Timing of Informativeness on Motivated Reasoning*

Hyundam Je[†] Sora Youn[‡]

November 2024

Abstract

We present the first experimental evidence that the timing of informativeness influences motivated reasoning: Participants place more weight on information that aligns with their desired outcomes when they learn about a signal's informativeness after receiving the signal, compared to when the order of information is reversed. This order effect occurs when prior beliefs are not extreme, regardless of the signal's informativeness. Our findings suggest that revealing informativeness earlier could be a simple yet effective way to mitigate motivated reasoning.

*We thank Syngjoo Choi, Alexander Coutts, Catherine Eckel, Christine Exley, Charles Holt, Daeyoung Jeong, Paolina Medina, Kirby Nielsen, Hyoeun Park, and Emanuel Vespa for their helpful comments. We also thank participants at the 2021 North-American ESA Meeting, the 2023 North-American ESA Meeting, the 2023 SEA Annual Meeting, the 2024 KEBEA Winter Workshop, and the NUS Applied Economics Student Workshop. This research was funded by the Department of Economics, Texas A&M University, and the College of Economics and Finance, Hanyang University. The study was pre-registered in AsPredicted (126181) and had ethics approval from Texas A&M University's Institutional Review Board (IRB2021-1435M). All errors are our own.

[†]Je: School of Economics, University of Sydney, hyundam.je@sydney.edu.au

[‡]Youn: Korea Information Society Development Institute, youn.soraaaaa@gmail.com

1 Introduction

Consider two scenarios: In the first, you come across an online news article claiming that a politician has committed a crime. Later, you find out whether the article's source is reliable. In the second scenario, you read the same article, but this time, you know the source's reliability before you begin reading. In both cases, you receive the same information: the news article (signal) and the reliability of the source (the signal's informativeness). The only difference is the order of information presentation: In the first scenario, the signal precedes the informativeness, while in the second scenario, the informativeness comes first.

Motivated reasoning is the tendency for individuals to process information in a way that aligns with their desired outcomes (Kunda, 1990; Bénabou and Tirole, 2002; Epley and Gilovich, 2016). Previous studies have provided abundant evidence of this tendency, whether the signal is informative (Eil and Rao, 2011; Möbius et al., 2022; Coutts, 2019a) or uninformative (Gonçalves et al., 2021; Ecker et al., 2022; Wittrock et al., 2023; Thaler, 2023). In the scenarios above, motivated reasoning suggests that an individual who dislikes the politician is more inclined to believe the news, regardless of its informativeness. This paper presents novel evidence that the order in which a signal and its informativeness are revealed could affect motivated reasoning.

There is extensive literature on the effects of the order in which information is presented. One well-documented phenomenon is the primacy effect, where individuals tend to pay more attention to and be more influenced by the first piece of information they receive. This effect has been observed across various fields, including social psychology (Jones et al., 1968), law (Ahlering and Parker, 1989), political science (van Erkel and Thijssen, 2016), and behavioral economics (Hogarth and Einhorn, 1992; Saccardo and Serra-Garcia, 2023). In most of the literature, the primacy effect has been studied in the context of the influence of the order of information when receiving two pieces of information (e.g., signals). In contrast, our framework examines situations where individuals receive two heterogeneous types of information: a signal and its informativeness. Therefore, the order of information in our framework is equivalent to the timing of informativeness. By separating the signal from its informativeness, we can understand how the timing of informativeness affects belief updating and gain deeper insights into how individuals process information. To the best of our knowledge, the timing of informativeness on belief updating has not yet been studied.

To assess the effect of the timing of informativeness, we conduct an online experiment using a variant of the classic box-and-ball setup. The box-and-ball setup is useful for this study for several reasons. First, it provides a neutral and controlled environment, reducing the potential noise from individual heterogeneity in real-world contexts. In real life, people's varied experiences and backgrounds could introduce variability, obscuring clear patterns in belief updating or motivated reasoning. By using a simplified experimental design, we can focus on cognitive processes without the confounding effects of preexisting personal biases or complexity. Second, the box-and-ball framework allows for the clear revelation of the timing of informativeness. In real-life scenarios, it is often difficult to manipulate or control the informativeness of signals because individuals may interpret a signal's informativeness differently based on their prior knowledge or preferences. The experimental setup in this study ensures that the signals' informativeness is consistent and easily understood, enabling clear analysis of how the timing of informativeness affects belief updating.

There are two boxes: Box A, containing one yellow and one blue ball, and Box B, containing two balls that can be either yellow or blue, though its composition is unknown to the participants. Participants are asked to predict the color of a Target Ball drawn from Box B. To assist in their decision, a Clue Ball is drawn with replacement, independently of the Target Ball. The Clue Ball is randomly selected from either Box A or Box B, with equal probability.

If the Clue Ball is drawn from Box A, its color is uninformative for predicting the color of the Target Ball. However, if the Clue Ball is drawn from Box B, its color provides information about the composition of Box B and, thus, the color of the Target Ball.

Our study employs a 2x2 between-subject structure. The first factor we vary is the order in which information is presented: Participants learn the color of the Clue Ball (the signal) and which box the Clue Ball is drawn from (the signal's informativeness), with the order depending on the treatment. In the Box First treatment, participants first learn which box the Clue Ball comes from, followed by the color of the Clue Ball, meaning the informativeness precedes the signal. In contrast, in the Color First treatment, participants first learn the color of the Clue Ball, followed by which box it comes from, meaning the signal precedes its informativeness. Importantly, in both treatments, participants receive the same information; the only difference is the order in which the information is presented.

The second factor we manipulate is the reward scheme, which allows us to verify whether motivated reasoning is driven by monetary incentives as intended. We implement two reward schemes for each treatment. At the beginning of the experiment, before participants receive instructions about the experiment, we ask them to choose between yellow and blue as their preferred color. In the Variable Reward scheme, if participants win (i.e., correctly predict the color of the Target Ball), their reward depends on whether their preferred color matches the color of the Target Ball: they receive \$1.5 if the colors match, and \$0.5 otherwise. This scheme is designed to induce motivated reasoning: participants are likely to believe that the Target Ball's color will match their preferred color. Conversely, the Fixed Reward scheme offers participants a fixed winning reward of \$1, independent of whether the Target Ball matches their preferred color. The Fixed Reward scheme allows us to confirm that the monetary-driven incentive functions in the Variable Reward scheme operate as intended.

To the best of our knowledge, our study is the first to provide evidence that the timing of informativeness affects motivated reasoning. In the Box First treatment, where informativeness is revealed earlier, whether the favored signal or not, i.e., whether the color of the Clue Ball matches the preferred color, does not affect participants' belief updating. In contrast, in the Color First treatment, where informativeness is revealed later, participants update their beliefs more substantially in line with the favorable signal compared to unfavorable signal. It implies that later revelation of informativeness stimulates motivated reasoning. This result is intriguing because, even when the content of the signal is the same, the timing of the signal's informativeness revelation could influence belief updating.

Furthermore, the experimental results identify two key features for the effect of the timing of informativeness on motivated reasoning. First, the order effect is independent of the signal's informativeness: we observe this effect not only when the signal is uninformative (Thaler, 2023), but also when the signal is informative. Second, the order effect occurs only when prior beliefs are not extreme. Participants with extreme priors tend to rely more on their prior beliefs and pay less attention to the signal's informativeness, aligning with the findings of (Bisière et al., 2015). As a result, the effect of the timing of informativeness is marginal when the priors are extreme.

Our paper makes two key contributions. First, we are the first to uncover that the timing of informativeness affects motivated reasoning. While many studies have documented evidence of motivated reasoning, the analysis of when and how it occurs has been limited (See Kunda, 1990). By employing our novel design, we find that the delayed revelation of informativeness could be factor that can stimulate motivated reasoning. Note that in our environment, we elicit motivated reasoning with only a $1 difference in rewards. However, in reality, motivated reasoning often stems from substantially greater disparities, such as strong political biases. Our findings suggest the possibility that if the context of motivated reasoning were more grounded in these real-world cases, the timing of informativeness could play a more significant role.

The second, more practical contribution is in reducing motivated reasoning. Previous studies have examined the negative effects of motivated reasoning and various strategies to mitigate it (Kunda, 1990; Beattie and Snider, 2019; Carpenter, 2019; Cotter et al., 2020). Motivated reasoning can lead to several issues across various domains, fostering bias and polarization by reinforcing preexisting beliefs and deepening ideological divides (Klaczynski, 2000; Redlawsk, 2002; Taber et al., 2009). It also skews decision-making, resulting in partisan outcomes (Bolsen and Palm, 2019), and contributes to overconfidence, as individuals dismiss contradictory evidence, reinforcing their existing beliefs (Thaler, 2021). Our results indicate that the timing of informativeness could play a crucial role: We propose that earlier disclosure of a signal's informativeness could be an effective strategy to curb motivated reasoning.

The paper is organized as follows: Section 2 introduces the experimental design. Section 3 details the framework and predictions regarding the experimental results. Section 4 presents the experiment's results, and Section 5 provides the conclusion.

2 Design

There are two boxes: Box A and Box B, each containing two balls. Box A always contains one yellow ball and one blue ball. The composition of Box B is unknown; each ball could be yellow or blue with equal probability. That is, Box B could contain two yellow balls with a 1/4 chance, two blue balls with a 1/4 chance, or one yellow ball and one blue ball with a 1/2 chance.[1] The experimental setting is summarized in Figure 1.

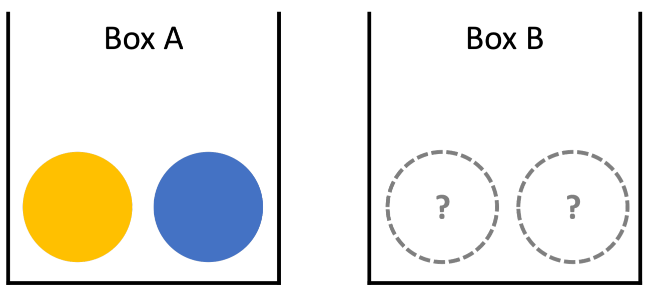


Figure 1: Experimental Setting

The computer independently draws two balls: a Target Ball and a Clue Ball. A *Target Ball*

[1]To confirm that ambiguity does not significantly affect the results, we conduct two versions of the instructions. In one version, participants are not informed about the distribution of Box B: "Box B contains two balls; however, the composition is unknown." In the other version, we provide additional information about the distribution: "Each ball in Box B could be yellow or blue with equal probability." In the main analyses, we combine both versions because we found no behavioral differences between them. See Appendix D for further information.

is drawn from Box B. The main experimental task for participants is to report the percentage chance (ranging between 0 and 100) that they believe the Target Ball is yellow or blue. To assist their belief updating, the computer draws a *Clue Ball* from either Box A or Box B with equal probability. If the Clue Ball is drawn from Box A, its color conveys no information about the composition of Box B or the color of the Target Ball. On the contrary, if the Clue Ball is drawn from Box B, the color of the Clue Ball hints about the color of the Target Ball: a yellow Clue Ball suggests a higher probability that the Target Ball is yellow, while a blue Clue Ball indicates a greater chance that the Target Ball is blue. Hence, the *color* of the Clue Ball serves as a signal about the color of the Target Ball, and the *box* from which the Clue Ball is drawn represents the informativeness of the signal.

To overview, our experiment follows a 2x2 between-subjects design. One dimension relates to the timing of informativeness, varying the order in which the order in which the *color* of the Clue Ball and the *box* from which it is drawn are presented. The other dimension involves the reward schemes, verifying that motivated reasoning is primarily driven by the monetary domain.

2.1 Timing of the Informativeness

Participants report their beliefs about the color of the Target Ball at three stages: the prior, interim, and posterior. Between each stage, the computer reveals information about the Clue Ball. In the *Box First* treatment, participants first learn about the box from which the Clue Ball is drawn, followed by the color of the Clue Ball. In the *Color First* treatment, participants receive the same information but in the reverse order. See Table 1 for the timeline of each treatment.

