AEA Papers and Proceedings 2025, 115: 631–636
https://doi.org/10.1257/pandp.20251107

FORMING MENTAL MODELS

The Role of Representational and Computational Complexity in Belief Formation†

By Cuimin Ba, J. Aislinn Bohren, and Alex Imas*

A growing body of theoretical and experimental research explores how individuals respond to information in complex environments. Previous studies have explored how people simplify and distort information in specific cognitively demanding situations, but less focus has been placed on distinguishing between different sources of complexity.

This paper builds on the two-stage belief updating model of Ba, Bohren, and Imas (2024) to study the interaction between representational and computational complexity. We consider a standard belief updating problem where an agent learns about an unknown state from a signal. The agent uses a two-stage process to form her subjective belief. In the *representational* stage, the agent forms a mental model of the information structure, which we refer to as a mental representation. In the *processing* stage, the agent processes the signal by applying a noisy version of Bayes' rule to her mental representation.

As the number of states increases, the cognitive demands of representing and processing information also increase. Specifically, a larger state space has greater representational complexity, as constrained attention makes it more difficult to attend to and mentally represent the information structure. It also has greater computational complexity, as processing constraints make it more difficult to carry out the computation necessary for Bayes' rule. Thus, both stages may be subject to more distortions.

Ba, Bohren, and Imas (2024) focus on environments where representativeness is present as a salience cue. They find that people mitigate these cognitive demands by channeling attention to a limited number of salient states. This dominates the impact of processing constraints, leading to excess movement of beliefs—that is, overreaction.

This paper focuses on environments where there are no obvious salience cues, and hence, the interaction between computational complexity and processing constraints may play a more prominent role in belief updating. We hypothesize that in such environments, participants will underreact. Moreover, cognitive imprecision (i.e., processing noise) will increase in the complexity of the environment, leading to more underreaction. We test this hypothesis in an online experiment. Participants were assigned to either a *simple* information environment with two states or a *complex* environment with five states. Consistent with our hypothesis, participants underreacted in both conditions, and more so in the complex one. Structural estimates suggest that this pattern is driven by higher cognitive imprecision in the complex condition.

I. Model of Belief Updating

An agent learns about an unknown state ω drawn from state space $\Omega \equiv \{\omega_1, \dots, \omega_N\} \subset (0,1)$ with $\omega_1 < \dots < \omega_N$ according to uniform prior $p_0 = (1/N, \dots, 1/N)$.[1] The agent observes a draw

* Ba: Department of Economics, University of Pittsburgh (email: bacuimin@gmail.com); Bohren: Department of Economics, University of Pennslyvania (email: abohren@sas.upenn.edu); Imas: Booth School of Business, University of Chicago (email: alex.imas@chicagobooth.edu).

† Go to https://doi.org/10.1257/pandp.20251107 to visit the article page for additional materials and author disclosure statement(s).

[1] We assume a uniform prior for simplicity; the results naturally extend to more general priors.

of a binary signal $s \in \{s_1, s_2\}$ distributed according to $\pi(s_2|\omega_i) = \omega_i$ and $\pi(s_1|\omega_i) = 1 - \omega_i$ in state ω_i. We refer to Ω as the *information structure*, since the values of the state pin down the signal distribution. We focus on symmetric information structures—that is, $\omega_i \in \Omega$ implies that $1 - \omega_i \in \Omega$. By Bayes' rule, the objective posterior probability of state ω_i following signal realization s_2 is $p_B(\omega_i|s_2) = \omega_i / \sum_{\omega_k \in \Omega} \omega_k$ and analogously $p_B(\omega_i|s_1) = (1 - \omega_i)/\sum_{\omega_k \in \Omega}(1 - \omega_k)$ following s_1. Let $p_B(s_j) = (p_B(\omega_1|s_j), \dots, p_B(\omega_N|s_j))$ denote the objective posterior distribution and $E_B[\omega|s_j] \equiv \sum_{\omega_k \in \Omega} \omega_k p_B(\omega_k|s_j)$ denote the objective posterior expected state, with an analogous definition for the prior expected state $E_0[\omega|s_j]$.

