Learning Gaps? Attitudes and Beliefs Under Described and Experienced Signals

David Gonzalez-Jimenez[a]

[a] *ERP, SEPS, University of St. Gallen* *

March 25, 2026

Abstract

We show that ambiguity attitudes vary systematically with the informativeness of available signals, using an experimental design that varies informativeness and a method that separately measures attitudes and beliefs. When signal-event associations are uncertain and must be inferred from sampling (experience), aversion decreases as signals become more informative, and insensitivity drops sharply under the strongest signal. By contrast, when associations are fully known (description), aversion is essentially unchanged across informativeness levels, while insensitivity shows at most a modest response. This contrast suggests the informativeness effect reflects responses to ambiguity about signal-event associations. Separately measuring attitudes and beliefs proves consequential: the two can move in different directions, and failure to separate them overstates conservatism in updating. For beliefs, subjects in description underweight signal information relative to the Bayesian benchmark, while subjects in experience cannot be distinguished from it under our baseline prior specification, though this finding is sensitive to the assumed prior.

Keywords: Ambiguity Attitude; Signal Informativeness; Decision from Experience; Description-Experience Gap; Belief Updating; Learning.

JEL codes: C91, D81, D83.

---

*Rosenbergstrasse 22, CH-9000 St.Gallen, Switzerland. Email: davidricardo.gonzalezjimenez@unisg.ch. I thank Peter Wakker, Aurelien Baillon and Chen Li for their support and guidance in this project, and Mohammed Abdellaoui for his hospitality during a visit to the department of Economics and Decision Sciences at HEC where an early version of this project benefited from his insightful remarks. Preregistered in OSF: https://doi.org/10.17605/OSF.IO/MUZFH

1 Introduction

The study of decision-making under ambiguity has produced a rich understanding of how individuals respond to unknown probabilities. Since Ellsberg (1961), a large body of work has documented that individuals exhibit systematic attitudes toward ambiguity—patterns in how they evaluate uncertain prospects that go beyond what subjective expected utility would predict (Schmeidler, 1989; Gilboa and Schmeidler, 1989; Klibanoff et al., 2005). These attitudes are predominantly studied in settings where a decision-maker faces an unknown distribution over payoff-relevant events and where the informational environment is held fixed. A notable departure from this view is source preference theory, which has shown that attitudes vary across different sources of uncertainty, reflecting differences in familiarity, competence, or the nature of the domain (Tversky and Fox, 1995; Abdellaoui et al., 2011; Wakker, 2010). This body of work established that the context in which ambiguity is encountered matters for how it is evaluated. Yet even with this insight, a fundamental question remains open: does the informativeness of available signals modulate ambiguity attitudes? This question arises naturally in most economically relevant situations, where individuals have access to signals—observable cues that are informative about, but do not directly determine, the events that affect payoffs (Milgrom, 1981). The informativeness of these signals varies: a diagnostic test may be highly predictive of a disease, while a set of vague symptoms may carry little information. Despite the centrality of signals to economic decision-making, almost nothing is known about how the informativeness of available signals shapes ambiguity attitudes.

This gap in the literature is consequential. If ambiguity attitudes respond to the informational environment—in particular, to the strength of the association between signals and payoff-relevant events—then they are shaped by features of the decision problem that a principal or institution can influence. This has implications for information design, mechanism design, and any applied setting where a principal chooses how informative a signal to disclose to an ambiguity-sensitive agent. Understanding this relationship, however, requires overcoming a fundamental measurement challenge: when a decision-maker receives an uncertain signal and makes a choice, her behavior reflects both her beliefs about the payoff-relevant event and her attitudes toward the uncertainty she faces. Without separating the two, one cannot determine whether a change in signal informativeness affects how she *feels* about the uncertainty, how she *updates* her beliefs, or both.

Recent work has begun to examine how uncertain signals affect decision-making. Epstein and Halevy (2024) study settings where signals are ambiguous, finding that posterior beliefs frequently violate the martingale property, which they interpret as evidence of signal ambiguity. Shishkin and Ortoleva (2023) document dilation effects when signals are uncertain, and Kops and Pasichnichenko (2023) test for negative value of information under ambiguity. However, a common limitation across these studies is the inability to separate beliefs from attitudes. Epstein and Halevy (2024) explicitly acknowledge that the updating patterns they observe may be distorted by unmeasured ambiguity attitudes. Liang (2019) attempts to address this confound indirectly, deriving theoretical predictions about how certainty equivalents should behave under different models and comparing these to experimental data, but this approach cannot quantify the degree of under- or overreaction to information. The belief-attitude confound is not merely a technical nuisance; it can reverse substantive conclusions. A decision-maker who appears to underreact to a signal may in fact be updating her beliefs correctly, with the apparent underreaction reflecting ambiguity attitudes that distort how those beliefs translate into choice.

This paper resolves the belief-attitude confound by applying the belief hedge method of Baillon et al. (2021) to a setting with uncertain signals of varying informativeness. This method provides independent measures of two components of ambiguity attitudes—aversion, which captures the tendency to avoid uncertain prospects, and insensitivity, which reflects a failure to discriminate between different levels of belief—without requiring knowledge of the decision-maker's beliefs, and it allows the recovery of additive beliefs (a-neutral probabilities) that are corrected for attitudinal confounds (Li et al., 2019). These two components distort choice in distinct ways: ambiguity aversion systematically reduces willingness to bet on uncertain events, while insensitivity compresses distinctions between events with different subjective likelihoods, making them appear more similar. To our knowledge, this is the first study to separately identify ambiguity attitudes and beliefs in a setting where the signal-event association itself is unknown. This joint measurement has a dual payoff. First, it allows us to study how signal informativeness shapes each component of ambiguity attitudes, which is the paper's primary contribution. Second, it produces clean measures of beliefs that can be used to assess the quality of updating, corrected for probability weighting due to compound risk and ambiguity attitudes—for example, insensitivity can make beliefs appear to underweight new information even when they update correctly.

To study these questions, we design an experiment in which subjects face urns containing colored balls marked with letters. The colors (blue, yellow, red) are the payoff-relevant events, each occurring with known and equal probability of 1/3. The letters (A, B) serve as signals, also with known marginal probability of 1/2. What varies across three urns is the joint distribution of colors and letters—that is, the informativeness of the signal. In the no-information urn, letters and colors are independent. In the weak-signal urn, there is a moderate association. In the strong-signal urn, the association is pronounced. By fixing the marginal distributions and varying only the signal-event correlation, we isolate the effect of informativeness on attitudes and beliefs. The experiment employs two treatment arms drawn from the Description-Experience paradigm (Hertwig, 2015): in the Description arm (DfD), subjects are told the exact composition of each urn; in the Experience arm (DfE), they sample with replacement and must infer the signal-event association from their observations. The DfD arm serves as a known-probability benchmark in which the only source of complexity is the compound nature of the risk. The DfE arm introduces genuine ambiguity about the signal-event relationship, making it the primary setting for studying how informativeness shapes ambiguity attitudes.

Our findings reveal three sets of results. First, and most centrally, focusing on the Experience arm where subjects must infer signal-event associations from sampling, we show that ambiguity attitudes respond to the informativeness of available signals, and that the two components of attitudes respond differently. Aversion is clearly lower for the strong signal, with weaker evidence for a decline already at the weak signal. Insensitivity, by contrast, requires a stronger signal to decline—only the strong signal reduces it. These asymmetric responses suggest that the two components of ambiguity attitudes respond differently to the degree of informational content. In the Description arm, where urn compositions are fully known and decision-making occurs under compound risk, we find no change in aversion across informativeness levels, while insensitivity shows a similar but weaker pattern, declining only for the strongest signal. This contrast suggests that the informativeness effects observed in Experience reflect responses to ambiguity about signal-event associations rather than responses to the compound structure alone. To our knowledge, this is the first demonstration that ambiguity attitudes vary systematically with the informational content of signals when those associations must be inferred. The broader ambiguity literature has predominantly studied these attitudes in settings where the information environment is fixed; source preference theory has shown that attitudes differ across sources of uncertainty (Tversky and Fox, 1995; Wakker, 2010). Our findings complement both perspectives by showing that the informativeness of available signals modulates attitudes—and that it does so differently for aversion and insensitivity.

Second, the joint measurement of attitudes and beliefs reveals that the two can move in different directions, confirming that the confound identified in prior studies is empirically consequential. In the Description arm, using raw matching probabilities rather than a-neutral probabilities yields a weight on information of 0.72, compared to 0.81 after correcting for attitudes. Both estimates indicate conservatism relative to the Bayesian benchmark, but the uncorrected estimate is further from it, suggesting that failure to separate attitudes from beliefs may overstate the degree of underreaction. This result validates the methodological approach and demonstrates that studies of belief updating that do not control for attitudes, towards compound risk or uncertainty, may draw misleading conclusions about the quality of inference.

Third, comparing across the Description and Experience arms reveals systematic differences along both dimensions. Because these arms represent fundamentally different learning environments rather than isolated experimental manipulations, we interpret the cross-arm patterns as descriptive characterizations of how attitudes and beliefs differ when signal-event associations are described versus experienced. Regarding attitudes, aversion is lower in Experience across all signal strengths, while insensitivity is higher in Experience for non-informative and weak signals but converges for strong signals. This dual pattern—that the two components move in opposite directions across conditions—would be masked by any measure that conflates them. Regarding beliefs, subjects in Description underweight signal information relative to the Bayesian benchmark, while subjects in Experience assign a weight that cannot be distinguished from it. The pooled comparison is consistent with a learning gap under the baseline prior specification, though this finding is sensitive to the assumed prior.

This paper contributes to several literatures. To the ambiguity literature, we provide the first evidence that ambiguity attitudes respond to signal informativeness when signal-event associations must be inferred from experience, with aversion and insensitivity responding differently to the degree of informativeness. This effect is absent when signal-event associations are fully known. Our findings are consistent with the spirit of source preference theory (Tversky and Fox, 1995; Abdellaoui et al., 2011; Wakker, 2010), which established that attitudes vary across sources of uncertainty and hypothesized that the amount of knowledge or competence a decision-maker has about a source is a key driver of this variation; whether differences in signal informativeness operate within a source or effectively define distinct sources is an open question we return to in the discussion. To the literature on belief updating (Benjamin, 2019; Ambuehl and Li, 2018; Augenblick et al., 2024), we contribute a method that jointly measures attitudes and beliefs. The broader updating literature has addressed the confound between probability weighting under risk and beliefs by employing belief elicitation methods that are robust to such distortions. However, these methods retrieve probabilistic beliefs at the cost of discarding information about attitudes — one cannot measure both simultaneously. The belief hedge method resolves this tradeoff, producing beliefs that are corrected for both probability weighting and attitudes while also delivering independent measures of the latter. To the Description-Experience gap literature (Hertwig, 2015; Wulff et al., 2018; Bohren et al., 2024), we extend the framework from payoff-relevant events to signal learning. Our cross-arm comparison, while subject to the multiple structural differences inherent to this paradigm, reveals that the two components of attitudes move in opposite directions across learning environments, a pattern that would be masked by measures that conflate them.