Note that the *Box First* treatment and the *Color First* treatment ultimately convey the same information; they differ only in the order in which the information is presented. Thus, Bayes' rule predicts that posterior beliefs will be identical across both treatments when controlling for prior beliefs. Any differences observed between the treatments would therefore suggest an effect of the order of information—specifically, the timing of informativeness—on belief updating.

2.1.1 Box First Treatment

In the *prior* stage, participants report their prior beliefs about the color of the Target Ball. They use a slider bar to indicate the percentage chance, ranging from 0 to 100, that they believe the Target Ball is yellow or blue.[2]

In the *interim* stage, participants are informed about the box from which the Clue Ball is drawn, but the color is not yet revealed. They receive the following message: "The Clue Ball is drawn from [Box A/Box B]. You don't see the color of the Clue Ball at this point." After receiving the message, they report their interim beliefs about the color of the Target Ball, using the same slider bar as in the prior stage.

In the *posterior* stage, participants are provided with the complete information: "The color of the Clue Ball is [Yellow/Blue]. The Clue Ball is drawn from [Box A/Box B]." They then report their posterior beliefs about the color of the Target Ball using the same slider bar.

2.1.2 Color First Treatment

The *Color First* treatment is identical to the *Box First* treatment, except for *interim* stage. In the *interim* stage of the *Color First* treatment, participants receive information about the color of the Clue Ball, but not about the box: "The color of the Clue Ball is [Yellow/Blue]. You don't know which box it was drawn from at this point."

[2]Appendix E includes screenshots of the interface presented to participants during the slider task.

Box First	Color First
The Target Ball and the Clue Ball are drawn	
↓	↓
Prior belief elicitation	Prior belief elicitation
↓	↓
"The Clue Ball is from **Box A/B**."	"The Clue Ball is **Yellow/Blue**."
↓	↓
Interim belief elicitation	Interim belief elicitation
↓	↓
"The Clue Ball is **Yellow/Blue**."	"The Clue Ball is from **Box A/B**."
↓	↓
Posterior belief elicitation	Posterior belief elicitation
↓	
The color of the Target Ball is revealed	

Table 1: Timeline of two treatments

2.2 Reward Schemes

Participants could win a prize based on the accuracy of their guess about the Target Ball's color. The more accurate their guess, the higher the chance of winning the prize. (The detailed process will be discussed in the next section.) The prize varies depending on the reward scheme in which they participate, either the Variable Reward scheme or the Fixed Reward scheme.

At the beginning of the experiment, before receiving any instructions, participants are asked to choose either yellow or blue as their preferred color. In the Variable Reward scheme, the prize is higher if the color of the Target Ball matches the participant's preferred color. Specifically, if the Target Ball is yellow (blue), a participant who chose yellow (blue) has a chance to win \$1.5, while a participant who chose blue (yellow) can win a lower prize of \$0.5. In the Fixed Reward scheme, the winning prize is \$1, regardless of the color of the Target Ball or their preferred color.

Note that the prize amount is affected by whether the Target Ball's color matches the participant's preferred color only in the Variable Reward scheme. Therefore, we expect participants to have a stronger incentive to believe that the Target Ball matches their preferred color in the Variable Reward scheme, but not in the Fixed Reward scheme. In other words, we predict that motivated reasoning occurs solely in the Variable Reward scheme.

We implement the two reward schemes for the following reasons. First, in the Variable Reward scheme, we link the preferred color to the prize amount to induce motivated reasoning. However, even if motivated reasoning is observed, it may not be entirely driven by monetary incentives; participants' inherent color preferences could play a role. For example, participants who choose yellow as their preferred color might be inclined toward the Target Ball being yellow, not because it offers three times the prize of the blue Target Ball, but simply because they personally favor yellow. To ensure that the monetary-driven incentive functions as intended, we implement the Fixed Reward scheme.

Second, the Fixed Reward scheme creates an environment free from motivated reasoning. Our framework suggests that the order of information influences belief updating when motivated reasoning is present, but has no effect when it is absent (see Section 3 for the details). To verify this, it is essential to create an environment identical to the Variable Reward scheme, but where motivated reasoning is not expected to occur.

2.3 Belief Elicitation

There are a total of 5 independent rounds, each including three belief elicitation questions (the prior, interim, and posterior, as described in the previous section).[3] This results in a total of 15 reported beliefs. To incentivize belief elicitation, we employ a binarized scoring rule (Hossain and Okui, 2013). At the end of the experiment, the computer randomly selects

[3]To prevent potential spillover effects, we inform participants at the beginning of the experiment that it consists of two parts. During the first part, which covers the first round, participants are unaware of the task in the second part. Upon entering the second part, they are instructed to repeat the same task as in the first round four more times.

one of reported beliefs for payment. The accuracy of the selected belief in the chosen round determines the chance of winning a prize. This winning probability increases as participants report a higher probability of the Target Ball being yellow when it actually is yellow, and blue when it is blue. This mechanism has been widely used in studies because it ensures that any decision-maker aiming to maximize their chance of winning a prize will report their beliefs truthfully, regardless of their risk preferences.[4]

We do not provide participants with the mathematical details of the scoring rule unless requested, as Danz et al. (2022) revealed that providing these details is not necessary for participants to reveal their true beliefs. Instead, we emphasize to the participants that the payment rule is designed to allow them to maximize their chances of winning the prize by reporting their most accurate beliefs.

2.4 Procedure

The experiment was conducted on Prolific, an online platform for recruiting research participants. A total of 299 participants took part in the experiment, with 200 assigned to the Variable Reward scheme (151 in the Color First treatment and 97 in the Box First treatment) and 99 assigned to the Fixed Reward scheme (48 in the Color First treatment and 51 in the Box First treatment). On average, participants spent 12 minutes to complete the experiment and earned \$2.7, including the participation fee.

3 Framework

In this section, we present a framework to validate our hypothesis on how the order of information affects the degree of motivated reasoning.

Our framework is built on the literature on the primacy effect — a well-established cogni-

[4] The implementation of the binarized scoring rule is detailed in Appendix A.

tive bias where individuals tend to pay more attention to the first piece of information they encounter compared to information presented later (Asch, 1946; Anderson and Barrios, 1961; Anderson, 1965; Yates and Curley, 1986) (for a review, see Benjamin, 2019).

We assume individuals are more likely to "encode" the informativeness of the signal when it is presented first, compared to when the signal itself is presented first.

3.1 Setup

Our framework assumes that individuals always encode the *color* of the Clue Ball (i.e., the signal), but may not always encode the *box* from which it is drawn (i.e., the informativeness of the signal). This assumption reflects that individuals might fail to recall the source of a story while correctly remembering its contents, aligning with previous research indicating that individuals may choose to ignore the informativeness of a signal to maintain their preferred beliefs (Kieren and Weber, 2021; Thaler, 2023).

Without loss of generality, in this section, we assume the individual is informed that the Clue Ball is yellow. We will examine the individual's posterior belief regarding the probability that the Target Ball is yellow, denoted as $\pi^{post}(y)$, after the information assessment process.

$$\underbrace{\left[\frac{\pi^{post}(y)}{1-\pi^{post}(y)}\right]}_{\text{Posterior Odds}} = \underbrace{\left[\frac{\pi^{prior}(y)}{1-\pi^{prior}(y)}\right]}_{\text{Prior Odds}} \times \underbrace{\left[\frac{Pr(Clue\ Ball = y \mid \tilde{I}, Target\ Ball = y)}{Pr(Clue\ Ball = y \mid \tilde{I}, Target\ Ball = b)}\right]}_{\text{Information Assessment}}. \tag{1}$$

Equation 1 shows how the information assessment process affects the odds of beliefs based on Bayes' rule. $I \in \{Box\ A,\ Box\ B\}$ indicates the box from which the Clue Ball is drawn, representing the signal's informativeness. To incorporate a possibility that the individual fails to encode the informativeness I, we denote the perceived informativeness by $\tilde{I} \in \{Box\ A,\ Box\ B\}$. If the informativeness is encoded, the individual identifies the box correctly, i.e., $\tilde{I} = I$. If the informativeness is not encoded, the individual assigns non-degenerated probabilities to the two boxes. The detailed probability assignment process will be discussed in the following section.

Let $Pr_{treatment}(\tilde{I} = I)$ represent the probability of successfully encoding informativeness for each treatment. We assume individuals exhibit the primacy effect, paying more attention to the first piece of information. Since informativeness is present in the first piece of information in the Box First treatment but not in the Color First treatment, we expect $Pr_{BoxFirst}(\tilde{I} = I) > Pr_{ColorFirst}(\tilde{I} = I)$.

The information assessment term $Pr(\mathit{Clue\ Ball} = y \mid \tilde{I}, \mathit{Target\ Ball} = y\ (b)$ refers to the probability that the Clue Ball is yellow given that it is drawn from $\tilde{I}$, and the Target Ball is yellow (blue). It determines the extent to which the posterior odds are adjusted relative to the prior odds. We discuss the information assessment in detail in the following section, with a particular focus on perceived informativeness.

3.2 Information Assessment

To discuss the perceived informativeness, we expand the information assessment term in Equation 1, with the expanded expression presented as follows:

$$\begin{aligned} &\frac{Pr(\mathit{Clue\ Ball} = y \mid \tilde{I}, \mathit{Target\ Ball} = y)}{Pr(\mathit{Clue\ Ball} = y \mid \tilde{I}, \mathit{Target\ Ball} = b)} = \\ &\qquad Pr(\text{Box A}) \times \left[\frac{Pr(\mathit{Clue\ Ball} = y \mid \text{Box A}, \mathit{Target\ Ball} = y)}{Pr(\mathit{Clue\ Ball} = y \mid \text{Box A}, \mathit{Target\ Ball} = b)}\right] \\ &\qquad + Pr(\text{Box B}) \times \left[\frac{Pr(\mathit{Clue\ Ball} = y \mid \text{Box B}, \mathit{Target\ Ball} = y)}{Pr(\mathit{Clue\ Ball} = y \mid \text{Box B}, \mathit{Target\ Ball} = b)}\right] \end{aligned} \tag{2}$$

Pr(Box A (B)) captures the probability that the Clue Ball is drawn from box Box A (B). This probability depends on whether the informativeness is encoded. If the informativeness is encoded, individuals correctly identify the probability. For example, $Pr(\text{Box A}) = 1$ and $Pr(\text{Box B}) = 0$ when the individual is informed that the Clue Ball is from Box A. (See the first two rows in Table 2.) In this case, the odds ratio of the information assessment term is 1 because she knows the color of the Target Ball is independent of the color of the Clue

Ball.[5] On the other hand, when she is informed that the Clue Ball is from Box B and the informativeness is encoded, the odds ratio would be 3.[6]

When the informativeness is not encoded, individuals assign probabilities to whether the Clue Ball is from Box A or B. This process of assigning probabilities can be distorted to reinforce individuals' preferred beliefs (Thaler, 2023). To capture this formally, we introduce a parameter that represents the level of motivated reasoning, denoted by $m \in (0,1)$. When $m = 1$, the individual correctly perceives that the Clue Ball comes from Box A or Box B with equal probability. However, as m approaches 0, the probability assignment becomes biased, favoring the individual's preferred beliefs. Specifically, when the signal is good news, the individual is more likely to believe it is informative: When the color of the Clue Ball is the preferred color, she is more likely to believe it is from Box B. Conversely, when the signal is bad news, she is more likely to believe it is uninformative: If the color of the Clue Ball is the non-preferred color, she is more likely to think it is from Box A. The last two rows in Table 2 summarize these cases. When the informativeness is not encoded, the odds ratio of the information assessment term varies between 1 and 3.