Following Ba, Bohren, and Imas (2024), the agent uses a two-stage process to form beliefs. First, she forms a mental model $\hat{\pi}$ of the information structure, which we refer to as a *mental representation*. This representation simplifies the information environment by channeling attention to certain states and neglecting others. Second, the agent processes the signal by applying a noisy version of Bayes' rule to her mental representation.

In the first stage, attention constraints interact with the *representational complexity* of the information environment. Representational complexity is increasing in the number of states N, as it captures the number of objects the agent needs to consider to form an accurate mental representation. While the belief updating model in Ba, Bohren, and Imas (2024) primarily focuses on salience-channeled attention, it can also be applied to environments with no salience cues where the agent channels attention as-if randomly. In this variation of the model, the agent channels attention to the first state she considers. Each state $\omega_i \in \Omega$ is equally likely to be attended to first, and all states that are not attended to first receive equal attention. This is modeled as a mental representation where the likelihood of signal realization s_j in the attended-to-first state ω_i is scaled up by attention weight $\alpha_1 > 1$, $\hat{\pi}_i(s_j|\omega_i) = \alpha_1 \pi(s_j|\omega_i)$, and the likelihood of s_j in not-attended-to-first states $\omega_k \neq \omega_i$ is scaled down by attention weight $\alpha_2 \in (0,1)$, $\hat{\pi}_i(s_j|\omega_k) = \alpha_2 \pi(s_j|\omega_k)$. The extent to which attention is constrained, and therefore channeled differentially across states, is captured by the extent to which parameters α_1 and α_2 differ from one.

In the second stage, processing constraints interact with the *computational complexity* of belief formation. Computational complexity is also increasing in the number of states N, as it captures the number of operations the agent needs to compute when implementing the relevant algorithm (Oprea 2020)—in this case, Bayes' rule. This interaction leads the agent to exhibit cognitive imprecision when using her mental representation to update beliefs. Following Woodford (2020) and Khaw, Li, and Woodford (2021), we model cognitive imprecision as the agent using a noisy cognitive signal of the underlying posterior to form her subjective posterior. Here, the relevant underlying posterior $p_{R,i}(s_j)$ is the Bayesian update with respect to the agent's mental representation given observed signal realization s_j and first-attended-to state ω_i. Specifically, letting $\alpha \equiv \alpha_1/\alpha_2$,

$$(1) \qquad p_{R,i}(\omega_i|s_2) = \frac{\omega_i}{\omega_i + \frac{1}{\alpha}\sum_{k \neq i} \omega_k} \text{ and}$$

$$(2) \qquad p_{R,i}(\omega_k|s_2) = \frac{\omega_k}{\alpha\omega_i + \sum_{k \neq i} \omega_k}, \; k \neq i.$$

Following the model of cognitive imprecision from Ba, Bohren, and Imas (2024), this results in a mean observed posterior

$$(3) \qquad \hat{p}_i(s_j) \equiv \lambda p_{R,i}(s_j) + (1 - \lambda)\bar{p}_0 \in \Delta(\Omega),$$

when the agent observes realization s_j and attends to state ω_i first, where $\lambda \in [0,1]$ measures the agent's cognitive precision and $\bar{p}_0 \in \Delta(\Omega)$ is the *cognitive default*, which is taken to be uniform as in Enke and Graeber (2023) and Ba, Bohren, and Imas (2024).[2] We refer to $\hat{p}_i(s_j)$ as the *subjective posterior*. Cognitive imprecision is decreasing in λ: $\lambda = 1$ corresponds to no cognitive imprecision, where the agent's subjective posterior is the Bayesian update with respect to her mental representation, while $\lambda = 0$ corresponds to full cognitive imprecision, where the agent's belief remains centered at her cognitive default following both signal realizations.