To illustrate, consider a doctor diagnosing a patient. She observes a persistent cough — a symptom shared by many conditions, from a mild infection to something more serious — so the association between this symptom and any particular disease is far from clear. When she decides whether to prescribe medication, that decision reflects both what she believes about the likelihood of the disease and how she feels about the uncertainty surrounding the signal. Suppose she appears to underreact to the symptom, hesitating to prescribe despite informative evidence. A natural interpretation is that she has not updated her beliefs enough. But it is equally possible that her beliefs have moved appropriately and her hesitation instead reflects aversion to the residual uncertainty — she knows the disease is likely but is reluctant to act on uncertain evidence. These two explanations lead to different conclusions: one is a failure of inference, the other a feature of preferences. Without separating attitudes from beliefs, one cannot tell which is at work, nor whether a stronger signal-event association would change her behavior by improving her inference, by reducing her aversion, or both.

The remainder of the paper is organized as follows. Section 2 describes the experimental design. Section 3 introduces the measures of ambiguity attitudes and beliefs, the hypotheses, and the statistical framework. Section 4 presents the results, and Section 5 discusses implications and directions for future research.

2 Experimental design

2.1 Procedure

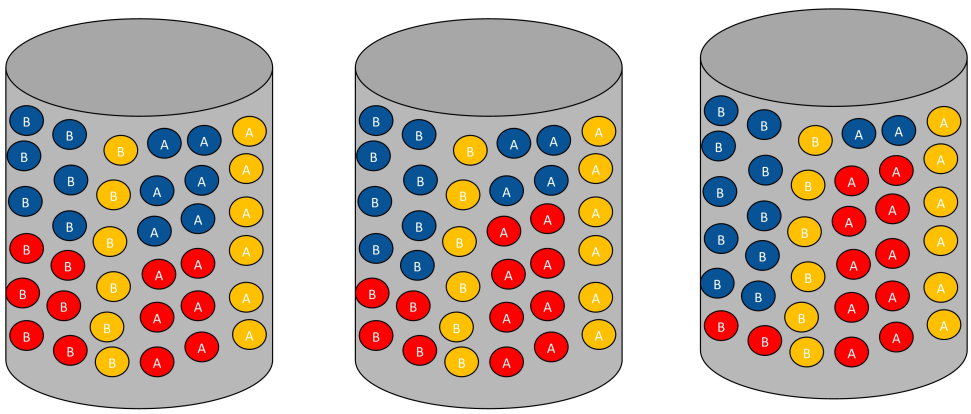
The experiment varies the informativeness of signals across three urns within a between-subjects design with two arms: the Experience arm (DfE) and the Description arm (DfD). Subjects in both treatments faced three urns that differed in the level of informativeness associated with the underlying signal: no (for no information), weak, and strong signal urns. Following Shishkin and Ortoleva (2023); Epstein and Halevy (2024) we fix the prior of PR-events, and in contrast to these studies we fix the prior for the distribution of the signals. By fixing both marginal distributions, we ensure that the only source of variability across urns is the correlation between signals and PR-events. This design choice is central to isolating the effect of signal informativeness on attitudes: any variation in attitudes across urns cannot be attributed to differences in the probability of events or signals themselves.

More concretely, in both conditions subjects were informed about the marginal distribution of color of the balls (PR-events) and Letters written on the balls (signals). The subjects were told that the urns contained balls that were colored blue, yellow and red in equal proportions of 1/3, and marked with a letter, A or B in equal proportions of 1/2. The joint probability of Colors and letters differed by urn, hence varying the degree of informativeness of the signal. The varying signal probabilities conditional on the colors of each urn are shown in Table 1:
In informative cases (the “weak” and “strong” urn), letter A was always a signal favoring red and blue disfavoring, and letter B was the opposite. The letters never signaled anything about yellow.

Subjects were paid a show-up fee of \$5.00 USD and had the opportunity to win a bonus of \$20

	Urn Informativeness		
	No	*Weak*	*Strong*
$P(A\|Blue)$	1/2	1/3	1/6
$P(B\|Blue)$	1/2	2/3	5/6
$P(A\|Yellow)$	1/2	1/2	1/2
$P(B\|Yellow)$	1/2	1/2	1/2
$P(A\|Red)$	1/2	2/3	5/6
$P(B\|Red)$	1/2	1/3	1/6

Table 1: Signal likelihood conditional on color, by urn



(a) No (b) Weak (c) Strong

Figure 1: Urns color and letter distribution

USD. One of their choices was selected at random to determine the bonus payment. The question implemented for real at the end was selected before any choices were made (but kept hidden) to ensure that uncertainty attitudes did not affect the incentive compatibility of the payment.

The procedure was as follows: First, subjects received common information about the three urns they would encounter. This information included the proportions of balls of colors and letters in the urns. Subjects were also informed that each color had at least one ball marked with each letter.

The subjects then received information on the joint distribution of colors and letters in the first urn according to their assigned arm, each arm is explained in 2.2. The computer then performed a random draw from the urn. The color of the drawn ball was kept hidden, but the letter on it was announced. Subjects then made a series of choices between two types of prospects: one prospect paid \$20 if the color of the retrieved ball matched a specified color ($20_{Col}0$), and the other prospect was a lottery with a specified probability p of winning \$20 ($20p0$). To explain the mechanics of the lottery, let us look at a particular decision in which the subject chose between the event and a lottery with a chance of winning of 50 out of 100. Before the decision is made, a random number between 1 and 100 was picked but it was not revealed. If the subject chose the lottery and the selected number was smaller or equal to 50 (that is, the chances of winning), then the subject would receive the bonus of \$20. Probabilities are presented as chances, as it is well known that probabilities are easier to understand if expressed as relative frequencies (Gigerenzer, 1996).

For example, as in Figure 2, a participant might choose between winning the bonus if the ball marked A is red and playing a lottery with a chance of winning 50%. The matching probability is the probability p at which the participant is indifferent between the two options $20_{Col}0 \sim 20p0$. It will depend on the participant's belief that the event would occur and the participant's attitudes towards uncertainty. The matching probabilities between the lottery and the bet on the color were derived using the bisection method, and the matching probabilities are defined as the midpoint between switching points. A further explanation of the bisection method can be found in Appendix A. Six matching probabilities were elicited: three for single-event prospects (e.g., the retrieved ball is red) and three for their complement (e.g., the retrieved ball is yellow or blue). Monotonicity was enforced in the choices: the matching probability of a composite event could not be lower than that of any of its single events. After the six matching probabilities were elicited, subjects continued to the next urn until they completed this process for all three urns. The order of presentation of the urn was randomized.

Figure 2: Example of a matching probability

2.2 Information arms

As mentioned above, the treatment arms differed only in the type of information provided on the contents of the urn. In the DfD arm, subjects were given exact information about how many balls of each color were marked with A and how many were marked with B before the ball that would determine the state of the world was drawn. This condition provides a setting where the signal-event association is fully known, so any attitudes observed reflect the compound nature of the risk rather than ambiguity about the association. Figure 3 gives an example.

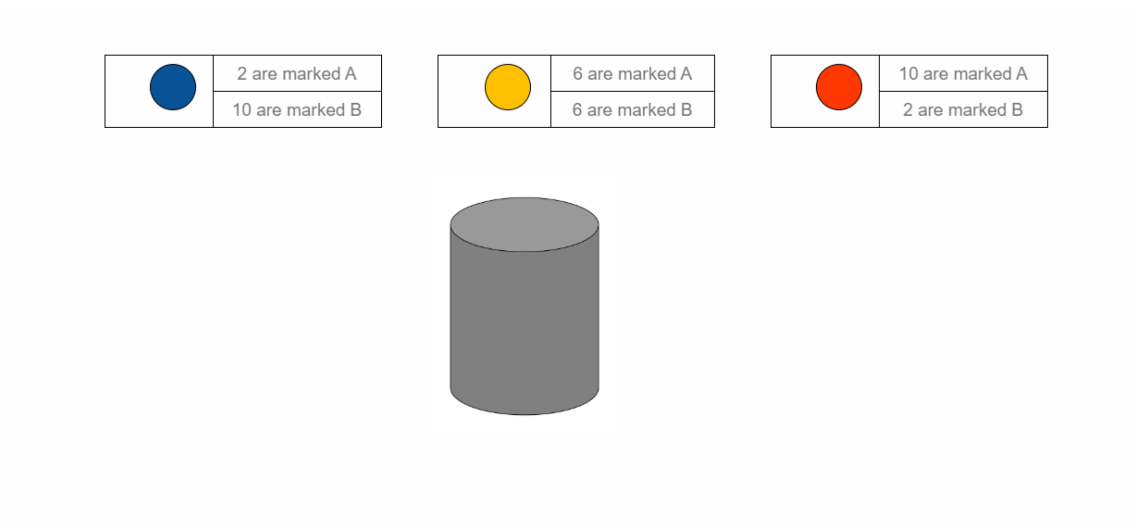


Figure 3: Interface of the information in the urn in the DfD arm

In the DfE arm, subjects were not informed about the urn's exact composition. Instead, they could sample from the urn with replacement as many times as they wished, consistent with the standard sampling paradigm in decisions from experience (Hertwig, 2015; Olschewski et al., 2024). In this condition, the signal-event association must be inferred from observed draws, so subjects face uncertainty about this association. Our analysis focuses on ambiguity attitudes and belief updating; however, for completeness, Appendix F summarizes sampling behavior and presents additional analyses of it. In line with the DfD–DfE literature highlighting the role of memory in experience-based choice (Cubitt et al., 2022), we provided a memory aid that displays the complete sampling history throughout, such a tool is also present when subjects made choices. Figure 4 shows an example of the interface used in the Experience arm, with the colored sequences of the top being the sampling history.