		I	Preferred Color		Non-Preferred Color	
			Pr(Box A)	Pr(Box B)	Pr(Box A)	Pr(Box B)
$\tilde{I}$	Encoded	Box A	1	0	1	0
		Box B	0	1	0	1
	Not Encoded	Box A	$1-0.5^m$	0.5^m	0.5^m	$1-0.5^m$
		Box B	$1-0.5^m$	0.5^m	0.5^m	$1-0.5^m$

Table 2: Perceived informativeness

In summary, our framework assumes that individuals do not always encode informativeness accurately, and we expect the probability of successful encoding to be higher in the Box

[5] $Pr(\textit{Clue Ball} = y \mid \text{Box A}, \textit{Target Ball} = y) = Pr(\textit{Clue Ball} = y \mid \text{Box A}, \textit{Target Ball} = b) = 0.5$

[6] $Pr(\textit{Clue Ball} = y \mid \text{Box B}, \textit{Target Ball} = y) = 0.75$, and $Pr(\textit{Clue Ball} = y \mid \text{Box B}, \textit{Target Ball} = b) = 0.25$.

First treatment than in the Color First treatment. Additionally, when informativeness is not successfully encoded, we anticipate that individuals are more likely to encode it as if they had received their preferred signal. Overall, our framework suggests a higher likelihood of biased encoding of informativeness in the Color First treatment, compared to the Box First treatment, leading to motivated reasoning.

4 Results

4.1 Prior Beliefs

Our analysis begins with an examination of participants' prior beliefs. Since participants do not receive any information at the prior stage, regardless of the treatment, we do not expect differences in prior beliefs between the Box First and Color First treatments. As expected, we find no difference. On average, participants in the Box First treatment believe the Target Ball is yellow with a probability of 50.8%, while the mean prior belief in the Color First treatment is 49.3%. The F-test confirms that the difference is not statistically significant (F-statistic: 0.9, p-value: 0.343).[7]

On the other hand, systematic differences in prior beliefs may exist across reward schemes and participants' choices of their preferred color (Coutts, 2019b). In the Variable Reward scheme, participants may be more likely to believe in scenarios where their reward is larger. Thus, participants who chose yellow as their preferred color (yellow choosers) would believe that the Target Ball is more likely to be yellow, while those who chose blue as their preferred color (blue choosers) would believe it is more likely to be blue. However, we do not expect this "wishful thinking" to occur in the Fixed Reward scheme, where the prize size is the same regardless of participants' preferred color.

Figure 2 presents the box plots of prior beliefs reported by participants, grouped by reward

[7] We control for round dummies, and the standard errors are clustered at the individual level.

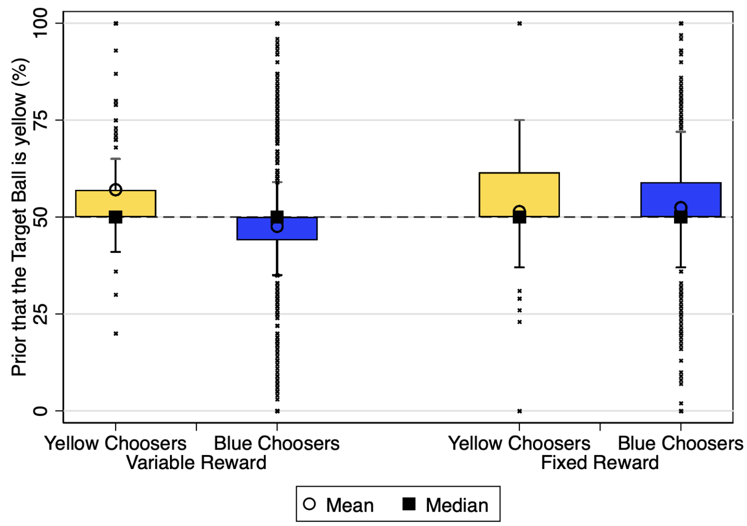


Figure 2: Prior beliefs that the Target Ball is yellow

schemes and their choices of preferred color. Among the 200 participants in the Variable Reward scheme, 30 (15%) were yellow choosers, and 170 (85%) were blue choosers. Yellow choosers reported a mean prior belief of 57.0% that the Target Ball is yellow, while blue choosers had a mean prior belief of 47.6%. The difference is statistically significant (F-statistic: 13.93, p-value < 0.001), confirming the presence of wishful thinking in the Variable Reward scheme. In the Fixed Reward scheme, 12 (12.1%) were yellow choosers and 87 (87.8%) were blue choosers. Their prior beliefs were 52.4% and 51.4%, respectively. As expected, the difference is not statistically significant (F-statistic: 0.07, p-value $= 0.788$).

Figure 2 highlights another key observation: priors are widely distributed, with some participants holding extreme beliefs. This large variance could present a potential issue, as participants with stronger priors and those with more moderate priors may behave systematically differently. For instance, participants who believe the Target Ball is 90% likely to be yellow, despite having no information about Box B, may discount the information from the signal more than those whose prior is 50% (Bisière et al., 2015).

To investigate this potential issue, we classify participants into two groups based on the strength of their prior beliefs: those with moderate priors (within the 25th–75th percentile) and those with extreme priors (outside this range). Of the total 1,495 decisions (299 participants, each making 5 rounds), 802 (53.6%) are classified as moderate priors, while 693 (46.4%) are classified as extreme priors. We find no statistically significant association between whether participants' priors are moderate or extreme and the order of information.[8]

For the remainder of the paper, we define beliefs in a way that aligns with the signal participants receive to simplify our analysis. When a participant observes a yellow (blue) Clue Ball, their belief corresponds to the belief that the Target Ball is yellow (blue). This approach has two advantages. First, regardless of whether the signal is yellow or blue, we will observe participants' beliefs moving in a single direction (unless they respond contrary to the signal). Therefore, in the analysis, we need only consider updates in one direction. Second, it allows us to examine how a preferred signal is associated with the extent of belief updating. Motivated reasoning suggests that participants update their beliefs more in line with the signal when it matches their preferred color compared to when it does not. Thus, this approach not only simplifies our analysis but also enables us to examine motivated reasoning in belief updating.

4.2 Belief Updating Patterns

Our central question of this paper is how the timing of informativeness influences motivated reasoning. As a first step in exploring this, we examine belief updating patterns based on the order in which information is presented. Since we expect motivated reasoning to emerge only in the Variable Reward scheme, the discussion here is focused solely on that scheme.[9]

4.2.1 Belief Updating with Uninformative Signal

When the signal is uninformative, a Bayesian's posterior belief should remain the same as their prior belief. However, in reality, individuals may overweight their preferred signal even

[8] The F-statistic testing the association between having moderate priors and assignment to the Color First or Box First treatment is 0.1 (p-value =0.75).

[9] As expected, we found no evidence of motivated reasoning in the Fixed Reward scheme, so the timing of informativeness had no effect. For more details, see Appendix B.

when the signal is uninformative (Thaler, 2023). In this section, we examine how the timing of informativeness influences belief updating when the signal is uninformative, with a particular focus on motivated reasoning.

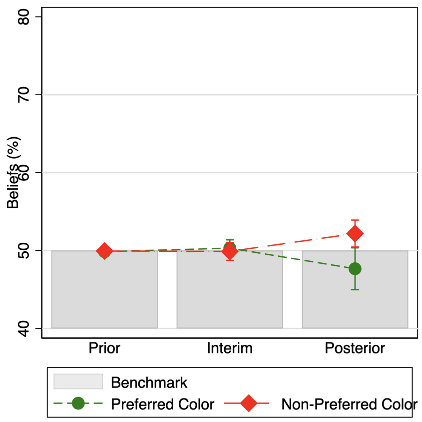
Figure 3 illustrates the changes of participants' average prior beliefs, interim beliefs, and posterior beliefs when the Clue Ball is from Box A, i.e., when the signal is uninformative. Note that the Y-axis represents the belief that the Target Ball is yellow for participants who observe a yellow Clue Ball, and blue for those who observe a blue Clue Ball.

The gray bars represent Bayesian benchmarks. At the prior stage, a Bayesian believes the Target Ball is equally likely to be yellow or blue. In the Box First treatment (Figure 3a and 3b), interim beliefs are expected to remain the same as the prior since the color of the Clue Ball has not yet been revealed. On the other hand, in the Color First treatment (Figure 3c and 3d), the color of the Clue Ball is revealed at the interim stage, prompting the Bayesian benchmarks to adjust the belief based on the signal. However, because the informativeness of the signal is still unknown, the update is less certain than with complete information. At the posterior stage, in both treatments, the Bayesian benchmarks return to the prior, as the signal is revealed to be uninformative.

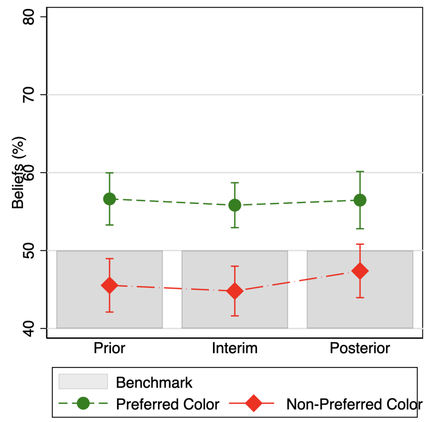
Green dots represent reported beliefs when the signal —the color of the Clue Ball — matches to the preferred color, while red dots represent beliefs when the signal indicates a non-preferred color. Regardless of whether their prior beliefs are moderate or extreme, participants generally update their beliefs in their interim and posterior in a way that aligns with the Bayesian benchmark.[10] For participants with extreme priors, their prior beliefs differ substantially depending on whether the signal aligns with the preferred color. However, their belief updating patterns —the shift from prior to posterior belief in response to the signal — show no noticeable differences.

On the other hand, participants with moderate priors, whose prior do not vary regardless

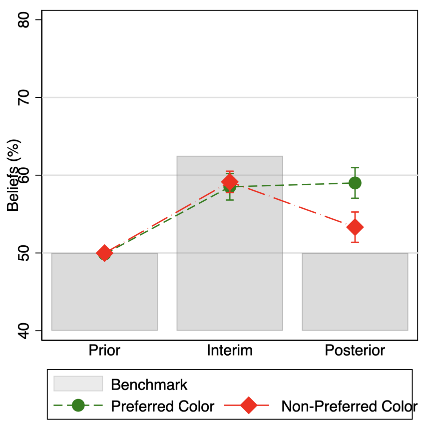
[10]See AppendixC for figures with pooled priors.



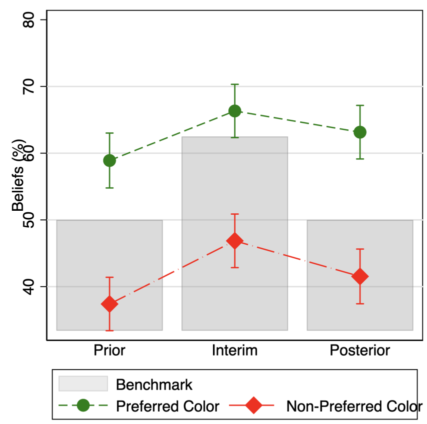
(a) Box First (Moderate)



(b) Box First (Extreme)



(c) Color First (Moderate)



(d) Color First (Extreme)

Figure 3: Belief updating with uninformative signal under the Variable Reward scheme

of whether the signal aligns with the preferred color, display distinct belief updating patterns: they adjust their posterior beliefs more toward the preferred signal in the Color First treatment compared to the Box First treatment. This difference in patterns suggests that the timing of informativeness could influence motivated reasoning. This will be examined in detail in Section 4.3. In the next section, we will explore whether this difference between treatments holds when the signal is informative.