As in Ba, Bohren, and Imas (2024), we measure the magnitude of over- or underreaction using the *overreaction ratio*,

$$r_i(s_j) \equiv \frac{\hat{E}_i[\omega|s_j] - E_B[\omega|s_j]}{E_B[\omega|s_j] - E_0[\omega|s_j]}, \tag{4}$$

where $\hat{E}_i[\omega|s_j] \equiv \sum_{\omega_k \in \Omega} \omega_k \hat{p}_i(\omega_k|s_j)$ denotes the subjective posterior expected state. The agent's reaction depends on the state attended to first. We focus on the average reaction across all random attention allocations,

$$r(s_j) \equiv \frac{1}{N} \sum_{\omega_i \in \Omega} r_i(s_j). \tag{5}$$

We say that the agent *overreacts* if $r(s_j) > 0$, *underreacts* if $r(s_j) \in [-1,0)$, and reacts in the *wrong direction* if $r(s_j) < -1$.[3]

Prediction 1 compares how the agent's average reaction varies with the complexity of the state space. It shows that if cognitive precision declines sufficiently as complexity increases, then underreaction increases. The proof is in the Supplemental Appendix.

PREDICTION 1: *Consider two information structures Ω and Ω' with different complexities, $N < N'$, and the same extreme states, $\omega_1 = \omega'_1$ and $\omega_N = \omega'_{N'}$. Given attention parameters $\alpha = \alpha' > 1$ and cognitive precision parameters $\lambda, \lambda' \in [0,1]$, the agent underreacts to both signal realizations in both environments, $r(s_j) \in [-1,0)$ and $r'(s_j) \in [-1,0)$ for $j = 1,2$. Moreover, there exists an $\varepsilon \in (0,\lambda)$ such that if the agent exhibits sufficiently more cognitive imprecision in the more complex environment Ω', $\lambda' < \lambda - \varepsilon$, then the agent underreacts more in Ω' than Ω, $r'(s_j) < r(s_j)$ for $j = 1,2$.*

The agent's cognitive constraints interact with the representational and computational complexity of the information environment to determine her reaction to information. The state the agent attends to first can lead to overreaction or underreaction. For example, following signal realization s_2 that is most likely in state ω_N, attending to ω_N first leads the agent to overweigh it—and hence, to overreact. In contrast, attending to ω_1 first leads the agent to overweigh it—even though it is least likely to generate s_2—and hence, to underreact. However, since the mass on ω_1 should decrease following s_2 while the mass on ω_N should increase, overweighing ω_1 results in a larger distortion than overweighing ω_N. This implies that, on average, random attention leads to underreaction. Cognitive imprecision reinforces this underreaction. As a result, if computational complexity increases cognitive imprecision, this leads to greater underreaction.[4]

[2] Formally, the model of cognitive imprecisions is as follows. The cognitive signal of $p_{R,i}(s_j)$ is drawn from a multinomial distribution with event probabilities $p_{R,i}(\cdot|s_j)$, N categories, and $\eta \geq 0$ trials. The cognitive prior over $p_{R,i}(\cdot|s_j)$ follows a Dirichlet distribution with N categories and concentration parameters $\nu \cdot \bar{p}_0$. The agent uses Bayes' rule to form a cognitive posterior about the distribution of $p_{R,i}(s_j)$ given her cognitive prior and the cognitive signal. The mean observed posterior is the expectation of the average cognitive posterior conditional on $p_R(s_j)$, where $\lambda = \eta/(\eta + \nu)$.

[3] If $r(s_j) = 0$, the subjective and objective posterior expected states coincide.

[4] Note that random attention on its own does not necessarily generate this comparative static: When there is no cognitive imprecision, whether increasing the number of states leads to more or less underreaction depends on the information environment.