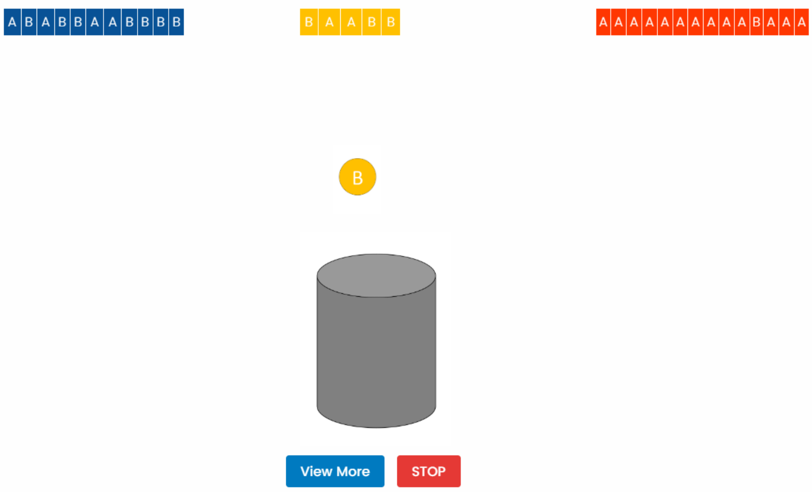


Figure 4: Interface for the Sampling stage of DfE arm

3 Analysis

Eliciting the six matching probabilities allows us to adopt the belief hedge method (Baillon et al., 2021). In particular, we use the specification found in Baillon et al. (2018). This method allows for the measurement of ambiguity attitudes without knowing the subjects' beliefs and it also allows for the derivation of such beliefs. This separation is essential for the questions this paper addresses. To study how signal informativeness shapes ambiguity attitudes, we need to ensure that variation in matching probabilities across urns reflects genuine attitudinal differences and is not confounded by differences in beliefs. Similarly, to study how beliefs react to information, we need measures that are not contaminated by attitudes. For example, in the case of the strong signal, if the letter is A, individuals should change their beliefs toward the ball being Red with a higher probability, while decreasing the probability of Blue. At the same time, the ease with which subjects can associate the letter with a color in such a high-information signal should reduce the uncertainty they perceive. Such an effect should be more noticeable especially in the Experience arm where distributions are not known, thus reducing ambiguity attitudes. On the other hand, in a low or no information situation, individuals are found to overreact to signals, changing their beliefs when they should not (Ba et al., 2024; Augenblick et al., 2024), and their attitudes are more pronounced. Disentangling attitudes and belief effects allows us to properly estimate the reaction of both of these behavioral constructs to informational content.

3.1 Ambiguity attitudes

Individuals in both the Experience and Description conditions face three urns $u \in \{N, W, S\}$, where N stands for the no information urn, W for the urn with the weak signal and S for the urn with the strong signal. For each urn, a ball is drawn, and then the letter on the ball σ is announced, $\sigma \in \{A, B\}$. The event of the ball being of a particular color is represented by $Col = k$ for simple events and $Col = k \cup l$ for composite events, where $k \neq l$ and $k, l \in \{Red, Yellow, Blue\}$. For simplicity, we write $m_{u,k}$ instead of $m(k|\sigma_u)$, as all matching probabilities elicited are conditioned on the same letter for a given urn.

If individuals were Subjective Expected Utility Maximizers, then the matching probability would just give us their belief about the event $m_u(Col) = P(Col|\sigma_u)$. Such beliefs would be congruent with probabilities, so $m_u(Col) + m_u(Col^c) = 1$. Typically, individuals do not comply with SEU meaning that their elicited beliefs are contaminated by their attitudes. For example, an ambiguity averse individual would reveal $m_u(Col) + m_u(Col^c) < 1$. It was the case that in order to affirm that the subjects' beliefs were contaminated by attitudes, it was necessary to know the uncontaminated beliefs beforehand; this is not necessary now when using Baillon et al.'s method.

Let $\overline{m}_{S_u} = (m_{u,Blue} + m_{u,Red} + m_{u,Yellow})/3$ be the average matching probability for all single events and let all composite event matching probabilities average be $\overline{m}_{C_u} = (m_{u,Blue\cup Yellow} + m_{u,Yellow\cup Red} + m_{u,Blue\cup Red})/3$. The proposed index of ambiguity aversion is as follows

$$b_u = 1 - (\overline{m}_{S_u} + \overline{m}_{C_u})$$

For Subjective Expected Utility maximizers we have $\overline{m}_{S_u} = 1/3$ and $\overline{m}_{C_u} = 2/3$ and thus $b = 0$ the neutrality point. Typically, the index is in $[-1, 1]$. Under ambiguity neutrality, the matching probability of an event and its complement sum to 1; under ambiguity aversion, this sum falls below 1. The index b captures the average of this shortfall: when $b > 0$, individuals dislike the uncertain events, preferring even lotteries with low winning probabilities over betting on the event. The opposite, $b < 0$, means a systematic preference for the uncertain events. In the DfD arm, positive values of b reflect avoidance of the complex (i.e. compound) urn's risk in favor of the simpler risk of the lottery, whereas in the DfE arm they reflect avoidance of uncertainty in favor of known risk.

The aversion index captures a motivational aspect, that is, pursuance or avoidance of uncertain events. The next index captures the cognitive side of uncertainty attitudes. In particular, it captures to what extent ambiguity confuses individuals, ultimately making them have an all-or-nothing attitude in which different levels of belief are treated alike. It has been called the a-insensitivity[1] index (Dimmock et al., 2016).

[1]a-insensitivity stands for Ambiguity generated likelihood insensitivity

$$a_u = 3 \times \left(\frac{1}{3} - (\overline{m}_{C_u} - \overline{m}_{S_u}) \right)$$

Expected utility maximizers have $a_u = 0$. Most of the indices will lie in $[0, 1]$. Under ambiguity neutrality, composite events are judged as more likely than single events, and the difference between their matching probabilities is $\overline{m}_{C_u} - \overline{m}_{S_u} = 1/3$. As individuals become confused by ambiguity, this difference shrinks: single and composite events are treated more alike. The index a measures how much this difference falls short of $1/3$. In the extreme case, individuals make no distinction between single and composite events ($\overline{m}_{C_u} = \overline{m}_{S_u}$), yielding $a = 1$.

As this work uses measurements of attitudes that the Description-Experience gap has not used before, it is worth relating them to the parametric forms used in this literature. The most popular was proposed by Goldstein and Einhorn (1987). In this form there are two parameters, an elevation parameter and a curvature parameter. The aversion index used in this work relates to the elevation parameter in probability weighting functions, while the a-insensitivity index is closely related to the curvature parameter. A crucial difference, however, is that in the Goldstein-Einhorn family elevation and curvature are not independent, whereas the indices used here are orthogonal (Baillon et al. (2021), Theorem 12). This independence is important for our analysis because, as we will show, aversion and insensitivity move in opposite directions between Description and Experience.

Within-arm analysis: does informativeness shape attitudes?

The primary question is whether attitudes vary with the informativeness of the signal. If signal informativeness shapes ambiguity attitudes, then the indices should differ across the No, Weak, and Strong urns within the same treatment arm:

$$\overline{b}_N^j \neq \overline{b}_W^j \neq \overline{b}_S^j$$

$$\overline{a}_N^j \neq \overline{a}_W^j \neq \overline{a}_S^j$$

with $j \in \{DfD, DfE\}$. We expect that the indexes decrease as the signal becomes more informative, because a stronger signal-event association should make the event more appealing. Whether the two components respond similarly or differently to informativeness is an open question. Prior work suggests they capture distinct aspects of ambiguity attitudes — insensitivity has been found to respond to cognitive factors such as competence with the source and time pressure (Kilka and Weber, 2001; Baillon et al., 2018), while aversion has not. However, these findings do not directly speak to the effect of signal informativeness, which is novel to this study. The orthogonality of the two indices (Baillon et al., 2021) means that any differential response can be detected, should it exist.

This within-arm comparison is informative in both treatment arms, but for different reasons. In the Experience arm, where the signal-event association is not known, variation across urns would indicate that attitudes respond to the informativeness of the signal as it is inferred from sampling. In the Description arm, where the association is fully known, the same comparison provides a benchmark: if attitudes also vary with informativeness under full information, it would suggest that stronger signal-event associations make individuals more comfortable betting on events regardless of whether uncertainty is present, rather than this being a phenomenon specific to ambiguity.

Between-arm analysis: do attitudes differ across Description and Experience?

Following the literature from the Description-Experience gap that has found that individual decisions differ under both conditions (Wulff et al., 2018; Hertwig, 2015), we also perform a comparison between arms. The general averages of both indexes should differ:

$$\overline{a}^{DfD} \neq \overline{a}^{DfE}$$

$$\overline{b}^{DfD} \neq \overline{b}^{DfE}$$

In addition, we compare each urn with its counterpart in the other arm:

$$\overline{a_u}^{DfD} \neq \overline{a_u}^{DfE}$$

$$\overline{b_u}^{DfD} \neq \overline{b_u}^{DfE}$$

This between-arm comparison serves a dual purpose. It extends the DfD-DfE literature to a setting with signals. It also helps contextualize the within-arm effects: if informativeness shapes attitudes in Experience but not in Description, the effect is specific to conditions where the signal-event association is not fully known.

3.2 Beliefs

The method used in this work allows for measurement of the a-neutral probabilities as well. They are additive probabilities that capture beliefs once corrected for ambiguity attitudes[2] (Li et al.,

[2] Li et al. (2019) characterize these probabilities as the belief that the ambiguity neutral twin of our subjects will have and show the assumptions necessary to obtain them

2019). The a-neutral probabilities conditional on the signal are:

$$\pi(Col = i|\sigma_u) = \frac{m_{u,i\cup l} + m_{u,i\cup k} - m_{u,k} - m_{u,l}}{2(1-a_u)} \tag{1}$$

where $i, k, l \in \{Blue, Yellow, red\}$.

The analysis of the updating of a-neutral probabilities follows methods in the literature that estimate under- or overweighting of probabilities.