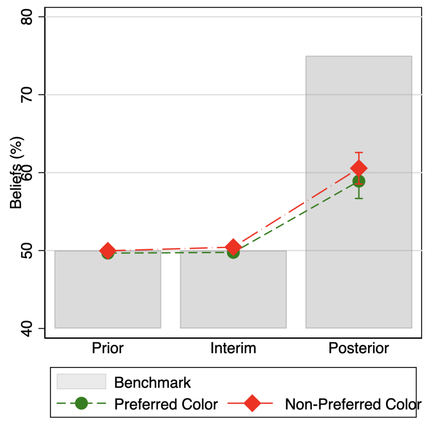
4.2.2 Belief Updating with Informative Signal

Previous studies on motivated reasoning suggest that when the signal is informative, individuals tend to update their beliefs more if the signal aligns with their preferences (Eil and Rao, 2011; Coutts, 2019b). In this section, we focus on how the timing of informativeness affects belief updating specifically when the signal is informative, particularly in the context of motivated reasoning.

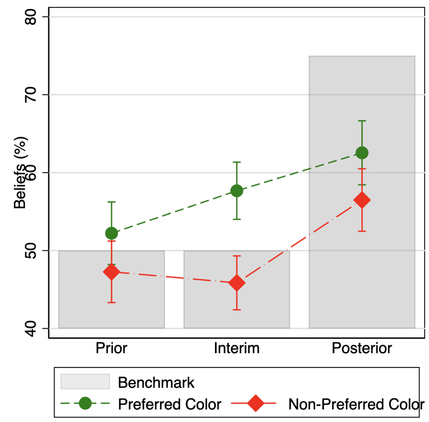
Figure 4 illustrates the belief updating patterns when the Clue Ball is from Box B, i.e., when the signal is informative. At the prior and interim stages, the Bayesian benchmark is identical to when the signal is uninformative. However, at the posterior stage, since the information from the Clue Ball is informative, the Bayesian updates their belief to align with the signal.

Similar to the case of the uninformative signal, the reported beliefs in Figure 4 when the signal is informative also suggest the possibility of motivated reasoning for participants with moderate priors: they update their posterior beliefs more in the direction of the preferred signal in the Color First treatment compared to the Box First treatment. Additionally, as expected, no such difference between treatments is observed for participants with extreme priors: there appears to be no significant difference in how posterior beliefs shift from priors between treatments.

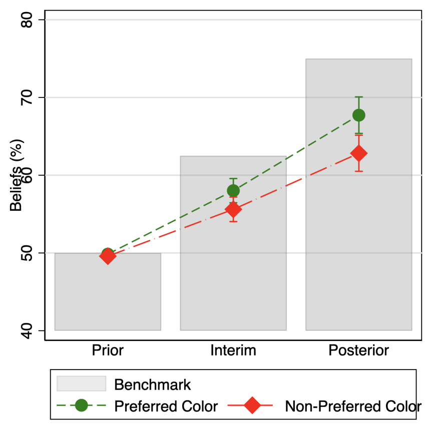
These results imply two key conditions for the effect of the timing of informativeness on motivated reasoning: first, the effect is independent of the signal's informativeness, and second, it occurs only for participants with moderate priors. A detailed analysis will be provided in the next section.



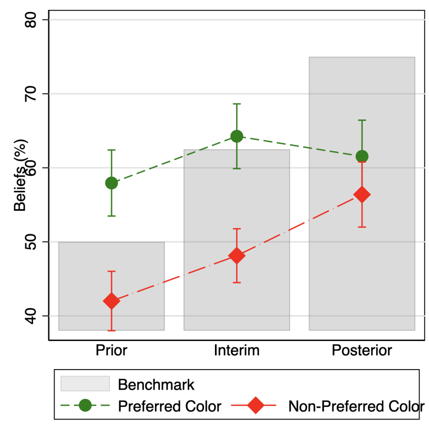
(a) Box First (Moderate)



(b) Box First (Extreme)



(c) Color First (Moderate)



(d) Color First (Extreme)

Figure 4: Belief updating with informative signal under the Variable Reward scheme

4.3 Effect of the Timing of Informativeness on Motivated Reasoning

In this section, we formally investigate the order effect, i.e., the effect of the timing of informativeness on motivated reasoning, using the following regression model:[11]

$$\begin{aligned} LogOddsPosterior_i &= \alpha_0 + \alpha_1 LogOddsPrior_i + \alpha_2 PreferredColor_i \\ &+ \alpha_3 ColorFirst_i + \alpha_4 PreferredColor_i \times ColorFirst_i + Controls_i + \epsilon_i. \end{aligned} \tag{3}$$

[11]We use log odds beliefs instead of raw beliefs for better interpretability, especially for probabilities near 0 or 1, and to enable consistent analysis of belief updates. We replaced 0 (%) with 0.1 (%) and 100 (%) with 99.9 (%) to compute the log ratio, as 0 and 100 cannot be directly used in logarithmic calculations.

As discussed in the previous section, we construct belief variables to align with the color of the Clue Ball. For participants observing a yellow Clue Ball, the odds of beliefs represent the ratio of the posterior probability that the Target Ball is yellow to the posterior probability that it is blue. For those observing a blue Clue Ball, the ratio reflects the posterior probability of the Target Ball being blue relative to it being yellow. This ensures that $LogOddsPosterior_i$ moves in the same direction controlling for $LogOddsPrior_i$, regardless of the color of the Clue Ball.

$PreferredColor_i$ indicates whether the color of the Clue Ball matches the participant i's preferred color, and $ColorFirst_i$ denotes whether participant i is in the Color First treatment. Therefore, the interaction term $PreferredColor_i \times ColorFirst_i$ therefore captures the effect of the timing of informativeness on the degree of motivated reasoning. If motivated reasoning is stronger in the Color First treatment than in the Box First treatment, the coefficient α_4 would be positive. $Controls$ includes round dummies and the indicator of the chosen color.

	DV: Log Odds Posterior					
	(1) All	(2) All	(3) Moderate	(4) Moderate	(5) Extreme	(6) Extreme
Log Odds Prior	0.535***	0.533***	-0.420	-0.562	0.534***	0.532***
	(0.076)	(0.076)	(2.823)	(2.861)	(0.076)	(0.077)
Preferred Color	0.201	-0.056	0.211	-0.198	0.191	0.073
	(0.158)	(0.238)	(0.133)	(0.195)	(0.309)	(0.449)
Color First		0.170		0.273		0.044
		(0.226)		(0.185)		(0.430)
Preferred Color × Color First		0.482		0.696***		0.245
		(0.328)		(0.265)		(0.642)
Uninformative Box	-0.553***	-0.554***	-0.567***	-0.543***	-0.524*	-0.533*
	(0.154)	(0.153)	(0.163)	(0.154)	(0.278)	(0.280)
Observations	1000	1000	535	535	465	465
Num. of Individuals	200	200	139	139	124	124

Table 3: The order effect and motivated reasoning (Variable Reward scheme)

Table 3 presents the regression estimates in the Variable Reward scheme. Columns (1) and (2) show the overall results, suggesting that prior beliefs positively influence posterior beliefs, and participants update less when they receive an uninformative signal (i.e., the Clue Ball is from Box A). Columns (3) and (4) display the results for participants with moderate priors. Their first feature is that is that their posterior beliefs do not depend on the prior belief. Secondly, and most interestingly, as predicted in the previous section, the timing of informativeness influences motivated reasoning. Participants update their beliefs significantly more toward the preferred signal when they are informed of the informativeness first, i.e., in the Color First treatment. This pattern is not observed for participants with extreme priors, as indicated by the regression results in columns (5) and (6). Results show that they rely more on their prior beliefs and pay less attention to the signal's informativeness, compared to participants with moderate priors. Figure 5, which illustrates the estimated marginal effects, verifies our findings. In the Color First treatment, the positive signal increases the log posterior odds by 0.968 (p=0.003) for participants with moderate priors, while no significant effect is observed in other cases.

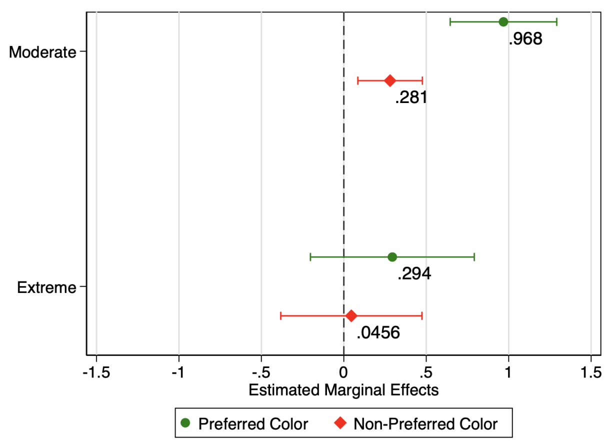


Figure 5: Marginal order effect (Variable Reward scheme)

Table 4 shows the the regression estimates in the Fixed Reward scheme. Unlike the results from the Variable Reward scheme, the order effect does not influence motivated reasoning here. Figure 6 further confirms this finding, showing that the marginal effects are not statistically significant in any case.

	DV: Log Odds Posterior					
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Moderate	Moderate	Extreme	Extreme
Log Odds Prior	0.350**	0.348**	-0.310	-0.492	0.331**	0.325**
	(0.136)	(0.136)	(4.208)	(4.341)	(0.134)	(0.132)
Preferred Color	-0.337	-0.186	-0.068	0.233	-0.611	-0.744
	(0.223)	(0.303)	(0.243)	(0.273)	(0.410)	(0.649)
Color First		0.047		0.389		-0.501
		(0.439)		(0.525)		(0.660)
Preferred Color × Color First		-0.327		-0.686		0.149
		(0.438)		(0.530)		(0.772)
Uninformative Box	-0.378*	-0.372*	-0.616**	-0.612***	-0.054	0.017
	(0.203)	(0.209)	(0.236)	(0.231)	(0.355)	(0.391)
Observations	495	495	267	267	228	228
Num. of Individuals	99	99	67	67	62	62

Table 4: The order effect and motivated reasoning (Fixed Reward scheme)

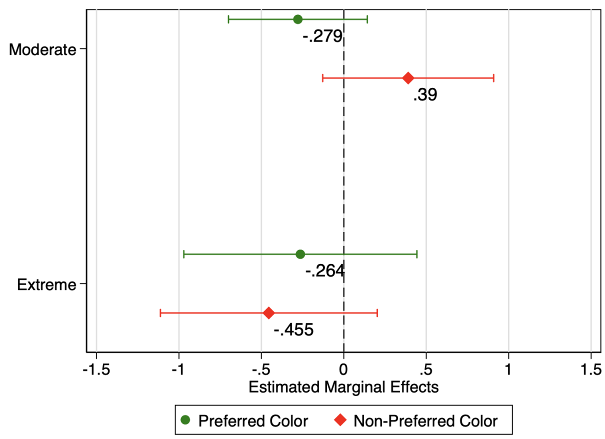


Figure 6: Marginal order effect (Fixed Reward scheme)

5 Conclusion

This paper examines the effect of the timing of informativeness on motivated reasoning. Our experiment, employing a variant of the classic balls-and-boxes setup, demonstrates that participants exhibit motivated reasoning when they encounter information before learning about its informativeness, compared to when the order is reversed. This finding highlights how the sequence of information influences belief updating.

Our results suggest that exposure to information without knowing its informativeness may encourage individuals to interpret it in a way that aligns with their preexisting preferences. This has broader implications for potential interventions to curb motivated reasoning, as simply altering the sequence of information could help mitigate its effects.

These insights pave the way for future research in several ways. Firstly, our findings can be extended to other types of motivated reasoning. For example, politicized motivated reasoning can occur even without monetary incentives (Kunda, 1990; Westen et al., 2006; Kim et al., 2010; Kahan, 2016; Thaler, 2023). In such cases, it would be valuable to investigate whether the timing of informativeness also influences the bias. Secondly, this research has practical implications in real-world settings for reducing cognitive biases. By strategically controlling the sequence in which information is presented, it may be possible to mitigate these biases, leading to better decision-making processes.

