \frac{\frac{(1-\omega_2)^{\theta+2}+\omega_2^{\theta+2}}{(1-\omega_2)^{\theta+1}+\omega_2^{\theta+1}} - 1/2}{(1-\omega_2)^2 + \omega_2^2 - 1/2} \\ &= \frac{\omega_2^{\theta+1} - (1-\omega_2)^{\theta+1}}{2(\omega_2 - 1/2)((1-\omega_2)^{\theta+1} + \omega_2^{\theta+1})}.\end{aligned}$$

Therefore, $r_R(s)$ is decreasing in ω_2 if and only if $f(x) \equiv \frac{(x-1/2)((1-x)^{\theta+1}+x^{\theta+1})}{x^{\theta+1}-(1-x)^{\theta+1}}$ is increasing in x when $x > 1/2$. Differentiate,

$$f'(x) = \frac{x^{\theta+1}(x^{\theta+1} - (\theta+1)(1-x)^{\theta}) - (1-x)^{\theta+1}((1-x)^{\theta+1} - (\theta+1)x^{\theta})}{(x^{\theta+1} - (1-x)^{\theta+1})^2}.$$

Note that the numerator can be written as $g(x) - g(1-x)$, where $g(x) \equiv x^{\theta+1}(x^{\theta+1} - (\theta+1)(1-x)^{\theta})$. Since $g(x)$ is increasing in x, it follows that $f'(x) > 0$, and thus $r_R(s)$ is decreasing in ω_2.

When ω_2 approaches 1, we have $\lim_{\omega_2 \to 1} r_R(s) + 1 = 1$. By Eq. (13), $r(s)$ converges to $-(1-\lambda)$, which is negative when $\lambda < 1$. When ω_2 approaches 1/2, by the L'Hospital's Rule, we have

$$\lim_{\omega_2 \to 1/2} r_R(s) + 1 = \lim_{\omega_2 \to 1/2} \frac{(\theta+1)(\omega_2^{\theta} + (1-\omega_2)^{\theta})}{2((1-\omega_2)^{\theta+1} + \omega_2^{\theta+1})} = \theta + 1.$$

Therefore, $\lim_{\omega_2 \to 1} r(s) = \lambda(\theta+1) - (1-\lambda) > 0$ when θ is sufficiently large. □

Proof of Corollary 2 When Ω and Ω' contain no more than three states, Corollary 2 immediately follows from Prediction 4. When $|\Omega| = |\Omega'| \geq 4$, we construct $\Omega'' = \{\omega_1, \omega_2', ..., \omega_{N-1}', \omega_N\}$. By Prediction 3, we know that the agent reacts more under Ω'' than under Ω. Since Ω' is less extreme than Ω'' and $W(\Omega'') > W(\Omega') > 0$, by Prediction 4, we know that the agent reacts more under Ω' than under Ω''. It then follows that the agent reacts more under Ω' than under Ω. □

Proof of Prediction 5. Suppose p' is strictly more concentrated than p and both are symmetric. Since the priors have the same support, $E_R'(\omega|s)$ coincides with $E_R(\omega|s)$ when θ diverges to infinity. Thus, to show that $r'(s) > r(s)$ when θ is sufficiently large, it suffices to show that $|E'(\omega|s) - E(\omega)| < |E(\omega|s) - 1/2|$.

Suppose the signal realization is r, then $E'(\omega|s) > 1/2$ and $E(\omega|s) > 1/2$. It suffices to show $E'(\omega|s) < E_p(\omega|s)$. Let $\Delta_p(\omega) = p'(\omega) - p(\omega)$, then $\Delta_p(\omega) \geq 0$ for $\omega \in [1-c, c]$ and $\Delta_p(\omega) \leq 0$ for $\omega \in [0, 1-c] \cup [c, 1]$, and at least one inequality is strict. We have

$$\begin{aligned} E_{p'}(\omega|s) &= \frac{\sum_\omega p'(\omega)p(s|\omega)\omega}{\sum_{\omega'} p'(\omega')p(s|\omega')} \\ &= 2\sum_\omega p'(\omega)\omega^2 \\ &= E_p(\omega|s) + 2\sum_\omega \Delta_p(\omega)\omega^2. \end{aligned}$$