3.2.1 Updating Analysis

Given the additive beliefs, it is now possible to analyze how they react to information in both DfD and DfE. In order to estimate the parameter quantifying the weight individuals give to information, we assume the widely applied quasi-Bayesian model (Grether, 1992; Holt and Smith, 2009; Ambuehl and Li, 2018; Coutts, 2019; Möbius et al., 2022; Augenblick et al., 2024) and we use a log-odds specification to derive it. Such specifications are known as the Grether regressions, but in most studies using them the state space is binary, and log-odds completely capture it. Instead, here we calculate the Grether regressions for the log-odds of each color and its complement[3], as seen in equation 2, where we model how subject i uses signal σ from urn u to form the a-neutral belief π_i, with ρ capturing the weight given to new information. When $\rho = 1$ individuals update as a Bayesian would. when $\rho < 1$, individual underreacts to information and when $\rho > 1$ they overreact to it.

$$ln\left(\frac{\pi_i(Col|\sigma_u)}{1-\pi_i(Col|\sigma_u)}\right) = ln\left(\frac{P(Col)}{1-P(Col)}\right) + \rho ln\left(\frac{P_i(\sigma_u|Col)}{P_i(\sigma_u|Col^c)}\right) \tag{2}$$

Given that the unconditional probabilities of the colors are always 1/3, then the log odds of the priors is always $ln(1/2)$. Thus, the constant in this specification will just reflect the log-odds of the priors as seen in Eq. 3. Hence, it is important that individuals understand the composition of the urn. Because of this, a comprehension question[4] was added about the contents of the urn and subjects who did not correctly answer it were removed from the analysis.

$$ln\left(\frac{\pi_i(Col|\sigma_u)}{1-\pi_i(Col|\sigma_u)}\right) = \rho_0 + \rho ln\left(\frac{P_i(\sigma_u|Col)}{P_i(\sigma_u|Col^c)}\right) + \epsilon_i \tag{3}$$

[3]Ba et al. (2024) argue against using the Grether equations for more than three spaces because multiple log-odds ratios are needed to capture all possible combinations of states, thus implicitly assuming that agents distort each pair of prior odds and signal likelihoods in an identical way. We relax this by considering the log-odds of the state and its complement, thereby reducing the number of necessary log-odds combinations to equal the number of states. Moreover, the Dirichlet prior assumed for the DfE arm (Section 3.2.2) has the aggregation property: the marginal distribution of any event versus its complement is a Beta distribution. The log-odds of each color versus its complement thus corresponds to the log-odds of a natural binary partition, rather than an artificial binarization of the three-state space.

[4]The question can be found in appendix B Figure 11

This equation will be estimated for both experience and description independently, in order to estimate if there exists conservatism in the conditions. To explore any possible differences between arms the two arms are pooled together and the following regression is estimated

$$ln\left(\frac{\pi_i(Col|\sigma_u)}{1-\pi_i(Col|\sigma_u)}\right)=\rho_0+\rho_1 DfE_i+\rho_2 ln\left(\frac{P_i(\sigma_u|Col)}{P_i(\sigma_u|Col^c)}\right)+\rho_3 DfE_i * ln\left(\frac{P_i(\sigma_u|Col)}{P_i(\sigma_u|Col^c)}\right)+\epsilon_i \quad (4)$$

where DfE is a binary variable that indicates whether subject i was assigned to the DfE arm. Thus, ρ_1 captures whether baserate log-odds differ between DfD and DfE (i.e., a level difference in beliefs holding the information term fixed). The coefficient ρ_2 captures the average weight subjects place on the information in the Description arm, while ρ_3 captures whether this weight differs in the Experience arm. The main specification in the paper uses a random-effects model, but we also report a fixed effect specification and a clustered standard errors specification in Appendix I and Appendix J. These alternative specifications are particularly relevant for the DfE arm: because the information term is derived from subjects' sampling histories rather than fixed by design, unobserved traits that affect both sampling behavior and belief reports may induce correlation between the individual-specific component and the regressor[5].

3.2.2 Signal Probabilities

Estimating under- or overreaction by means of Eq. 3 and Eq. 4 requires the probabilities that individuals assign to the signal given each state, $P_i(\sigma_u|Col)$. In the case of the description arm, since individuals know the contents of the urn, it is natural that this probability corresponds to the actual proportion of letter σ_u in color Col of the urn, that is $P_i(\sigma_u|Col) = P(\sigma_u|Col)$. However, in the case of the Experience arm, individuals do not observe this information directly, thus $P_i(\sigma_u|Col) \neq P(\sigma_u|Col)$ as using the actual likelihood of the urn distribution can lead to bias in the estimation of reaction to information.

Deriving a likelihood from subjects' actual sampling histories under DfE requires characterizing the inference problem they face. Subjects were informed about all marginal distributions (Section 2.1), so what is unknown is the joint distribution of colors and letters. We model the problem *as if* subjects were inferring the composition of the urn they are sampling from, within the set of compositions consistent with the known marginals. Given the constraint that each color has 12 balls, that 18 balls are marked A and 18 are marked B, and that each color has at least one ball of each letter, there are 91 possible urn compositions.

The prior over the 91 compositions is a discrete distribution, with probability masses proportional to the Dirichlet density $D(\alpha_m)$, $m = \{1, 2, 3, 4, 5, 6\}$, evaluated at each composition's vector

[5]This concern does not arise in DfD, where the information term is pinned down by the urn design and is thus the same for all subjects facing the same urn.

of joint probabilities. All concentration parameters α are set equal. Such symmetry reflects the principle of insufficient reason (Savage, 1954): before sampling, subjects have no basis for believing that one color-letter combination is more prevalent than another. This guarantees that the prior expectation for each joint probability is $\sum_U \eta_U * P(Col_U \cap \sigma_U) = 1/6$. The symmetric Dirichlet thus involves two assumptions: that no color-letter combination is favored a priori (equal α parameters), and a specific degree of prior dispersion (the value of α). We maintain the first throughout, as there is no basis for breaking symmetry. We vary the second to assess robustness, since we cannot observe subjects' actual prior expectations about the urn compositions.

The implications of the value of α are as follows: when $\alpha < 1$, more probability mass is placed on extreme compositions, those with pronounced differences in color-letter proportions, than on moderate ones. When $\alpha > 1$, more weight is placed on compositions close to the symmetric urn with 1/6 in each color-letter combination. The special case $\alpha = 1$ is the uniform distribution, under which all 91 compositions are equally likely a priori.

How subjects learn from sampling depends on the interaction between this prior and the observed data. After each draw, posterior weights are updated via Bayes' rule: the likelihood of the observed color-letter combination under each composition multiplies its prior weight, and the weights are renormalized. Because only compositions consistent with the known marginals receive positive weight, these constraints are maintained throughout. If a subject observes more A-Red balls, the posterior places more mass on compositions with higher A-Red proportions. How quickly the posterior moves away from the prior depends on α: a low α concentrates mass on extreme compositions, so each draw is highly informative and the posterior shifts rapidly; a high α concentrates mass on moderate compositions, so more draws are needed before the posterior departs substantially from the prior. The primary analysis uses $\alpha = 1$ (uniform priors). To assess robustness, we also report results for $\alpha = 1.5$ and $\alpha = 0.15$, which correspond to priors that favor moderate and extreme compositions, respectively.

The signal probability used in the updating analysis, $P_i(\sigma_u|Col)$, is the posterior predictive probability: the weighted average of each composition's conditional probability of the signal given the color, where the weights are the posterior probabilities over compositions given subject i's sampling history.

By deriving $P_i(\sigma_u|Col)$ from each subject's sampling history, this work addresses two related concerns raised in the DfD-DfE literature. Aydogan (2021) showed that assumptions about prior beliefs can affect whether a gap between experience and description is detected. Abdellaoui et al. (2011) argued that comparisons of attitudes and beliefs should be based on the information subjects actually faced, not the underlying distribution of the urn. Both concerns point to the same principle: in DfE, the relevant signal probabilities are those implied by the subject's own observations, not the true urn composition. The Dirichlet model implements this principle — the posterior $P_i(\sigma_u|Col)$ reflects the sampled information and the assumed prior.

4 Results

4.1 Data

We recruited 400 subjects from prolific, comprising 194 males and 197 females[6]. The average bonus won was \$7.8. Individuals were assigned to the Decision from Description or Decision from Experience arms. Table 2 shows the composition of each group by demographic characteristic.

	All	DfD	DfE
Age			
Mean	38.98	38.35	39.61
SD	12.76	13.17	12.32
Education in %			
High school	33.08	32.18	34.17
Higher education (Bachelor)	49.00	48.02	50.25
Post-graduate education	17.66	19.80	15.58
Gender in %			
Female	49.00	48.02	50.25
Male	48.26	50.50	46.23
Non-binary	1.99	0.99	3.02

Table 2: Descriptive statistics for the complete sample

As stated in the pre-registration, the analysis was only performed using those individuals who passed both comprehension checks, both questions can be found in appendix B. The sample is then reduced to 295 subjects of which 142 subjects were assigned to the DfE arm and 153 subjects were assigned to the DfD arm. Table 3 shows the composition of the subjects who passed both comprehension checks and who will be used in the analysis.

When probabilities were recovered, there were five individuals whose probabilities did not lie inside the unit interval, all of them in the DfE arm. Attitudes and a-neutral probabilities were recalculated for these individuals, forcing their beliefs to lie in the unit interval. The analyses without these individuals do not differ qualitatively and are found in the Appendix H. The analyses without probabilities being in the unit interval do not differ in magnitude from the analysis presented here.