References

Ahlering, R. F. and Parker, L. (1989). Need for cognition as a moderator of the primacy effect. *Journal of Research in Personality*, 23:313–317.

Anderson, N. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of personality and social psychology*, 34:1–9.

Anderson, N. and Barrios, A. (1961). Primacy effects in personality impression formation. *The Journal of Abnormal and Social Psychology*, 63(2):346–350.

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, 41(3):258–290.

Beattie, P. and Snider, D. (2019). Knowledge in international relations: Susceptibilities to motivated reasoning among experts and non-experts. *Journal of Social and Political Psychology*.

Benjamin, D. J. (2019). Chapter 2 - Errors in probabilistic reasoning and judgment biases. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics - Foundations and Applications 2*, volume 2 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 69–186. North-Holland.

Bisière, C., Décamps, J.-P., and Lovo, S. (2015). Risk attitude, beliefs updating, and the information content of trades: An experiment. *Management Science*, 61(6):1378–1397.

Bolsen, T. and Palm, R. (2019). Motivated reasoning and political decision making.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Carpenter, C. J. (2019). Cognitive dissonance, ego-involvement, and motivated reasoning. *Annals of the International Communication Association*, 43:1 – 23.

Cotter, R., Lodge, M., and Vidigal, R. (2020). The boundary conditions of motivated reasoning.

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

Coutts, A. (2019b). Testing models of belief bias: An experiment. *Games and Economic Behavior*, 113:549–565.

Danz, D., Vesterlund, L., and Wilson, A. J. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review*, 112(9):2851–2883.

Ecker, U., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L., Brashier, N., Kendeou, P., Vraga, E., and Amazeen, M. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1:13–29.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–138.

Epley, N. and Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(2):133–140.

Gonçalves, D., Libgober, J., and Willis, J. (2021). Learning versus unlearning: An experiment on retractions. *Working Paper*.

Hogarth, R. M. and Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24(1):1–55.

Hossain, T. and Okui, R. (2013). The binarized scoring rule. *The Review of Economic Studies*, 80:984–1001.

Jones, E. E., Rock, L., Shaver, K. G., Goethals, G. R., and Ward, L. M. (1968). Pattern of performance and ability attribution: an unexpected primacy effect. *Journal of Personality and Social Psychology*, 10(4):317–340.

Kahan, D. (2016). The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. pages 1–16.

Kieren, P. and Weber, M. (2021). Expectation formation under uninformative signals. *Working Paper*.

Kim, S., Taber, C., and Lodge, M. (2010). A computational model of the citizen as motivated reasoner: Modeling the dynamics of the 2000 presidential election. *Political Behavior*, 32:1–28.

Klaczynski, P. A. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development*, 71(5):1347–1366.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, 108 3:480–98.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2022). Managing self-confidence: Theory and experimental evidence. *Management Science*, 68(11):7793–7817.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? testing the effects of motivated reasoning on political decision making. *The Journal of Politics*, 64(4):1021–1044.

Saccardo, S. and Serra-Garcia, M. (2023). Enabling or limiting cognitive flexibility? evidence of demand for moral commitment. *American Economic Review*, 113:396–429.

Taber, C. S., Cann, D., and Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior*, 31(2):137–155.

Thaler, M. (2021). Gender differences in motivated reasoning. *Journal of Economic Behavior and Organization*, 191(C):501–518.

Thaler, M. (2023). The fake news effect: Experimentally identifying motivated reasoning using trust in news. *American Economic Journal: Microeconomics, forthcoming.*

van Erkel, P. F. and Thijssen, P. (2016). The first one wins: Distilling the primacy effect. *Electoral Studies*, 44:245–254.

Westen, D., Blagov, P., Harenski, K., Kilts, C., and Hamann, S. (2006). Neural bases of motivated reasoning: An fmri study of emotional constraints on partisan political judgment in the 2004 u.s. presidential election. *Journal of Cognitive Neuroscience*, 18:1947–1958.

Wittrock, L., Strobel, M., and Tsakas, E. (2023). Belief updating with misinformation. *Working Paper*.

Yates, J. and Curley, S. P. (1986). Contingency judgment: Primacy effects and attention decrement. *Acta Psychologica*, 62(3):293–302.

A Binarized Scoring Rule

Let $\pi_y \in [0, 1]$ and $\pi_b \in [0, 1]$ denote the participant's reported beliefs about the Target Ball being yellow and the Target Ball being blue, respectively. i.e., $\pi_y + \pi_b = 1$. The participant's payoff is as follows:

i. If the Target Ball is yellow

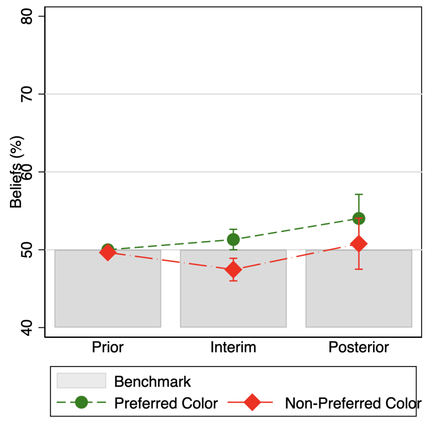
$$\text{Payoff} = \begin{cases} (\textit{High Prize}, 0;\ \ 1 - (1 - \pi_y)^2) & \text{if the preferred color is yellow} \\ (\textit{Low Prize}, 0;\ \ 1 - (1 - \pi_y)^2) & \text{if the preferred color is blue} \end{cases}$$

ii. If the Target Ball is blue

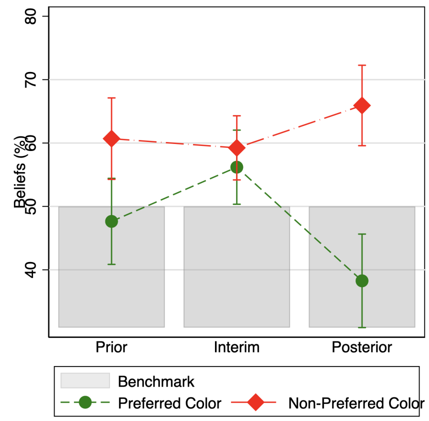
$$\text{Payoff} = \begin{cases} (\textit{High Prize}, 0;\ \ 1 - (1 - \pi_b)^2) & \text{if the preferred color is blue} \\ (\textit{Low Prize}, 0;\ \ 1 - (1 - \pi_b)^2) & \text{if the preferred color is yellow} \end{cases}$$

Note that *High Prize* is \$1.5 and *Low Prize* is \$0.5 in Variable Reward scheme, and both are \$1 in Fixed Reward scheme.

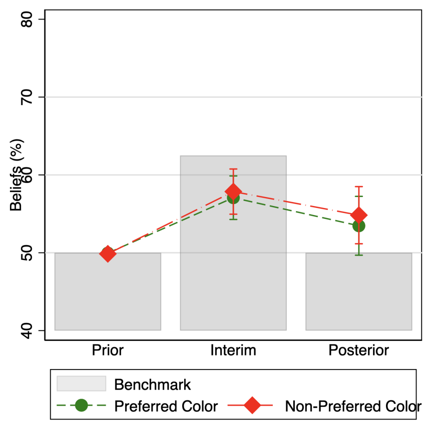
B Belief Updating under the Fixed Reward Scheme



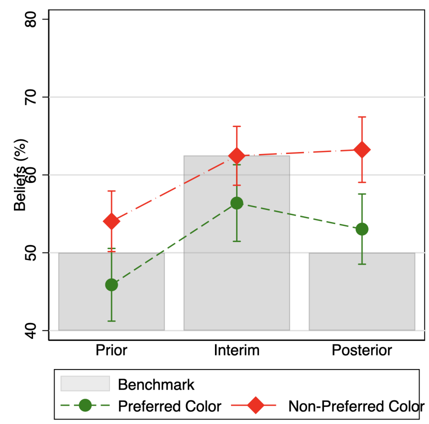
(a) Box First (Moderate)



(b) Box First (Extreme)



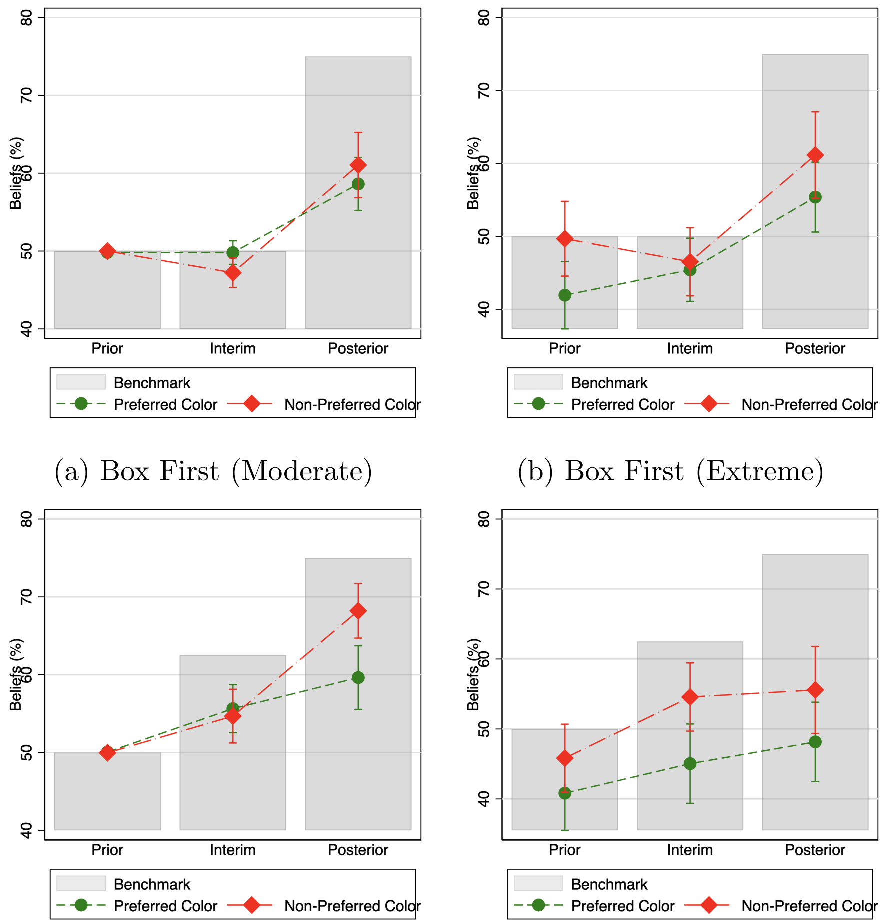
(c) Color First (Moderate)



(d) Color First (Extreme)

Figure 7: Belief updating with uninformative signal under the Fixed Reward scheme

Figure 7 and 8 show the belief updating patterns under the Fixed Reward scheme when the signal is uninformative and informative, respectively. Unlike in the Variable Reward scheme, even with moderate priors, noticeable differences in belief updating between treatments are not observed. This is explained further in Section 4.3.