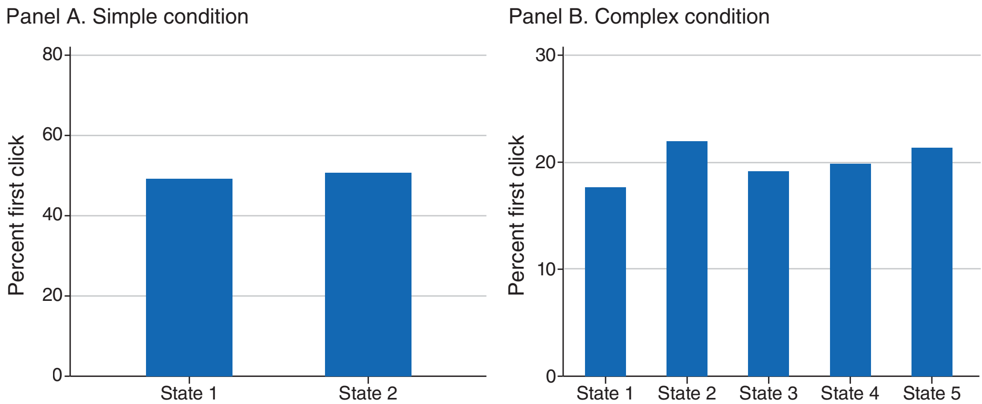

Figure 1. Participants' First Clicks Are as-if Random

II. Experimental Design

To test Prediction 1, we recruited 247 subjects from Prolific to participate in a controlled experiment. As in Ba, Bohren, and Imas (2024), we adopted the classic "bookbag-and-poker-chip" design and elicited posterior beliefs in various information environments. Participants first learned about the general design of the information environment but not the exact information structure. To learn the information structure, participants had to click on each state to display the probability of signal realization s_1 versus s_2 in that state. This procedure eliminated readily available salience cues such as representativeness.

We used the Mouselab design of Payne, Bettman, and Johnson (1988) to measure attention; they showed that the object that is clicked first is a valid measure of attention. Using this design, Ba, Bohren, and Imas (2024) showed that the state participants clicked first received increased weight in the posterior belief.

We compare participants' posterior beliefs in *simple* and *complex* information environments:

- *Simple*: $\Omega_2 = \{\omega_1, \omega_1\}$, where $\omega_1 \in \{0.1, 0.2, 0.3, 0.4\}$.
- *Complex*: $\Omega_5 = \{\omega_1, \omega_2, 0.5, 1 - \omega_2, 1 - \omega_1\}$, where $\omega_1 \in \{0.1, 0.2, 0.3, 0.4\}$, $\omega_2 \in \{0.2, 0.3, 0.4, 0.45\}$, and $\omega_1 < \omega_2$.

Per our preregistration, we exclude observations in which participants react in the wrong direction.

III. Results

First Clicks.—We first test the assumption that our paradigm removed salience cues and generated random attention. As shown in Figure 1, the fraction of participants who clicked each state first was evenly distributed across states in both the simple and complex conditions. Thus, attention was allocated randomly across states, as in the setup presented in Section I.

Over- and Underreaction.—Consistent with Ba, Bohren, and Imas (2024) and Prediction 1, when salience cues were absent, participants overwhelmingly underreacted on average in both the simple and complex conditions: The overreaction ratio was negative across all information structures in both conditions (see Figure 2). This contrasts with Ba, Bohren, and Imas (2024), where participants overwhelmingly overreacted in complex five-state environments when the representativeness salience cue was present.

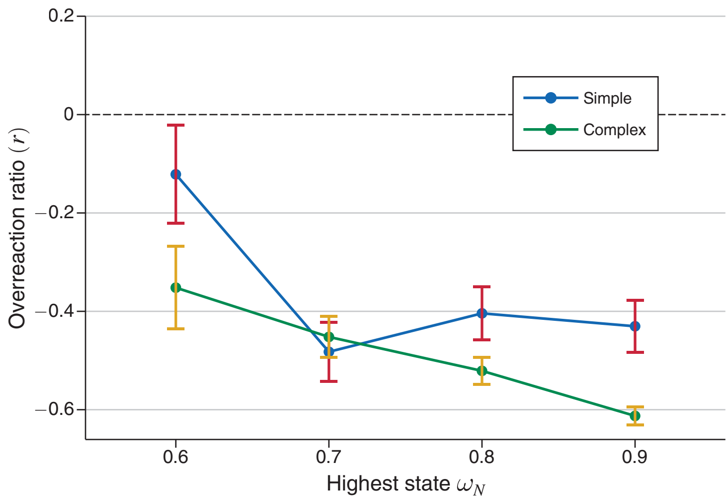

FIGURE 2. OVERREACTION RATIO VARIES WITH COMPLEXITY

Notes: Each data point aggregates all information environments with a given highest state ω_N in the relevant complexity condition.