[6]The rest of the subjects reported they were non-binary or preferred not to reveal their gender

	All	DfD	DfE
Age			
Mean	37.83	37.54	38.15
SD	12.36	12.89	11.79
Education in %			
High school	34.58	33.33	35.92
Higher education (Bachelor)	48.47	47.06	50.00
Post-graduate education	16.95	19.61	14.08
Gender in %			
Female	50.17	50.98	49.30
Male	47.12	48.37	45.77
Non-binary	2.37	0.65	4.23

Table 3: Descriptive statistics excluding individuals who did not pass the comprehension check

4.2 Attitudes

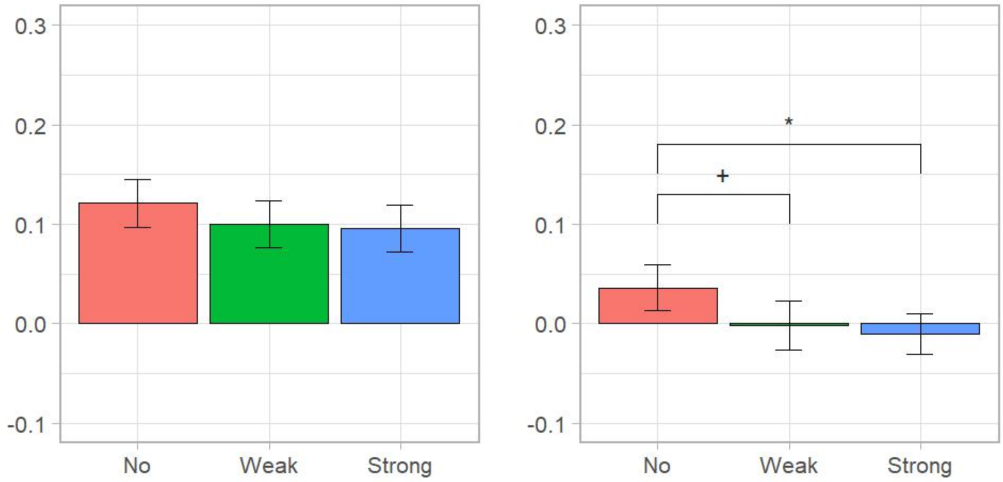
We begin with the primary question: does signal informativeness shape ambiguity attitudes when signal-event associations must be inferred from sampling? We then examine whether the same pattern holds in the Description arm, where associations are fully known, before turning to differences between the two conditions.

Attitudes across informativeness levels within each arm

Within-treatment repeated-measures ANOVA reveals that in the DfE arm, signal strength affects aversion ($F(2, 282) = 3.11, p = 0.05$). Post-hoc analysis shows that aversion is higher for the non-informative signal than for the strong signal ($M_{S-N} = -0.05, t = -2.40, p_{bonf} = 0.03$), and higher than for the weak signal at the 10% level ($M_{W-N} = -0.04, t = -1.89, p_{bonf} = 0.09$). In the DfD arm, by contrast, aversion does not differ across signal strengths ($F(2, 304) = 1.84, p = 0.16$). When the signal-event association is fully known, varying its strength does not change how much individuals avoid the event; as we show next, it affects only how well they discriminate between different levels of likelihood.

Indeed, unlike aversion, there is evidence that insensitivity responds to informativeness in both arms, though the evidence is weaker in DfD. In the DfE arm, signal strength affects insensitivity ($F(2, 282) = 10.58, p < 0.01$): the strong signal produces lower insensitivity than both the no-information signal ($M_{S-N} = -0.11, t = -4.31, p_{bonf} < 0.01$) and the weak signal ($M_{S-W} = -0.10, t = -3.85, p_{bonf} < 0.01$), while the no-information and weak signals do not differ ($p_{bonf} = 0.97$). In the DfD arm, a similar but weaker pattern emerges ($F(2, 304) = 2.83, p = 0.06$): insensitivity in the strong signal is lower than in the weak signal ($M_{S-W} = -0.05, t = -2.30, p_{bonf} = 0.03$). The complete pairwise post-hoc comparisons are in Table 9 and Table 10 in Appendix C.

The effects of varying signal informativeness on aversion and insensitivity differ in two respects. First, the two components respond differently to the degree of informativeness: for aversion, it is sufficient for the signal to carry some information; for insensitivity, only the strongest signal produces a decline. Second, aversion varies across urns only under uncertainty about the signal-event association (DfE), while insensitivity varies in both arms, though the evidence is weaker when the association is known (DfD). The within-arm patterns are evident in Figure 5 and Figure 6.



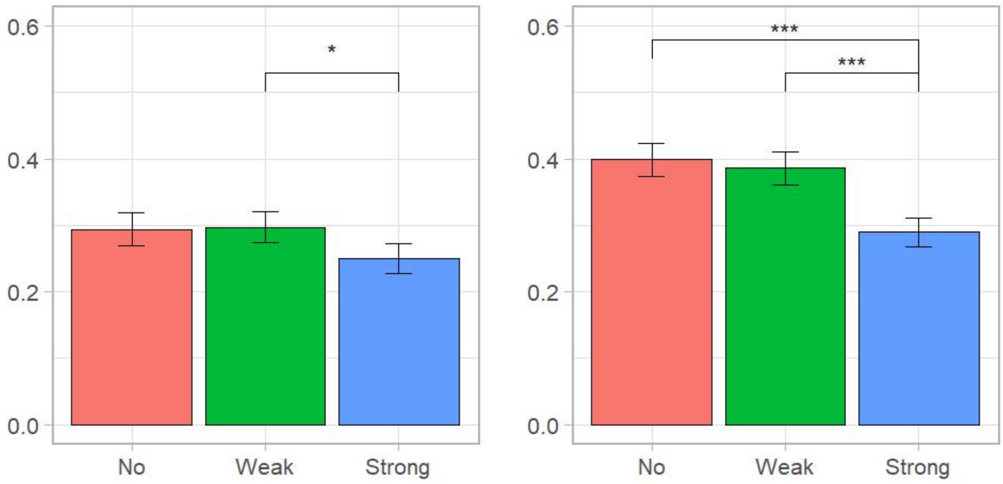
(a) Decision from Description arm

(b) Decision from Experience arm

Figure 5: Mean aversion index by urn. Symbols indicate Bonferroni-adjusted post-hoc pairwise differences in mean aversion across urns $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Do attitudes differ between Description and Experience?

Because these arms represent fundamentally different learning environments rather than isolated experimental manipulations, we interpret these comparisons as descriptive characterizations rather than causal estimates. Table 4 reports summary statistics for the ambiguity attitude indices by treatment arm. Averaged across urns, ambiguity aversion is higher in the DfD arm than in the DfE arm. The same pattern holds urn by urn: for each of the No, Weak, and Strong signal urns, aversion is higher in DfD than in DfE, and in DfE average aversion is close to neutrality across urns. By contrast, a-insensitivity is lower in DfD than in DfE both on average and within the No and Weak signal urns — the opposite direction of aversion — while for the Strong signal urn we find no statistically significant difference between arms. The aversion comparison is consistent with Ert and Trautmann (2014), who document that allowing subjects to sample beforehand increases the relative attractiveness of the ambiguous urn. However, sampling alone cannot explain the patterns



(a) Decision from Description arm (b) Decision from Experience arm

Figure 6: Mean a-insensitivity index by urn. Symbols indicate Bonferroni-adjusted post-hoc pairwise differences in mean a-insensitivity across urns $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

observed in Experience, as aversion varies across urns with different signal informativeness within the DfE arm, indicating that the informational content of the signal, not just the act of sampling, plays a role.

	DfE		DfD			
	Mean	SD	Mean	SD	t	p
Aversion						
Average	0.01	0.23	0.11	0.28	-3.28	0.00
No	0.04	0.27	0.12	0.30	-2.54	0.01
Weak	-0.00	0.29	0.10	0.29	-2.96	0.00
Strong	-0.01	0.24	0.10	0.29	-3.45	0.00
Insensitivity						
Average	0.36	0.22	0.28	0.25	2.85	0.00
No	0.40	0.29	0.29	0.31	2.97	0.00
Weak	0.39	0.30	0.30	0.29	2.60	0.01
Strong	0.29	0.25	0.25	0.29	1.26	0.21

Table 4: Comparison of means between urns in the DfE and DfD arms

Together, these between-arm patterns show that the two components of attitudes move in opposite directions: aversion is lower in DfE while insensitivity is higher. For the strong signal, insensitivity converges across arms (Table 4, $t = 1.26, p = 0.21$), consistent with the within-arm finding that strong signals reduce insensitivity in both conditions.

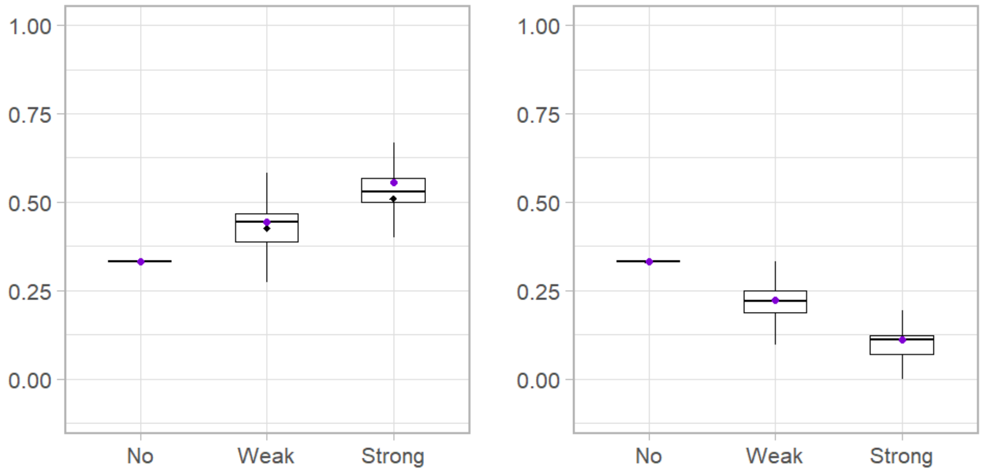
4.3 Beliefs

As described in Section 3.2, a-neutral probabilities were estimated for the individuals in both arms. Before presenting the results using these corrected beliefs, it is useful to illustrate why the correction matters. In the Description arm, running the Grether regression (Eq. 3) using raw matching probabilities as the dependent variable yields a weight on information of $\hat{\rho} = 0.72$ (Table 15 in Appendix G). Using a-neutral probabilities instead yields $\hat{\rho} = 0.81$ (DfD column in Table 5). Both estimates indicate conservatism relative to the Bayesian benchmark ($\rho = 1$), but the uncorrected estimate is further from it, suggesting that failure to separate attitudes from beliefs may overstate the degree of underreaction. The remainder of this section uses a-neutral probabilities throughout.

As seen in Section 2 Table 1 the pairs of conditional probabilities $P(A|Red), P(B|Blue)$ and $P(A|Blue), P(B|Red)$ are equal to each other within each urn, so the Bayesian posterior is the same within each pair. For the first pair, the signal increases the posterior as informativeness grows; for the second, it decreases it. We group them accordingly in Figure 7 for DfD and Figure 8 for DfE.