Since $\Delta_p(\omega)$ is symmetric around 1/2,

$$\sum_\omega \Delta_p(\omega)\omega^2 = \sum_{\omega < 1-c} \Delta_p(\omega)\omega^2 + \sum_{\omega \in (1-c, c)} \Delta_p(\omega)\omega^2 + \sum_{\omega > c} \Delta_p(\omega)\omega^2$$

$$=2\sum_{\omega\in(1/2,c)}\varDelta_p(\omega)(\omega-1/2)^2+2\sum_{\omega\in[c,1)}\varDelta_p(\omega)(\omega-1/2)^2<0,$$

where the inequality holds because $|\omega-1/2|<|\omega'-1/2|$ for any $\omega\in(1/2,c)$ and $\omega'\in(c,1)$. Therefore, $E'(\omega|s)<E_p(\omega|s)$. The proof is analogous for signal b. □

Proof of Prediction 6. For convenience, we denote the binary state space as $\varOmega=\{1-x,x\}$ where $x>1/2$ and the prior as $(1-p,p)$.

Part(i). First, consider a confirmatory signal s and assume $p>1/2$. We have

$$\begin{aligned}
E(\omega)&=(1-p)(1-x)+px,\ \overline{E}(\omega)=1/2,\\
E(\omega|s)&=\frac{(1-p)(1-x)^2+px^2}{(1-p)(1-x)+px},\\
E_R(\omega|s)&=\frac{(1-p)(1-x)^{\theta+2}+px^{\theta+2}}{(1-p)(1-x)^{\theta+1}+px^{\theta+1}}.
\end{aligned}$$

The agent has wrong direction updates at s when $\hat{E}(\omega|s)-E(\omega)=\lambda E_R(\omega|s)+(1-\lambda)\overline{E}(\omega)-E(\omega)<0$, which occurs if and only if

$$\begin{aligned}
&\lambda\frac{(1-p)(1-x)^{\theta+2}+px^{\theta+2}}{(1-p)(1-x)^{\theta+1}+px^{\theta+1}}+\frac{1}{2}(1-\lambda)<(1-p)(1-x)+px\\
\Leftrightarrow&\frac{px^{\theta+1}-(1-p)(1-x)^{\theta+1}}{px^{\theta+1}+(1-p)(1-x)^{\theta+1}}<\frac{2p-1}{\lambda}.
\end{aligned}$$

The agent overreacts to s when $\hat{E}(\omega|s)-E(\omega|s)=\lambda E_R(\omega|s)+(1-\lambda)\overline{E}(\omega)-E(\omega|s)>0$, which occurs if and only if

$$\begin{aligned}
&\lambda\frac{(1-p)(1-x)^{\theta+2}+px^{\theta+2}}{(1-p)(1-x)^{\theta+1}+px^{\theta+1}}+\frac{1}{2}(1-\lambda)>\frac{(1-p)(1-x)^2+px^2}{(1-p)(1-x)+px}\\
\Leftrightarrow&\frac{px^{\theta+1}-(1-p)(1-x)^{\theta+1}}{px^{\theta+1}+(1-p)(1-x)^{\theta+1}}>\frac{1}{\lambda}\frac{px-(1-p)(1-x)}{px+(1-p)(1-x)}.
\end{aligned}$$

The agent underreacts to s if and only if

$$\frac{2p-1}{\lambda}<\frac{px^{\theta+1}-(1-p)(1-x)^{\theta+1}}{px^{\theta+1}+(1-p)(1-x)^{\theta+1}}<\frac{1}{\lambda}\frac{px-(1-p)(1-x)}{px+(1-p)(1-x)}.$$

Let $t=x/(1-x)$ and $\ell(t)\equiv\frac{pt-(1-p)}{pt+(1-p)}$, then $\ell(t)$ is increasing in t. Note that wrong direction updating occurs if $\ell(t^{\theta+1})<\frac{2p-1}{\lambda}$, underreaction occurs if $\frac{2p-1}{\lambda}<\ell(t^{\theta+1})<\frac{\ell(t)}{\lambda}$, overreaction occurs if $\ell(t^{\theta+1})>\frac{\ell(t)}{\lambda}$.

First, note that $\lim_{t\to1}\ell(t^{\theta+1})=2p-1$ and $\lim_{t\to\infty}\ell(t^{\theta+1})=1$. Therefore, if $\lambda\le 2p-1$, then the agent updates in the wrong direction for all values of x. If $\lambda>2p-1$, then there exists a cutoff $c_1\in(1/2,1)$ such that $\ell((c_1/(1-c_1))^{\theta+1})=\frac{2p-1}{\lambda}$ and the agent updates in the wrong direction for all $x\in(1/2,c_1)$.