To study the role of computational complexity, we compare underreaction across both complexity conditions. In line with Prediction 1, agents *underreacted more* in the complex condition compared to the simple one: Averaging across all information environments of a given complexity, the overreaction ratio was -0.53 in the complex condition and -0.36 in the simple one ($p < 0.01$). Similarly, a higher share of participants (83.6 percent) underreacted in the complex condition compared to the simple condition (73.4 percent). Figure 2 shows that this pattern also generally holds when information environments are disaggregated by the value of the highest state (which corresponds to the highest probability of generating signal realization s_2): The overreaction ratio was significantly lower in the complex five-state environments compared to the analogous two-state environment for three of the four values considered (the exception is information environments with $\omega_N = 0.7$). This contrasts starkly with the finding in Ba, Bohren, and Imas (2024) when the representativeness salience cue was present: In this case, channeling attention to the most representative state resulted in *more overreaction* in the complex condition.

Taken together, these results provide strong evidence for the importance of both channeled attention and processing constraints in belief formation. As shown in Ba, Bohren, and Imas (2024), when representativeness is present as a salience cue, salience-channeled attention plays a dominant role and leads to overreaction. In contrast, when salience cues are absent, randomly channeled attention interacts with processing constraints to generate underreaction. In such environments, the impact of higher computational complexity on cognitive precision has a dominant effect, leading to greater underreaction.

Structural Estimation.—Finally, we structurally estimate the attentional distortion and cognitive precision parameters for each condition. The results are presented in Table 1. The attentional distortion parameter α is estimated at 1.06 for the simple condition and 1.20 for the complex condition. This difference is not statistically significant, suggesting similar attentional distortions across both conditions. In contrast, cognitive imprecision increased with complexity: The cognitive precision parameter λ was significantly higher in the simple condition than the complex condition (0.58 versus 0.42, $p < 0.01$). Taken together, these estimates suggest that the interaction between computational complexity and cognitive imprecision played a key role in generating more underreaction in the complex condition (as predicted by Prediction 1).

Table 1—Parameter Estimates

	α	95% CI	λ	95% CI
Simple	1.06	[1.00, 1.20]	0.58	[0.52, 0.65]
Complex	1.20	[1.01, 1.55]	0.42	[0.35, 0.46]

Notes: Parameter estimates minimize the average Kullback–Leibler divergence of the model-predicted subjective posteriors from the reported posteriors, aggregating across all information environments in the relevant complexity condition. The 95 percent confidence intervals are obtained from 300 bootstrap samples.

IV. Conclusion

We show that increased state-space complexity amplifies distortions in both the mental representation and processing stage of belief formation. Theoretically and empirically, we show that if attention is channeled as-if randomly, then cognitive imprecision reinforces attention distortions in shaping beliefs, leading to *more underreaction* as state-space complexity increases. This contrasts with the findings in Ba, Bohren, and Imas (2024), where representativeness salience cues are present. In their paper, representativeness-channeled attention distortions dominate cognitive imprecision in shaping beliefs, resulting in *more overreaction* as state-space complexity increases. These findings underscore the importance of considering the interaction between psychological frictions in belief updating. They also open avenues for further exploration of how other salience cues or learning may mitigate these effects.

REFERENCES

Ba, Cuimin, J. Aislinn Bohren, and Alex Imas. 2024. "Over- and Underreaction to Information." Penn Institute for Economic Research Working Paper 24-030.

Enke, Benjamin, and Thomas Graeber. 2023. "Cognitive Uncertainty." *Quarterly Journal of Economics* 138 (4): 2021–67.

Khaw, Mel Win, Ziang Li, and Michael Woodford. 2021. "Cognitive Imprecision and Small-Stakes Risk Aversion." *Review of Economic Studies* 88 (4): 1979–2013.

Oprea, Ryan. 2020. "What Makes a Rule Complex?" *American Economic Review* 110 (12): 3913–51.

Payne, John W., James R. Bettman, and Eric J. Johnson. 1988. "Adaptive Strategy Selection in Decision Making." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (3): 534–52.

Woodford, Michael. 2020. "Modeling Imprecision in Perception, Valuation, and Choice." *Annual Review of Economics* 12: 579–601.