(a) Box First (Moderate)

(b) Box First (Extreme)

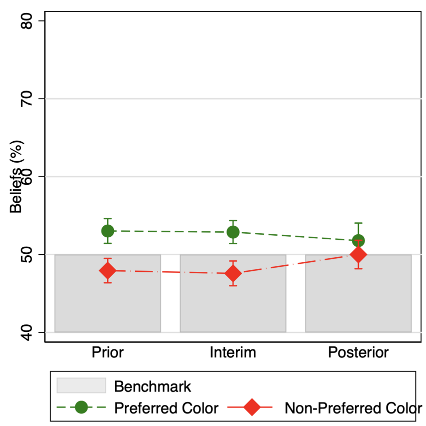
(c) Color First (Moderate)

(d) Color First (Extreme)

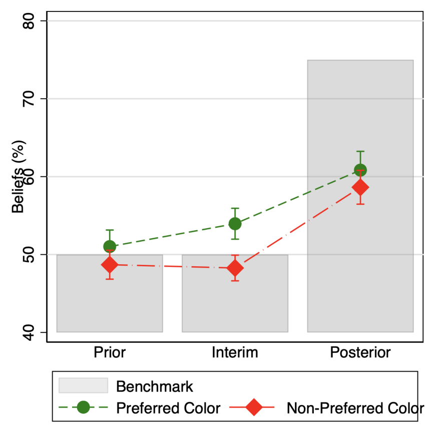
Figure 8: Belief updating with informative signal under the Fixed Reward scheme

C Belief Updating Patterns with Pooled Priors

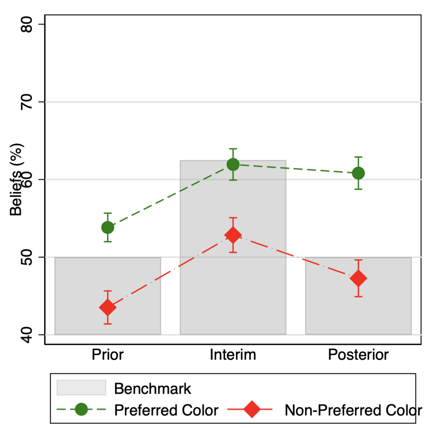
This section examines belief updating patterns for all participants, combining those with extreme and moderate priors. Overall, the results align with the belief updating patterns observed among participants with moderate priors: participants generally update their beliefs in line with Bayesian benchmarks, and the patterns suggest the presence of the order effect in the Variable Reward scheme, as discussed in Section 4.2.



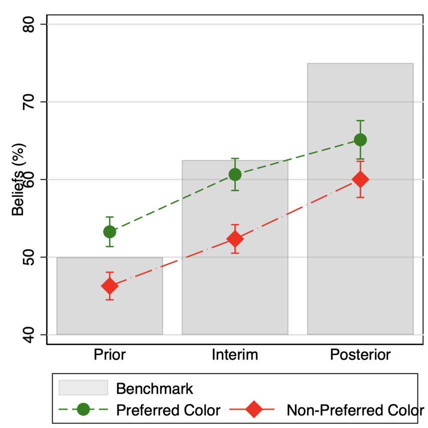
(a) Box First (uninformative)



(b) Box First (informative)

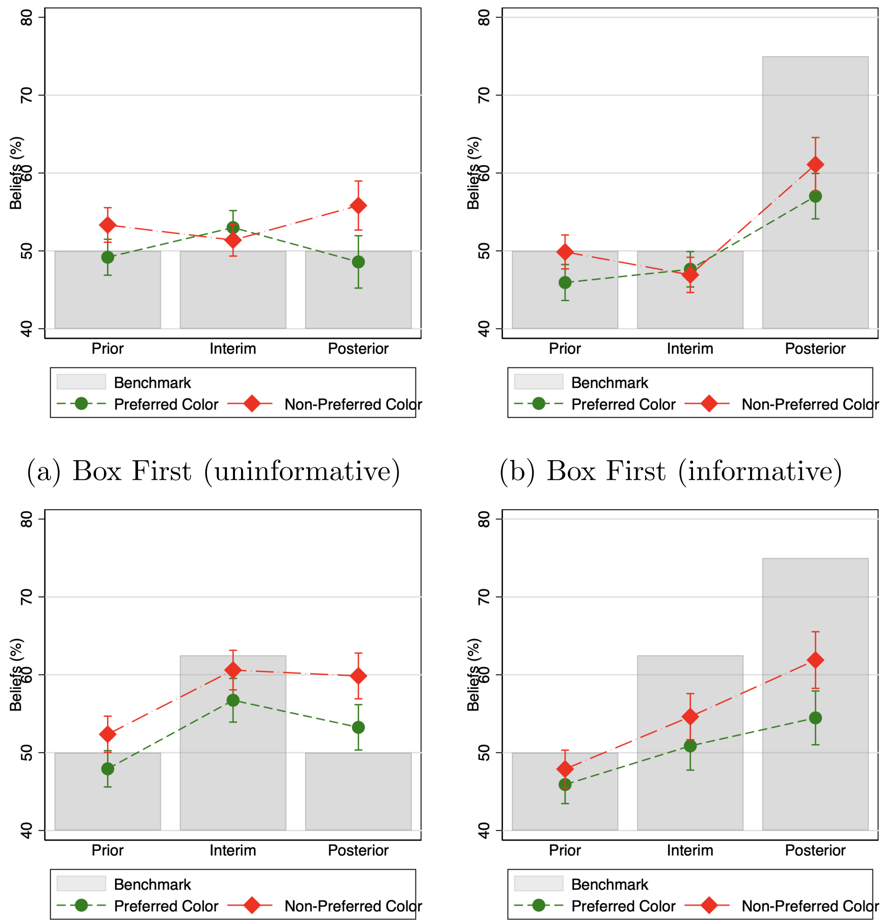


(c) Color First (uninformative)



(d) Color First (informative)

Figure 9: Belief updating under the Variable Reward scheme



(a) Box First (uninformative)

(b) Box First (informative)

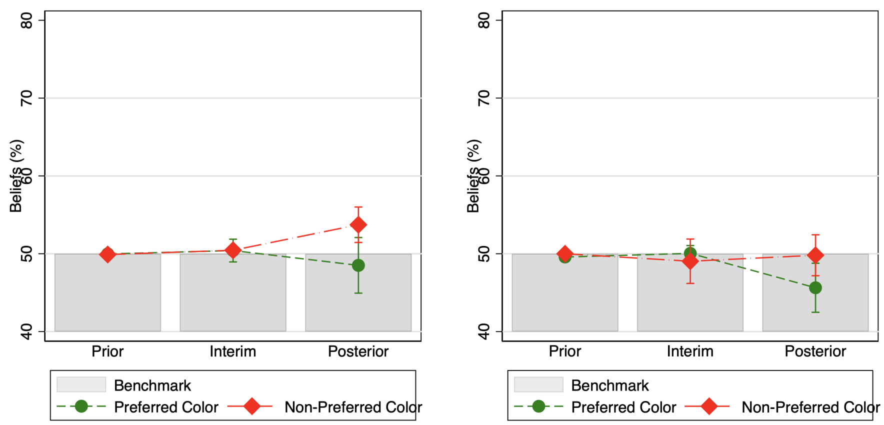
(c) Color First (uninformative)

(d) Color First (informative)

Figure 10: Belief updating under the Fixed Reward scheme

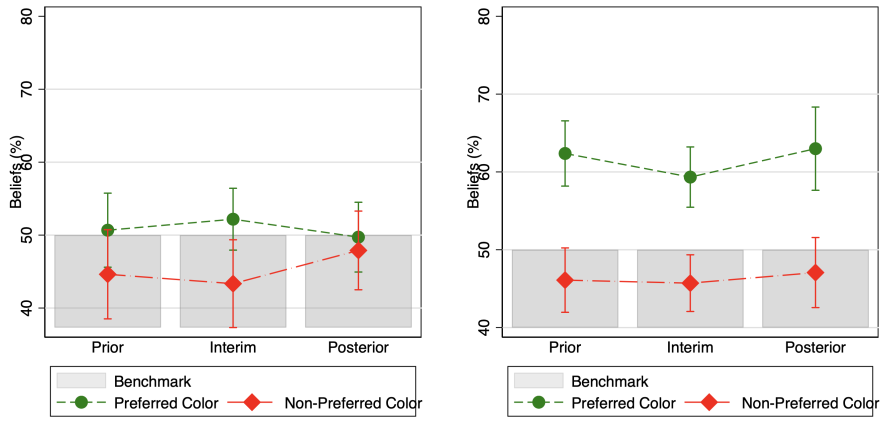
D Ambiguity in the Instructions

This section presents all the main figures for both versions of the instructions. In one version, participants are not informed about the composition of Box B: "Box B contains two balls; however, the composition is unknown." In the other version, we provide additional information about the distribution: "Each ball in Box B could be yellow or blue with equal probability." The figures show that, overall, there is no difference in belief updating patterns, regardless of whether ambiguity is present in the instructions or not.



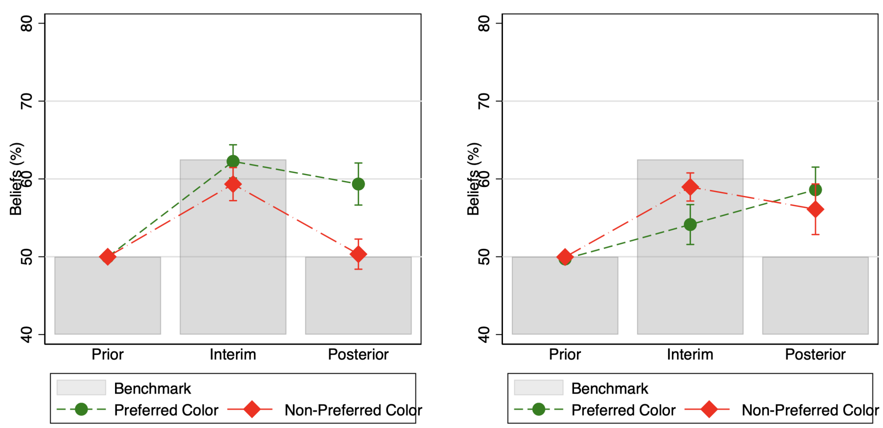
(a) with Ambiguity (b) without Ambiguity

Figure 11: Box First, Box A (Moderate)



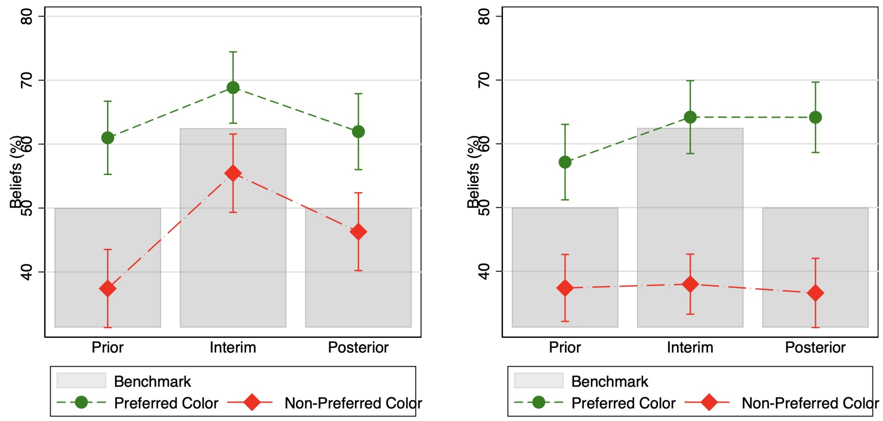
(a) with Ambiguity (b) without Ambiguity

Figure 12: Box First, Box A (Extreme)



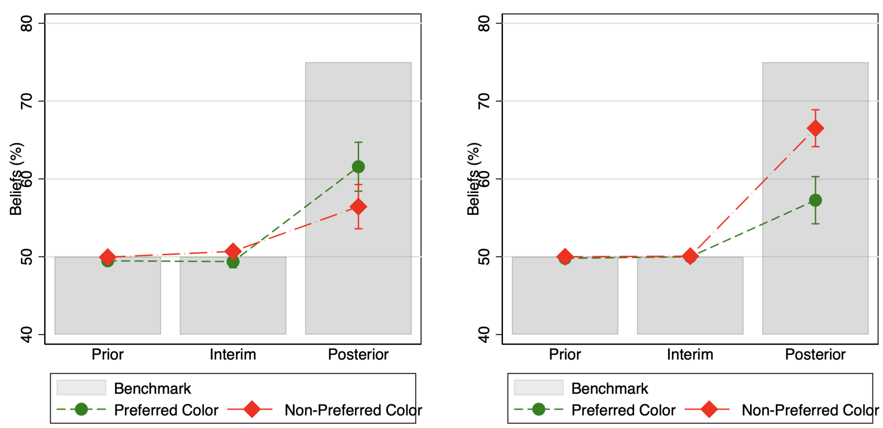
(a) with Ambiguity (b) without Ambiguity