Second, note that

$$\ell(t^{\theta+1})/\ell(t) = \frac{(pt+(1-p))(pt^{\theta+1}-(1-p))}{(pt-(1-p))(pt^{\theta+1}+(1-p))}$$
$$= 1 + \frac{2}{\frac{p^2t^{\theta+2}-(1-p)^2}{p(1-p)(t^{\theta+1}-t)} - 1}.$$

It is then easy to show that $\ell(t^{\theta+1})/\ell(t)$ is first increasing and then decreasing in t. Since $\frac{2p-1}{\lambda} = \ell((c_1/(1-c_1))^{\theta+1}) < \frac{\ell(c_1/(1-c_1))}{\lambda}$, by continuity we have $\frac{2p-1}{\lambda} < \ell(t^{\theta+1}) < \frac{\ell(t)}{\lambda}$ for t strictly larger than but sufficiently close to $c_1/(1-c_1)$. On the other hand, for t sufficiently large, both $\ell(t^{\theta+1})$ and $\ell(t)$ are close to 1, so we must have $\frac{2p-1}{\lambda} < \ell(t^{\theta+1}) < \frac{\ell(t)}{\lambda}$. Lastly, notice that for any $\lambda > 2p-1$, we have $\frac{1}{\lambda}\lim_{t\to 1} l(t) = \frac{2p-1}{\lambda} < \lim_{t\to 1}\lim_{\theta\to\infty} l(t^{\theta+1}) = 1$. Therefore, if θ sufficiently large, there exists x close to 1/2 such that the agent overreacts. Combining these observations, we know that there exist $c_1 \le c_2 \le c_3 \le 1$ such that the agent underreacts when $x \in (c_1, 1)\setminus(c_2, c_3)$ and overreacts when $x \in (c_2, c_3)$, and (c_2, c_3) is non-empty if θ is sufficiently large.

Part(ii). Next, consider a disconfirmatory signal s and assume $p < 1/2$, then $E(\omega|s) > E(\omega)$. As in Part (i), wrong direction updating occurs if $\ell(t^{\theta+1}) < \frac{2p-1}{\lambda}$, underreaction occurs if $\frac{2p-1}{\lambda} < \ell(t^{\theta+1}) < \frac{\ell(t)}{\lambda}$, overreaction occurs if $\ell(t^{\theta+1}) > \frac{\ell(t)}{\lambda}$. Since $l(t)$ is increasing, $\ell(t^{\theta+1}) > \ell(1) = 2p-1 > \frac{2p-1}{\lambda}$, so wrong direction updating is impossible. It remains to determine whether the agent overreacts or underreacts by comparing $\ell(t^{\theta+1})$ and $\frac{\ell(t)}{\lambda}$. Let $c_4 = 1-p$, then $\ell(c_4/(1-c_4)) = 0$. Note that when $t < c_4/(1-c_4)$, $\ell(t) < 0$. Thus $\frac{\ell(t)}{\lambda} < \ell(t) < \ell(t^{\theta+1})$ and the agent overreacts. On the other hand, for t sufficiently large, both $\ell(t^{\theta+1})$ and $\ell(t)$ are close to 1, which implies that $\ell(t^{\theta+1}) < \ell(t)/\lambda$ and so the agent underreacts. Therefore, there exists cutoff $c_4 \in (1/2, 1)$ such that the agent overreacts if $x \in (1/2, c_4)$ and underreacts if $x \in (c_4, 1)$. ☐

B Experiment Details and Additional Analyses

Original Analyses including wrong direction updates

Table 5. Belief-Updating by Complexity and Diagnosticity

	(1) Overreaction Ratio	(2) Overreaction Ratio
3 States	0.189**	-0.0473
	(0.0882)	(0.126)
5 States	0.185**	0.0137
	(0.0892)	(0.127)
Uninformative		0.909***
		(0.134)
3 States * Uninformative		-0.545***
		(0.130)
5 States * Uninformative		-0.582***
		(0.131)
Constant	0.0707	0.144
	(0.0846)	(0.125)
N	5683	5683
adj. R^2	0.003	0.013