For individuals in the DfD arm, the main distribution statistics are summarized in Figure 7. The black points in this figure are the measured a-neutral posterior beliefs, as explained in 3.2, and the purple points represent the Bayesian benchmark. As an example, for the no information sample, the Bayesian posterior is 1/3 for both signal groups. For the non-informative signal, both the Bayesian posterior and the measured posteriors juxtapose, meaning that both are very close. For weak and strong signals, the main pattern seems to be that the sample's mean posteriors (in black) are lower than the Bayesian benchmarks (in purple). In fact, for the strong urn more than half of individuals underestimate the probability with respect to the benchmark.

In the context of DfE the comparison with Bayesian benchmarks is not as straightforward. As subjects sample as much as they wish and see different sequences of information, the benchmark is not a single point. Furthermore, the benchmark posterior distribution is contingent on the parameter assumed for the priors. Figure 8 shows a comparison between the distributions of the measured a-neutral posteriors, denoted *Act.*, and the distribution of the Bayesian posterior derived assuming uniform priors ($\alpha = 1$), denoted *U. bench.* It shows that the measured posteriors have a more variable distribution than the uniform benchmark. Figure 8 shows comparable means in both signal groups between benchmark and the actual posteriors. However, when the signal decreases the posterior, the median of the strong signal is lower than the benchmark, indicating more overreaction, whereas when the signal increases the posterior this pattern is less evident. This underlines the importance of the analysis of the under- and overreaction using Eq. 3 to derive inferences about the subjects' behavior. Figure 12 for the comparison of the benchmarks with the

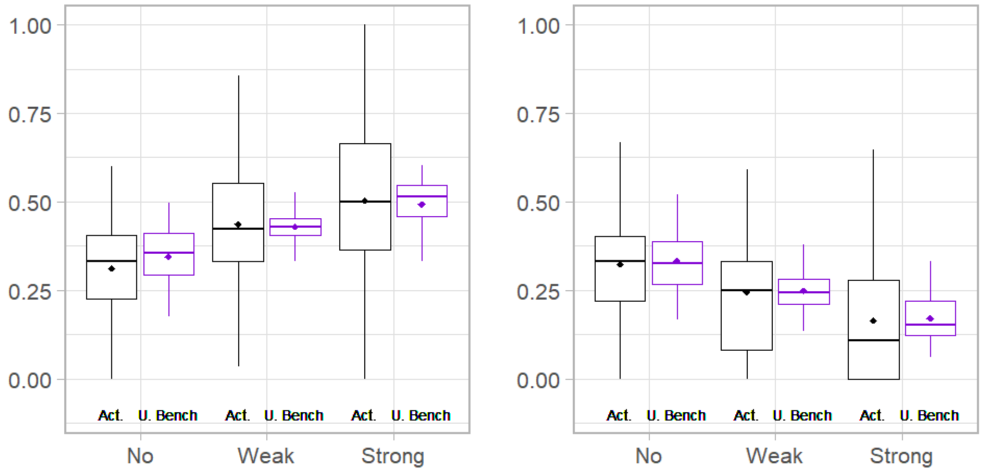


(a) Signal increases posterior

(b) Signal decreases posterior

Figure 7: Comparison actual posterior (Black) vs Bayesian benchmark (Purple) in DfD arm

values of $\alpha = 1.5$ and $\alpha = 0.15$ is available in Appendix D.



(a) Signal increases posterior

(b) Signal decreases posterior

Figure 8: Comparison actual posterior (Black) vs Bayesian benchmark for $\alpha = 1$ (Purple) in DfE arm

4.3.1 Uniform Prior Analysis

Table 5 shows the parameter that captures the weight individuals give to the information from the signal (weight of information) for both arms and for the pooled models. The hypothesis that $\rho = 1$ is rejected for the DfD arm ($t = -8.32, p < 0.01$), but cannot be rejected for DfE ($t = -1.58, p = 0.11$). That is, subjects in DfE assign a weight to information that is not distinguishable from

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	0.81^{***}	0.92^{***}	0.81^{***}	0.81^{***}
	(0.02)	(0.05)	(0.03)	(0.03)
W. Info x DfE			0.11^{*}	0.12^{*}
			(0.05)	(0.05)
DfE			0.02	0.02
			(0.03)	(0.03)
Male				0.01
				(0.03)
Bachelor				-0.00
				(0.03)
Post-Graduate				0.01
				(0.04)
Age				0.01
				(0.01)
Constant	-0.70^{***}	-0.68^{***}	-0.70^{***}	-0.70^{***}
	(0.01)	(0.03)	(0.02)	(0.03)
s_idios	0.52	0.94	0.74	0.74
s_id	0.00	0.00	0.00	0.00
R^2	0.48	0.22	0.31	0.31
Adj. R^2	0.48	0.22	0.31	0.31
Num. obs.	1324	1146	2470	2461

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Stars denote significance vs $\rho = 0$; tests of $\rho = 1$ are reported in the text.

Table 5: Reactiveness to information for the DfD and DfE arms and pooling both conditions

the Bayesian benchmark[7], while subjects in DfD assign significantly less weight than a Bayesian would.

Interestingly, in both the DfD and DfE arms the estimation of the constant is close to the value of the log-odds of the prior $ln(\frac{1/3}{2/3}) \approx -0.69$. In fact, in the DfD arm we have $t = -0.47, p = 0.63$, and in the DfE arm $t = 0.68, p = 0.49$. This experiment was not designed to capture base rate neglect, as prior probabilities were symmetrical and their log-odds did not change through the experiment. However, it gives some evidence of the adequacy of the estimated model[8].

Models (3) and (4) in Table 5 pool both conditions. The key term is $W.info \times DfE$ because it tells us whether people in the Experience arm react differently to the information in the signal than people in the Description arm. The results show clear underreaction in the Description arm. By contrast, participants in the Experience arm react more strongly to the signal, which brings

[7]This finding is likely conservative. Some a-neutral probabilities equal zero, producing undefined log-odds that are excluded from the regression. These zeros are more frequent in DfE than in DfD and are concentrated in the strong signal urn (see Appendix E). Because the dropped observations are those where the signal had its largest effect, their exclusion likely biases $\hat{\rho}$ downward in DfE.

[8]The model explains substantially more variance in DfD ($R^2 = 0.48$) than in DfE ($R^2 = 0.22$). This is expected: in DfD the information term is fixed by the urn composition and is identical for all subjects facing the same urn, so the regressor is noise-free. In DfE the information term is derived from each subject's sampling history and varies across individuals even within the same urn, reducing the model's explanatory power.

their updating closer to the Bayesian benchmark. Under the uniform-prior benchmark, this is consistent with a learning gap. Finally, the coefficient on DfE itself is not statistically significant (Models 3 and 4), suggesting no clear overall shift in beliefs between arms once we account for the information in the signal; the difference mainly shows up in how strongly participants respond to that information.

Because the Experience arm builds on information generated through sampling, it is natural to ask whether the estimated learning gap depends on how we account for individual differences and repeated observations. We therefore complement the main random-effects results with two robustness checks reported in the appendix: a fixed-effects specification (Table 21 of Appendix I) and inference with standard errors clustered at the subject level (Table 23 of Appendix J). Across these checks, the qualitative pattern remains the same: the pooled interaction is positive, indicating stronger responsiveness to the signal in Experience than in Description. The interaction is statistically significant at the 10% level under fixed effects and becomes insignificant under participant-clustered standard errors.

4.3.2 Comparison with different priors

As discussed in Section 3.2.2, the choice of the initial parameter α can matter for our analysis because it changes what is assumed about how informative subjects expected the urn to be before observing any draws. For example, if a subject placed substantial prior weight on very extreme joint distributions (e.g., α close to 0.15) and sampled from the strong urn, signals are highly diagnostic under such priors and their beliefs would move drastically after observing the signal. If we wrongly assume the baseline specification ($\alpha = 1$, implying a more uniform prior), we would be comparing these large belief updates to a more moderate Bayesian benchmark. To bridge this gap, the estimated weight placed on information (ρ) would be stretched upward, tending to overstate responsiveness to signals. Although there is no particular reason to expect such extreme initial beliefs, exploring how the assumed prior affects the estimated responsiveness to information is useful to assess robustness. It is worth emphasizing that the sensitivity to α concerns only the updating analysis in DfE: the measurement of ambiguity attitudes does not require any assumption about priors.

Table 6 shows the estimate of the average weight of information for each assumed concentration parameter α. As expected, if individuals expect ex ante that extreme urns are very probable, then signals are expected to be very informative and the implied information term becomes stronger. With $\alpha = 0.15$, the estimated weight of information is below the Bayesian benchmark ($\rho < 1$; test of $\rho = 1$: $t = -5.33$, $p < 0.01$), implying conservatism in the Experience arm under this prior. In contrast, with $\alpha = 1.5$ the estimate is close to one, thus closer to the Bayesian benchmark.

	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 0.15$
Weight of Info. (ρ)	0.92^{***}	1.02^{***}	0.76^{***}
	(0.05)	(0.06)	(0.04)
Constant	-0.68^{***}	-0.68^{***}	-0.67^{***}
	(0.03)	(0.03)	(0.03)
s_idios	0.94	0.93	0.95
s_id	0.00	0.00	0.00
R^2	0.22	0.22	0.20
Adj. R^2	0.22	0.22	0.20
Num. obs.	1146	1146	1146

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Random Effects estimation

Table 6: Reactiveness to information for different initial parameters α for the Experience arm

Table 7 reports the pooled regressions comparing Experience and Description under the same three priors. The qualitative pattern mirrors the previous results. For $\alpha = 1$ and especially for $\alpha = 1.5$, the interaction term $W.Info \times DfE$ is positive, indicating stronger responsiveness to information in Experience than in Description. When $\alpha = 0.15$, this interaction becomes small and we do not detect a differential impact of Experience on information weighting.

	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 0.15$
Weight of Info. (ρ)	0.81^{***}	0.81^{***}	0.81^{***}
	(0.03)	(0.03)	(0.03)
W. Info x DfE	0.11^{*}	0.21^{***}	-0.05
	(0.05)	(0.06)	(0.05)
DfE	0.02	0.02	0.03
	(0.03)	(0.03)	(0.03)
Constant	-0.70^{***}	-0.70^{***}	-0.70^{***}
	(0.02)	(0.02)	(0.02)
s_idios	0.74	0.74	0.75
s_id	0.00	0.00	0.00
R^2	0.31	0.32	0.30
Adj. R^2	0.31	0.31	0.30
Num. obs.	2470	2470	2470

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Random Effects estimation

Table 7: Pooled models of reactiveness to information for different parameters α for the Experience arm

As in the previous section, similar analyses with alternative model specifications are reported in Appendix I and J. The patterns are very similar: results for $\alpha = 1.5$ do not change, while for $\alpha = 0.15$ the pooled interaction remains close to zero and, with participant-clustered standard errors, the coefficient on DfE becomes significant.