Figure 13: Color First, Box A (Moderate)



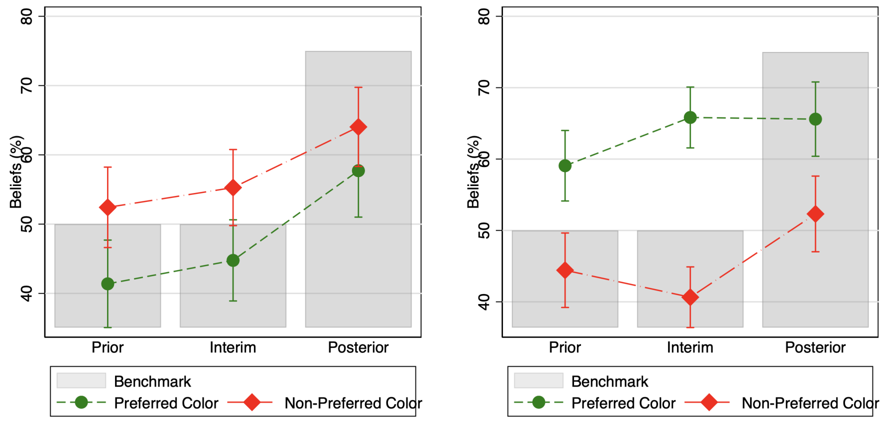
(a) with Ambiguity (b) without Ambiguity

Figure 14: Color First, Box A (Extreme)



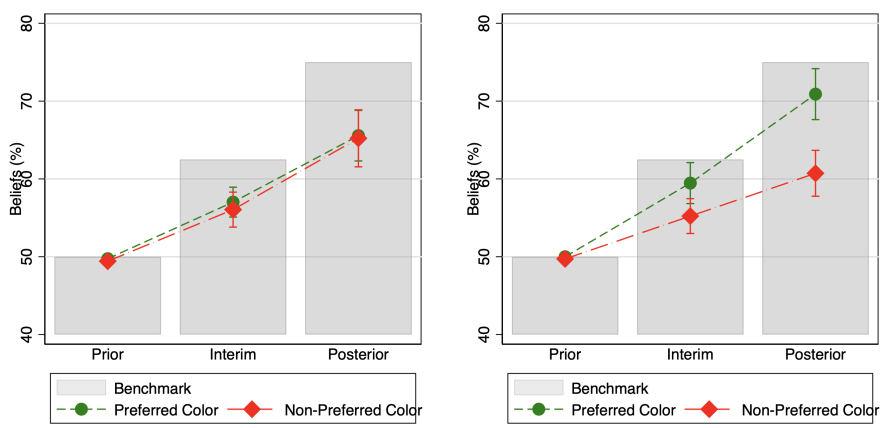
(a) with Ambiguity (b) without Ambiguity

Figure 15: Box First, Box B (Moderate)



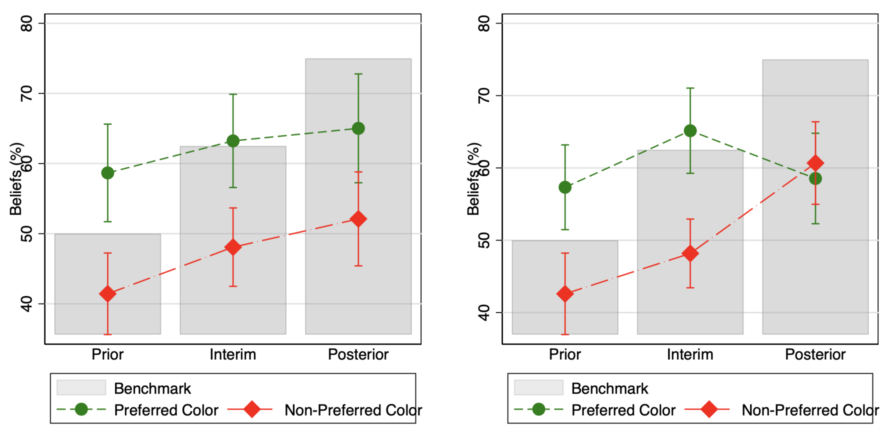
(a) with Ambiguity (b) without Ambiguity

Figure 16: Box First, Box B (Extreme)



(a) with Ambiguity (b) without Ambiguity

Figure 17: Color First, Box B (Moderate)



(a) with Ambiguity

(b) without Ambiguity

Figure 18: Color First, Box B (Extreme)

E Instructions and Screenshots

Figures 19 through 27 present the computer screen participants see in the experiment. Figure 19 is displayed only once at the start of the experiment. Upon entering Part 1, participants are asked to read the instructions carefully (Figures 20 - 22) and answer the comprehension questions. If participants fail to answer all the questions correctly, they will be sent back to the first page of the instructions and asked to review them. They cannot proceed until they answer all the questions correctly. After completing Part 1, participants proceed to Part 2 (Figure 27). The instructions (Figures 21 - 22) are displayed again at the beginning of Part 2.

Figure 19: Screenshot before starting the experiment

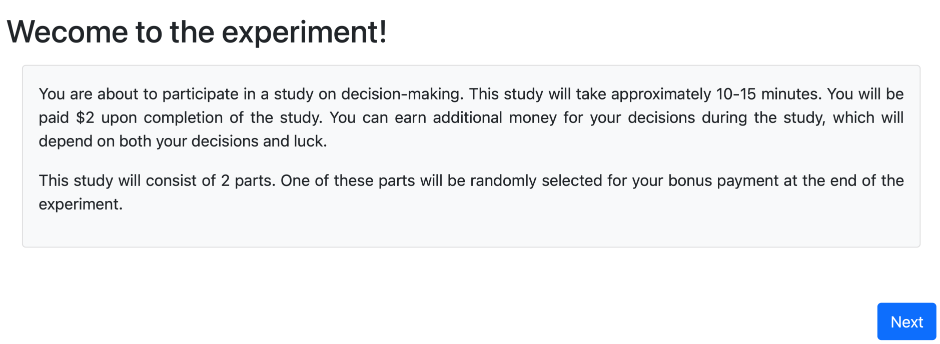


Figure 20: General instructions (1)

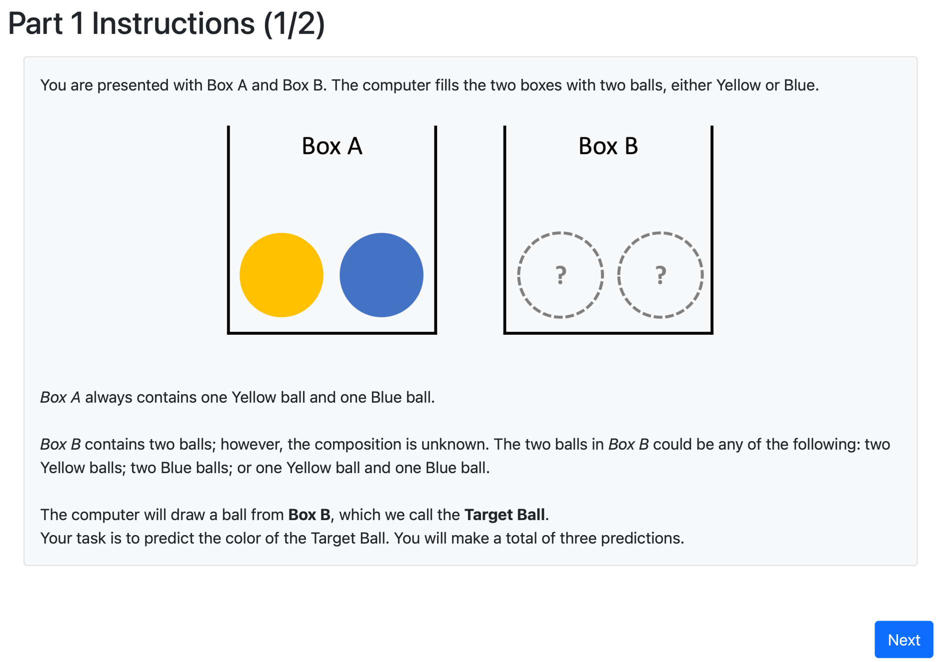


Figure 21: General instructions (2)

Part 1 Instructions (2/2)

The experiment will proceed as follows:

1. **First Guess.**

   The computer randomly draws a ball from *Box B*. The ball that is drawn is referred to as the *Target Ball*. You will provide your first prediction of the color of the *Target Ball*. The results of your prediction will not yet be displayed.

   After drawing the *Target Ball*, the computer returns it to *Box B*. The contents of *Box B* therefore remains the same.

2. **Computer Generates Clue.**

   After you complete your first guess, the computer will generate a clue for you. This clue is created by the computer randomly selecting either Box A or Box B. Both boxes have equal probability in being selected. Then, the computer randomly draws a ball from the selected box. Each ball in the selected box has equal probability of being picked.

   The ball that is selected is called the *Clue Ball*. After drawing the *Clue Ball*, the computer returns it to the selected box. The contents of the selected box therefore remains the same.

3. **Second Guess.**

   You will see the color of the *Clue Ball*.

   Then, you will provide your second prediction about the color of the *Target Ball*. The results of your prediction will not yet be displayed.

4. **Third Guess.**

   Lastly, you will see which box the *Clue Ball* was drawn from.

   Then, you will provide your third prediction about the color of the *Target Ball*. The results of your prediction will not yet be displayed.

5. **Payoff.**

   Based on the accuracy of your three predictions, you will have a chance to receive a prize. The prize amount depends on the color of the *Target ball* and the color you chose in the beginning of this experiment.

   If the color of the *Target ball* is Yellow, which *matches* your color choice, the prize is $1.5.
   If the color of the *Target ball* is Blue, which *does not match* your color choice, the prize is $0.5.

   The payment rule is designed so that you can secure the largest chance of winning the prize by reporting your most accurate guess. The precise payment rule details are available by request at the end of the experiment.

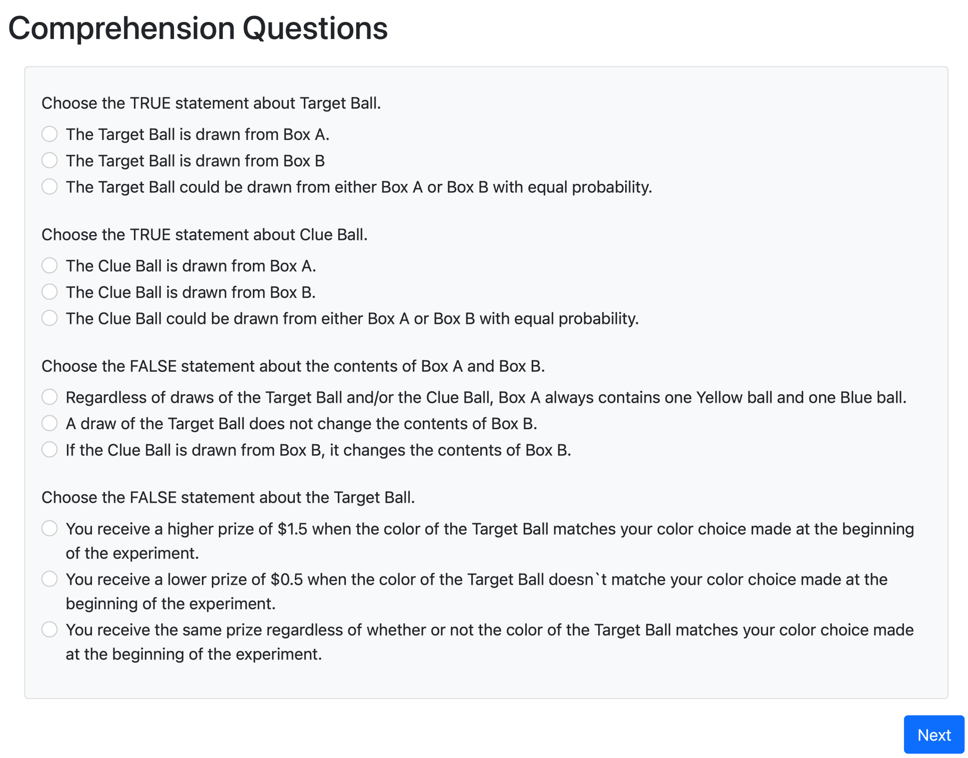
   At the end of the experiment, one of the three guesses will be randomly chosen for payment. Every guess has the same chance of being selected for payment.