Standard errors clustered at the individual level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6. Belief-Updating by Concentration and Diagnosticity

	(1) Overreaction Ratio	(2) Overreaction Ratio
Concentrated	0.209***	0.116***
	(0.0529)	(0.0400)
Diffuse	-0.219***	-0.218***
	(0.0313)	(0.0227)
Uninformative		0.513***
		(0.0628)
Concentrated * Uninformative		-0.187***
		(0.0572)
Diffuse * Uninformative		0.00151
		(0.0384)
Constant	0.260***	0.0963***
	(0.0250)	(0.0184)
N	4220	4220
adj. R^2	0.049	0.110

Standard errors clustered at the individual level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7. Belief-Updating by Signal Type and Diagnosticity

	(1) Overreaction Ratio	(2) Overreaction Ratio
Confirm	1.822***	2.510***
	(0.194)	(0.272)
Disconfirm	3.513***	5.211***
	(0.291)	(0.453)
Uninformative		-6.984***
		(0.526)
Confirm * Uninformative		2.266***
		(0.275)
Disconfirm * Uninformative		4.500***
		(0.451)
Constant	0.0707	0.144
	(0.0847)	(0.125)
N	4440	4440
adj. R^2	0.036	0.072

Standard errors clustered at the individual level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Experiment Instructions (3-state Example)

Page 1:

Introduction

Welcome to the study!

If you read the following instructions carefully, you will be able to earn additional money. The actual amount you will earn depends on your decisions. We will also test your understanding of them later.

There is a base fee of $2 for completing the study. **To complete the study, you will need to read all instructions carefully and answer the corresponding comprehension questions correctly. If you answer all the comprehension questions correctly, you will receive a bonus of $1.00.**

In addition, you can earn a potential bonus of **$10.00**. At the end of the study, one of the tasks will be randomly selected and your decision in this task will determine your bonus.

Important information

- You should think about each task **independently** of all other tasks in this study.
- You will note that we sometimes ask you work on similar-sounding tasks. These tasks might have similar answers, or very different ones. Please consider each individual task **carefully.**
- Whenever a task involves a random draw, then this random draw is **actually implemented by the computer** in exactly the way it is described to you in the task.

Page 2:

The Experiment

In this study you will be asked to complete **4 guessing tasks.**

In each guessing task, there are three bags, "Bag A," "Bag B," and "Bag C." Each bag contains 100 balls, some of which are **red** and some of which are **blue**. One of the bags will be selected at random by the computer as described below. You will not observe which bag was selected. Instead, the computer will then randomly draw a ball from the secretly selected bag, and will show this ball to you.

Your task is to **guess the probability that each bag was selected** based on the available information. The exact procedure is described below.

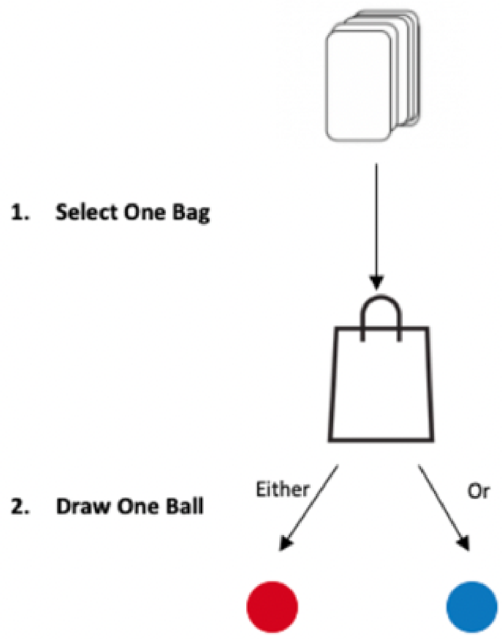
Task Setup

- There is a deck of cards that consists of 100 cards. Each card in the deck either has "Bag A," "Bag B," or "Bag C" written on it. You will be informed about **how many** of these 100 cards have "Bag A," "Bag B," and "Bag C" written on them.

- You will be informed about **how many red and blue balls** each bag contains.

These numbers are very important for making accurate guesses.

Page 2 (cont.):



Sequence of Events

1. The computer **randomly selects one** of the 100 cards, with equal probability.
   - If a "Bag A" card was drawn, Bag A is selected.
   - If a "Bag B" card was drawn, Bag B is selected.
   - If a "Bag C" card was drawn, Bag C is selected.
2. Next, the computer randomly draws **one of the 100 balls** from the secretly selected bag. Each of the 100 balls is equally likely to be selected.
3. The computer will then **inform you about the color** of the randomly drawn ball.