4.4 Summary of results

Table 8 summarizes the main findings.

Table 8: Summary of results

Panel A: Attitudes across informativeness levels within each arm

	Aversion		Insensitivity	
Comparison	DfD	DfE	DfD	DfE
No vs Weak	n.s.	+	n.s.	n.s.
No vs Strong	n.s.	*	n.s.	* * *
Weak vs Strong	n.s.	n.s.	*	* * *
ANOVA	n.s.	*	+	**

Panel B: Do attitudes differ between arms?

Urn	Aversion[a]	Insensitivity[b]
No	*	**
Weak	**	**
Strong	**	n.s.
Average	* * *	* * *

Panel C: Weight of information (ρ)

Prior	DfD	DfE	Difference[c]
$\alpha = 1$	0.81	0.92	0.11*
$\alpha = 1.5$	—	1.02	0.21***
$\alpha = 0.15$	—	0.76	n.s.

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Panel A: Bonferroni-corrected post-hoc pairwise comparisons. Panel B: two-sample t-tests.
[a] Aversion is higher in DfD than in DfE for all significant comparisons.
[b] Insensitivity is higher in DfE than in DfD for all significant comparisons.
[c] Coefficient on the interaction $W.Info \times DfE$ from the pooled regression (Table 7). Stars denote significance vs zero; tests of $\rho = 1$ are reported in Section 4.3. DfD is invariant to α.

The central finding, from the Experience arm where subjects infer signal-event associations from sampling, is that ambiguity attitudes vary with the informativeness of available signals, and that the two components respond differently. Aversion decreases with informativeness in DfE but not in DfD. Insensitivity decreases only for the strongest signal in both arms, though the evidence is weaker in DfD. Between arms, aversion is consistently higher in DfD while insensitivity is consistently higher in DfE. For beliefs, subjects in DfE cannot be distinguished from Bayesian updating under priors that do not place substantial weight on extreme compositions, while subjects in DfD consistently underreact to information.

5 Discussion

The central finding of this paper is that ambiguity attitudes vary systematically with the informativeness of available signals. In the Experience arm, where subjects must infer the signal-event association from sampling, both aversion and insensitivity decline as signals become more informative — but they do so differently. Aversion is clearly lower for the strong signal, with weaker evidence for a decline already at the weak signal. Insensitivity requires a stronger signal before it declines. In the Description arm, where the association is fully known, aversion does not respond to informativeness at all, while insensitivity shows a similar but weaker pattern, declining only for the strongest signal. These findings show that ambiguity attitudes respond to the informational content of signals, extending the literature by separately identifying attitudes and beliefs in a setting with uncertain signals.

Our findings are consistent with the spirit of source preference theory (Tversky and Fox, 1995; Abdellaoui et al., 2011; Wakker, 2010), which established that attitudes vary across sources of uncertainty and hypothesized that the amount of knowledge or competence a decision-maker has about a source is a key driver of this variation. Within the Experience arm, varying informativeness changes both components of attitudes, which is reminiscent of how attitudes differ across sources that vary in familiarity or competence (Kilka and Weber, 2001). Whether differences in signal informativeness operate within a single source of uncertainty or effectively define distinct sources is an open question. In our design, the three urns share the same marginal distributions and differ only in the strength of the signal-event association; one could view them as a single source with varying informativeness or as three sources defined by their informational content. Resolving this distinction is beyond the scope of this study.

A natural concern with our DfE design is that sampling is endogenous: subjects choose when to stop, so the information they carry into their decision is self-selected. Abdellaoui et al. (2024) address this concern by exogenously fixing the number of draws before choice and find that learning moves ambiguity attitudes toward neutrality — a result consistent with the within-arm patterns we observe. However, when the question is how attitudes respond to different levels of signal informativeness, exogenous sampling may introduce its own problems: if subjects are forced to stop before they can distinguish a weak association from a strong one, their attitudes may reflect the imposed constraint rather than their response to the informativeness of the signal; if they are given enough draws, the association may become effectively known, eliminating the ambiguity the design is meant to study. Free sampling means that any residual ambiguity at the point of decision reflects the amount of information subjects voluntarily gathered, rather than a constraint imposed by the experimenter[9].

[9]In our data, the correlation between the number of samples drawn and ambiguity attitudes is weak at best; see

There is also a less obvious reason why exogenously varying informativeness while holding residual ambiguity fixed may not be as straightforward as it appears in signal settings. Unlike classical ambiguity about the state distribution, ambiguity about the signal-event association can only be reduced by observing joint realizations of signals and events. In our design, both marginal distributions are known and the only unknown is the joint distribution of signals and events — so any information about the joint could in principle resolve both informativeness and residual ambiguity simultaneously. Whether this entanglement actually confounds identification is an open question, but it is one of the reasons we chose free sampling, where the residual ambiguity at the point of decision is a natural consequence of subjects' own information gathering. Now that we have established that signal informativeness shapes ambiguity attitudes, an interesting next step would be to explore how the mode of information acquisition — active sampling versus passively receiving observations — and the quantity of information interact with this relationship.

The attitudes observed in the Description arm deserve separate consideration. In DfD, subjects know the exact composition of each urn, so there is no ambiguity about the signal-event association. The positive aversion observed in DfD therefore reflects avoidance of compound risk — the complexity of evaluating a conditional probability structure — rather than ambiguity. This is consistent with findings that attitudes toward compound risks differ from those toward simple risks (Armantier and Treich, 2016; Halevy, 2007) and from those toward ambiguity (Abdellaoui et al., 2015). Moreover, the comparison of attitudes between DfD and DfE underscores the need to account for this distinction: studies of belief updating that use matching probabilities and assume that attitudes toward the compound signal-event structure are equivalent to those toward the simple lottery — as in Ambuehl and Li (2018) — may mischaracterize the quality of updating, since the compound risk literature shows precisely that these attitudes differ.

The type of uncertainty subjects face in this study also deserves clarification. Before making their choices, subjects know which letter is on the sampled ball, which means some uncertainty has already been resolved. This is a key difference from Epstein and Halevy (2024), where no signal has been drawn and the elicitation is done for both signals. We therefore do not address attitudes toward signal uncertainty directly, as their definition of attitudes toward signal ambiguity requires the violation of the martingale property of Bayesian updating, and thus the elicitation of conditional probabilities for both signals. In this experiment, since there is no uncertainty about the color composition of the urn, what introduces uncertainty is the unknown relationship between letters and colors — a setting with parallels to those explored by Kops and Pasichnichenko (2023) and Liang (2019). What we capture is the posterior ambiguity attitude after the realization of an uncertain signal whose conditional distribution on color is unknown.

Appendix F for details.

The between-arm comparison reveals that aversion is higher in DfD while insensitivity is higher in DfE — opposite directions for the two components. Detecting this pattern requires indices that are independent of each other. In the Goldstein-Einhorn probability weighting family, the elevation and curvature parameters are not independent, so opposite movements in the two components could partially offset each other and go undetected. The orthogonality of the indices used here (Baillon et al. (2021), Theorem 12) ensures that each component is measured without contamination from the other. This matters practically: a single-parameter measure that blends aversion and insensitivity would show a smaller or ambiguous difference between Description and Experience, potentially leading to the conclusion that the two conditions produce similar attitudes.

Turning to beliefs, in the Description arm subjects exhibit underreaction in their updating of the probabilities of payoff-relevant events, consistent with findings in the literature with binary state spaces (Holt and Smith, 2009; Ambuehl and Li, 2018; Coutts, 2019). This finding contrasts with Ba et al. (2024), who find overreaction when there are more than two states. Several differences between the studies may account for this: they compare over- and underreaction to an expected value, while our states are not collapsible to a single number, and we assume a log-odds functional form, which they avoid.

In the Experience arm, subjects assign a weight to information that is not distinguishable from the Bayesian benchmark under priors that do not place substantial weight on extreme compositions ($\alpha \geq 1$). This finding is consistent with Moreno and Rosokha (2016), who observe that new information is given greater weight under ambiguity, although in their study the unknown element was the distribution of states rather than the signal-event association. The result depends on the assumed prior: under $\alpha = 0.15$, which concentrates mass on extreme compositions, the learning gap disappears. The priors under which the gap is detected ($\alpha \geq 1$) are arguably more plausible: Aydogan et al. (2024) estimated prior parameters in a two-state model and found them to be approximately uniform, with individuals assigning greater weight to central values than to extreme ones. However, in our three-state setting, completely discarding extreme distributions is not possible, so the sensitivity to the prior remains a limitation. This study follows Aydogan (2021) in showing the importance of priors for detecting gaps in the DfD-DfE literature; however, since we use the belief hedge method, priors are not needed to measure attitudes — only to determine whether there is a learning gap.

A possible solution would be to measure prior beliefs before sampling and estimate parameters for initial priors. However, such a measure would need to be collected before sampling, which would require eliciting conditional behavior for both signals. This would mean comparing contingent elicitations of the prior with contingent elicitations of the posterior, and some literature suggests that contingent thinking itself introduces bias (Aina et al., 2024). Developing methods for prior elicitation that avoid this confound is a direction for future work.

Returning to the doctor example from the introduction, our findings show that the concern raised there is empirically consequential. When the signal-event association must be inferred — as it is for the doctor facing an uncertain symptom — attitudes and beliefs are both at work, and they respond differently to the strength of the association. Aversion declines as the association becomes more informative, while insensitivity requires a stronger association before it drops. A measure that conflates attitudes with beliefs would attribute her hesitation entirely to poor inference, when in fact her beliefs may be sound and the hesitation reflects her response to the residual uncertainty.

6 Conclusion

This paper asked whether the informativeness of available signals shapes ambiguity attitudes. Using the belief hedge method to independently measure attitudes and beliefs, we find that it does — and that the two components of attitudes respond differently. Aversion decreases with informativeness when the signal-event association must be inferred from sampling, but not when it is fully known. Insensitivity responds only to the strongest signal in both conditions, though the evidence is weaker when the association is known. These findings show that ambiguity attitudes vary systematically with the informativeness of available signals.