On the next screen, you are presented with comprehension questions. If any of your submitted answers are incorrect, you cannot proceed. Please review the instructions. Once you get familiar with the instructions, click “Confirm”.

Back

Confirm

Figure 22: General instructions (3)



Comprehension Questions

Choose the TRUE statement about Target Ball.

- The Target Ball is drawn from Box A.
- The Target Ball is drawn from Box B
- The Target Ball could be drawn from either Box A or Box B with equal probability.

Choose the TRUE statement about Clue Ball.

- The Clue Ball is drawn from Box A.
- The Clue Ball is drawn from Box B.
- The Clue Ball could be drawn from either Box A or Box B with equal probability.

Choose the FALSE statement about the contents of Box A and Box B.

- Regardless of draws of the Target Ball and/or the Clue Ball, Box A always contains one Yellow ball and one Blue ball.
- A draw of the Target Ball does not change the contents of Box B.
- If the Clue Ball is drawn from Box B, it changes the contents of Box B.

Choose the FALSE statement about the Target Ball.

- You receive a higher prize of $1.5 when the color of the Target Ball matches your color choice made at the beginning of the experiment.
- You receive a lower prize of $0.5 when the color of the Target Ball doesn`t matche your color choice made at the beginning of the experiment.
- You receive the same prize regardless of whether or not the color of the Target Ball matches your color choice made at the beginning of the experiment.

Next

Figure 23: Comprehension questions

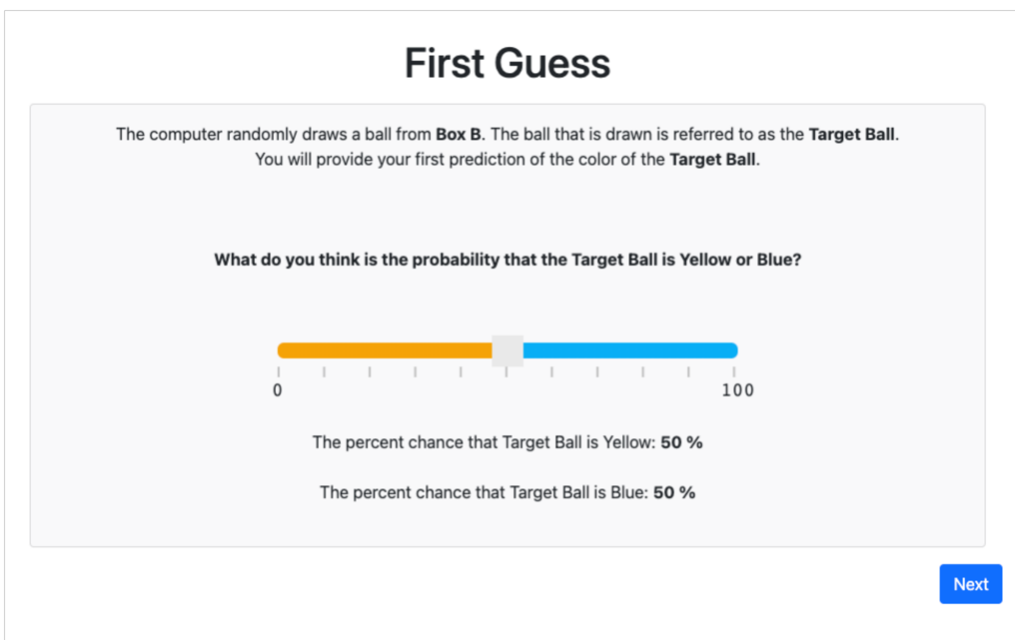


Figure 24: Screenshot of the prior belief elicitation

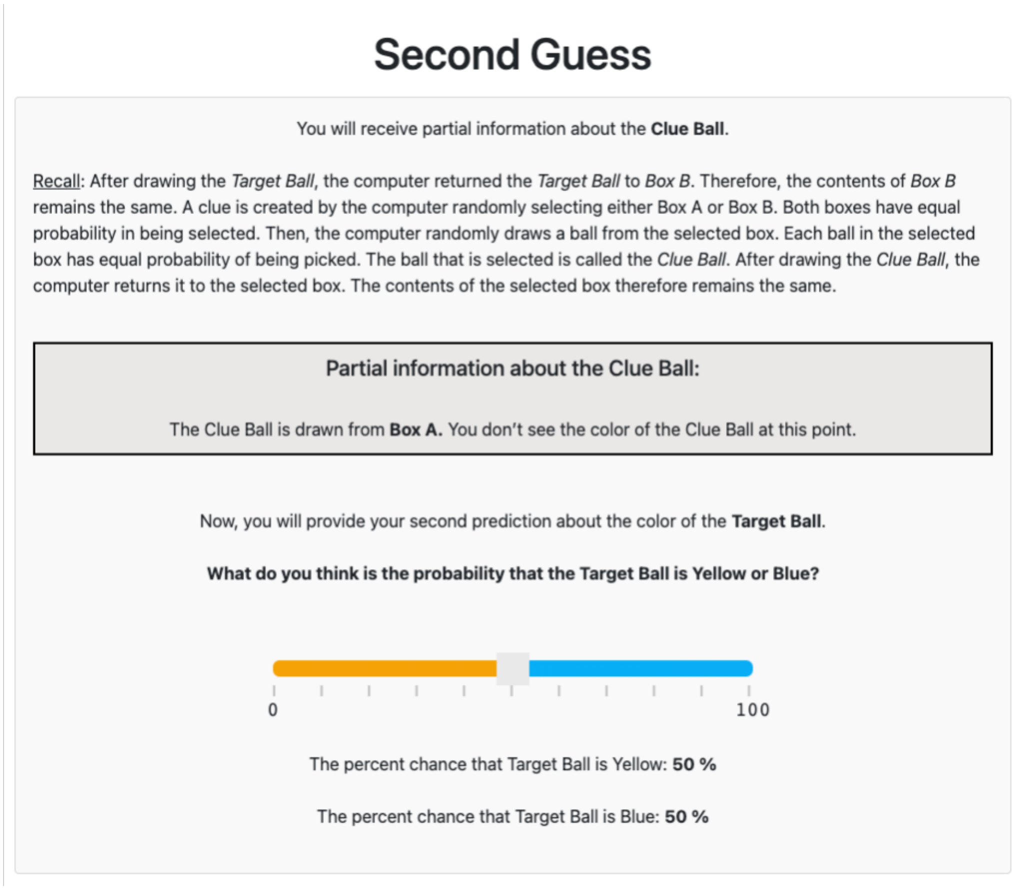


Figure 25: Screenshot of the interim belief elicitation

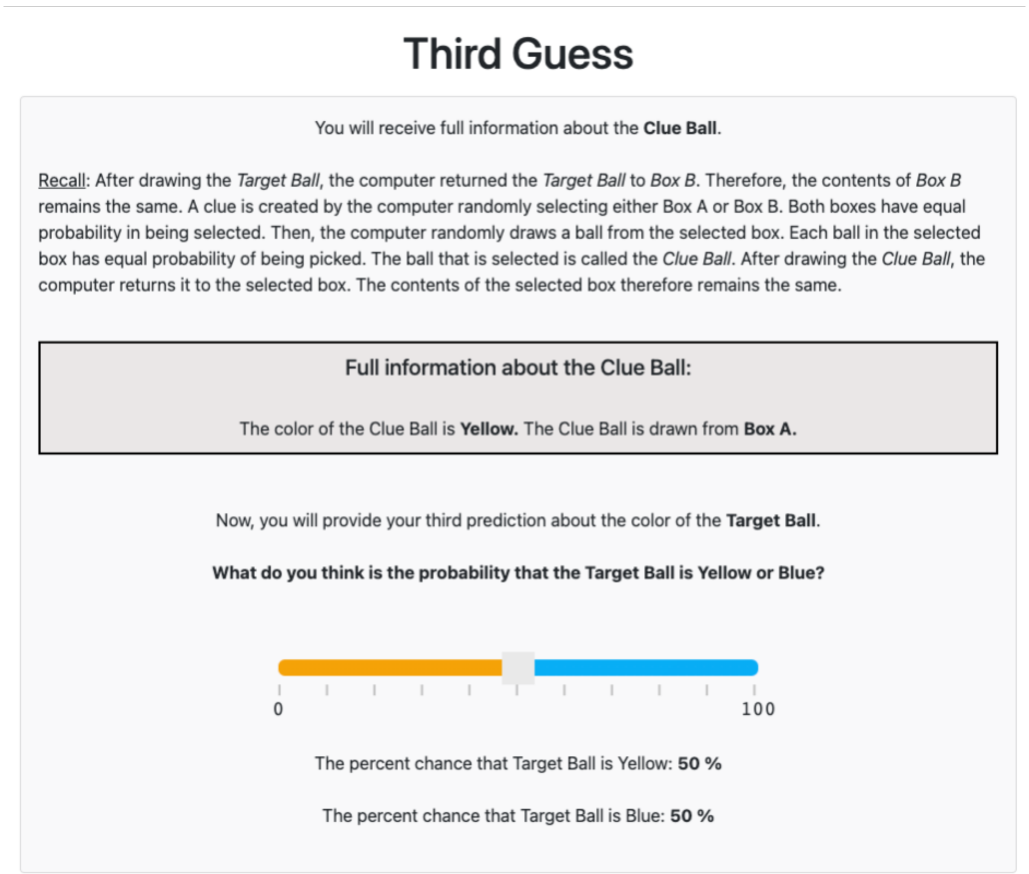


Figure 26: Screenshot of the posterior belief elicitation

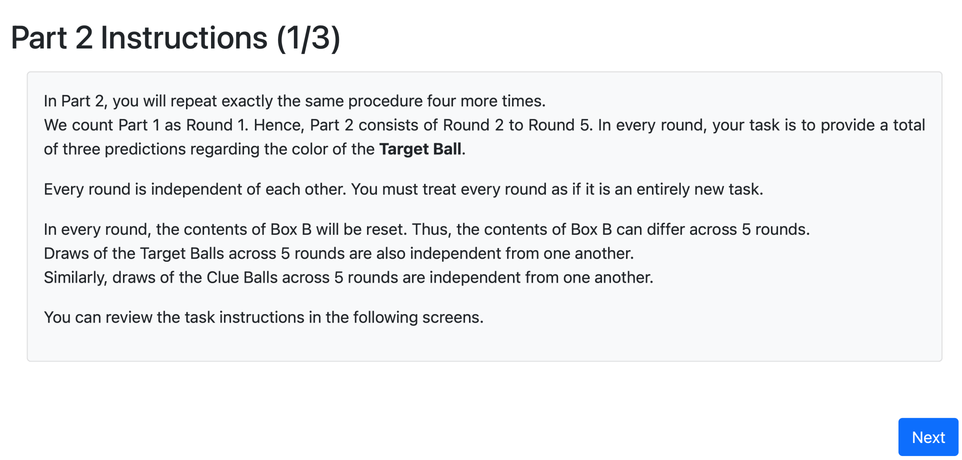


Figure 27: Instructions for Part 2