After seeing the color of the ball, you will make your guess by **stating a probability between 0% and 100%** that each of Bag A, Bag B, and Bag C was drawn. Note that the probabilities have to sum to 100.

One ball will be drawn from a bag and you will make one guess after the ball is drawn.

Please Note

- The number of "Bag A," "Bag B," and "Bag C" cards **can vary across tasks**.
- The number of red and blue balls in each bag **varies across tasks**.
- The computer **draws a new card for each task,** so you should **think about which bag was selected in a task independently of all other tasks.**

Page 3:

Your Payment

You can potentially earn an additional bonus of **$10.00.** At the end of the study, we will randomly select one of your guesses. Whether or not you receive the $10.00 depends on how much probability you assigned to the bag that was *actually drawn* in those tasks.

This means: if Bag A was selected, your chances of receiving $10.00 are greater the higher the probability you assigned to Bag A. If Bag A was not selected, your chances of receiving $10.00 are greater the lower the probability you assigned to Bag A. In case you're interested, the specific method that determines whether you get the prize is explained in the link here.

Page 3 (link):

Using the laws of probability, the computer determines a **statistically correct statement of the probability that a good asset was drawn**, based on all of the information available to you. This **optimal guess** does not rely on information that you do not have. It is just the best possible (this means: payoff maximizing) estimate given the available information. In technical terms, this guess is based on a statistical rule called Bayes' Law. If your guess is within 3% of the optimal guess, you will earn the additional $10.00.

All this means is that, in order to earn as much money as possible, you should try to give your best estimate of the probability that each asset was drawn. For example, if you are 50% sure that a good asset was selected, 30% sure that a neutral asset was selected, and 20% sure that a bad asset was selected, you should allocate probability 50% to a good asset, 30% to a neutral asset, and 20% to a bad asset.

Page 4:

Comprehension Questions

The following questions test your understanding of the instructions.

Important: If you fail to answer any one of these questions correctly, you will not earn the additional $1 bonus. You will have one chance to answer the questions correctly. Click here to review the instructions.

Which statement about the number of cards corresponding to each bag is correct?

- ○ The number of "Bag A" cards is always the same in all tasks.
- ○ The exact number of cards corresponding to each bag may vary across tasks.

Which statement about the allocation of red and blue balls in the bags is correct?

- ○ The exact fraction of red and blue balls in each bag may vary across tasks.
- ○ The fraction of red balls in each bag is the same in all tasks.

Which statement about the probabilities of each bag is correct?

- ○ In a given task, the probabilities that each bag was drawn must add up to 100.
- ○ In a given task, the probability that each bag was drawn is 100, summing up to 300 in total.

If Bag A has more red balls than blue balls and Bag B has more blue balls than red balls, and a red ball is drawn in the first round, which bag is more likely to have been chosen for this task? Write **Bag A** or **Bag B**.

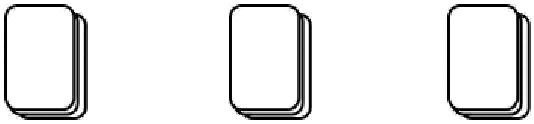
If Bag C has more blue balls than red balls and Bag A has more red balls than blue balls, and a red ball is drawn in the first round, which bag is more likely to have been chosen for this task? Write **Bag A** or **Bag C.**

Page 5:

Thank you for your responses. Please click continue to proceed on to the tasks.

Trial Example

In this task:



"Bag A" cards: **33** **"Bag B"** cards: **34** **"Bag C"** cards: **33**

Bag A contains 60 red balls and 40 blue balls.
Bag B contains 50 red balls and 50 blue balls.
Bag C contains 40 red balls and 60 blue balls.

Next, the computer **randomly selected one bag** by drawing a card from the deck.

The computer randomly drew the first ball from the selected bag:

Your task is to guess the probability that each bag was selected.

Select a probability (between 0 and 100) that expresses **how likely** you think that **Bag A**, **Bag B**, and **Bag C** have been selected. Note that the probabilities have to sum to 100.

Probability of **Bag A**:	0
Probability of **Bag B**:	0
Probability of **Bag C**:	0
Total	0