The joint measurement of attitudes and beliefs proves essential. Using raw matching probabilities rather than attitude-corrected beliefs overstates the degree of conservatism in updating, confirming that the belief-attitude confound is empirically consequential. Applying the method to the comparison of Description and Experience reveals that the two conditions differ along both dimensions, with opposite signs: aversion is lower in Experience while insensitivity is higher. For beliefs, subjects in Experience assign a weight to information that is indistinguishable from the Bayesian benchmark under plausible priors, while subjects in Description consistently underreact. This learning gap is sensitive to the assumed prior, highlighting the importance of prior elicitation in future work.

Several open questions remain. Whether varying signal informativeness defines distinct sources of uncertainty or operates within a single source is a question our design raises but cannot resolve. How the mode and quantity of information acquisition interact with the effects documented here is another direction for future research. More broadly, our results suggest that any applied setting in which a principal chooses how informative a signal to disclose to an ambiguity-sensitive agent should account for the possibility that the signal itself shapes the agent's attitudes — not only her beliefs.

References

Abdellaoui, M., Hill, B., Kemel, E., and Maafi, H. (2024). Learning under ambiguity: An experimental investigation. *Journal of Economic Theory*, 216:105797.

Abdellaoui, M., Klibanoff, P., and Placido, L. (2015). Experiments on Compound Risk in Relation to Simple Risk and to Ambiguity. *Management Science*, 61(6):1306–1322. Publisher: INFORMS.

Abdellaoui, M., L'Haridon, O., and Paraschiv, C. (2011). Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications? *Management Science*, 57(10):1879–1895.

Aina, C., Amelio, A., and Brutt, K. (2024). Contingent Belief Updating.

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

Armantier, O. and Treich, N. (2016). The Rich Domain of Risk. *Management Science*, 62(7):1954–1969. Publisher: INFORMS.

Augenblick, N., Lazarus, E., and Thaler, M. (2024). Overinference from Weak Signals and Underinference from Strong Signals. arXiv:2109.09871 [econ, q-fin].

Aydogan, I. (2021). Prior Beliefs and Ambiguity Attitudes in Decision from Experience. *Management Science*, page mnsc.2020.3841.

Aydogan, I., Baillon, A., Kemel, E., and Li, C. (2024). How much do we learn? Measuring symmetric and asymmetric deviations from Bayesian updating through choices. *Working Paper*.

Ba, C., Bohren, J. A., and Imas, A. (2024). Over- and Underreaction to Information: the Role of Complexity in Belief-Updating. *SSRN Electronic Journal*.

Baillon, A., Bleichrodt, H., Li, C., and Wakker, P. P. (2021). Belief hedges: Measuring ambiguity for all events and all models. *Journal of Economic Theory*, 198:105353.

Baillon, A., Huang, Z., Selim, A., and Wakker, P. P. (2018). Measuring Ambiguity Attitudes for All (Natural) Events. *Econometrica*, 86(5):1839–1858.

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

Bohren, J. A., Hascher, J., Imas, A., Ungeheuer, M., and Weber, M. (2024). A Cognitive Foundation for Perceiving Uncertainty.

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

Cubitt, R., Kopsacheilis, O., and Starmer, C. (2022). An inquiry into the nature and causes of the Description - Experience gap. *Journal of Risk and Uncertainty*, 65(2):105–137.

Dimmock, S. G., Kouwenberg, R., and Wakker, P. P. (2016). Ambiguity attitudes in a large representative sample. *Management Science*, 62(5):1363–1380. Publisher: INFORMS.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

Epstein, L. G. and Halevy, Y. (2024). Hard-to-Interpret Signals. *Journal of the European Economic Association*, 22(1):393–427.

Ert, E. and Trautmann, S. T. (2014). Sampling experience reverses preferences for ambiguity. *Journal of Risk and Uncertainty*, 49(1):31–42.

Gigerenzer, G. (1996). The psychology of good judgment: frequency formats and simple algorithms. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 16(3):273–280.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Goldstein, W. M. and Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, 94(2):236–254. Place: US Publisher: American Psychological Association.

Grether, D. M. (1992). Testing bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization*, 17(1):31–57.

Halevy, Y. (2007). Ellsberg Revisited: An Experimental Study. *Econometrica*, 75(2):503–536.

Hertwig, R. (2015). Decisions from Experience. In *The Wiley Blackwell Handbook of Judgment and Decision Making*, pages 239–267. John Wiley & Sons, Ltd.

Holt, C. A. and Smith, A. M. (2009). An update on Bayesian updating. *Journal of Economic Behavior & Organization*, 69(2):125–134.

Kilka, M. and Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty? *Management Science*, 47(12):1712–1726.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, 73(6):1849–1892.

Kops, C. and Pasichnichenko, I. (2023). Testing negative value of information and ambiguity aversion. *Journal of Economic Theory*, 213:105730.

Li, C., Turmunkh, U., and Wakker, P. P. (2019). Trust as a decision under ambiguity. *Experimental Economics*, 22(1):51–75.

Liang, Y. (2019). Learning from Unknown Information Sources. *SSRN Electronic Journal*.

Milgrom, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380–391. Publisher: [RAND Corporation, Wiley].

Moreno, O. M. and Rosokha, Y. (2016). Learning under compound risk vs. learning under ambiguity – an experiment. *Journal of Risk and Uncertainty*, 53(2):137–162.

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. Publisher: INFORMS.

Olschewski, S., Luckman, A., Mason, A., Ludvig, E. A., and Konstantinidis, E. (2024). The future of decisions from experience: Connecting real-world decision problems to cognitive processes. *Perspectives on Psychological Science*, 19(1):82–102. Place: US Publisher: Sage Publications.

Savage, L. J. (1954). *The foundations of statistics.* The foundations of statistics. John Wiley & Sons, Oxford, England. Pages: xv, 294.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57(3):571–587.

Shishkin, D. and Ortoleva, P. (2023). Ambiguous information and dilation: An experiment. *Journal of Economic Theory*, 208:105610.

Tversky, A. and Fox, C. R. (1995). Weighing Risk and Uncertainty. *Psychological Review*, 102(2):269–283.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity.* Cambridge University Press.

Wulff, D. U., Mergenthaler-Canseco, M., and Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, 144(2):140–176.

A Bisection Method

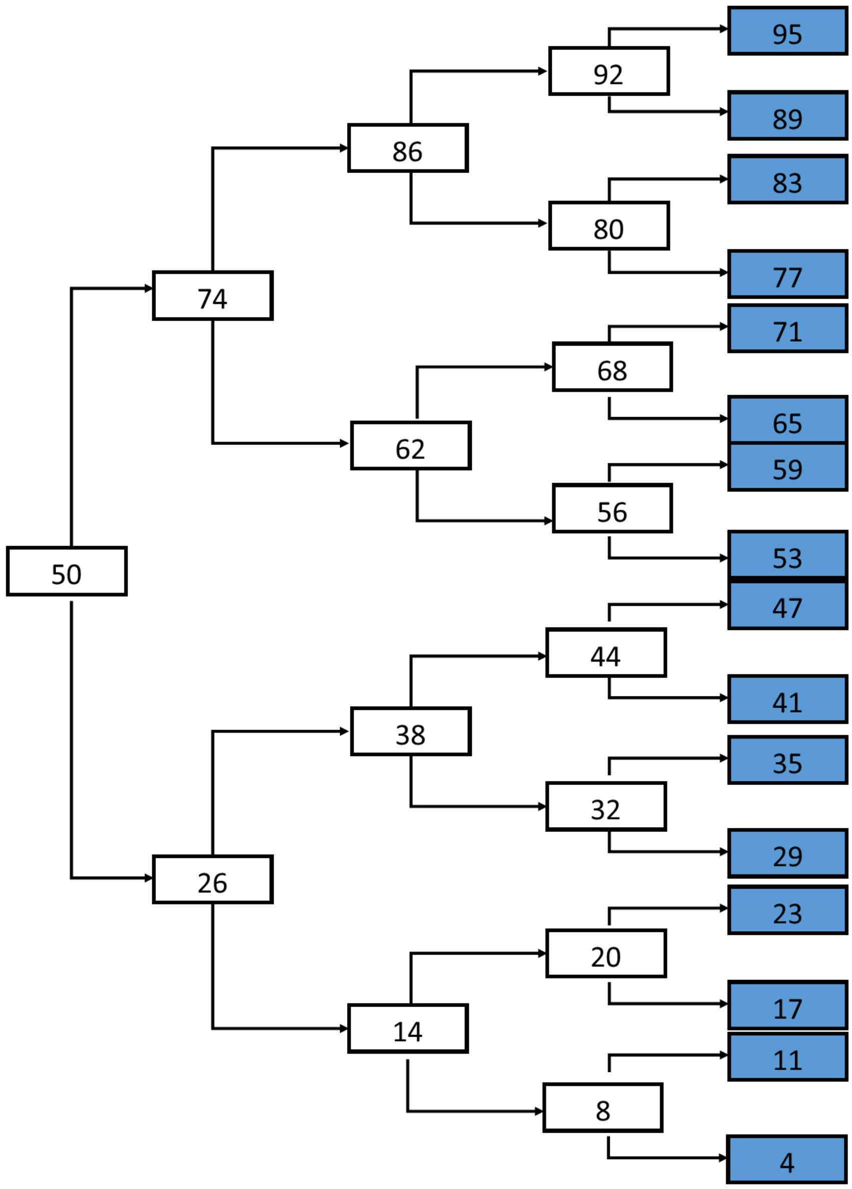


Figure 9: Sequence of lottery choices in the bisection method.

In this experiment, we implement the bisection method in the following way. Individuals first faced a choice, as seen in Figure 2. The choice was between betting on the color or betting on the lottery. Everyone started with a choice of 50% for single-color choices, and how the chances of winning changed in the lottery choice is illustrated in Figure 9. The mechanic was as follows if subjects chose to bet on the color, then in in the sequence they would move up, if they chose the lottery, then they would move down. This sequence would be repeated for 4 choices, the resulting matching probabilities are shown in blue.

By employing the bisection method, it is possible that the subjects did not encounter the question predetermined for payment. If this situation arose at the end of the experiment, the participants would have had the opportunity to make their decision.

B Comprehension Questions

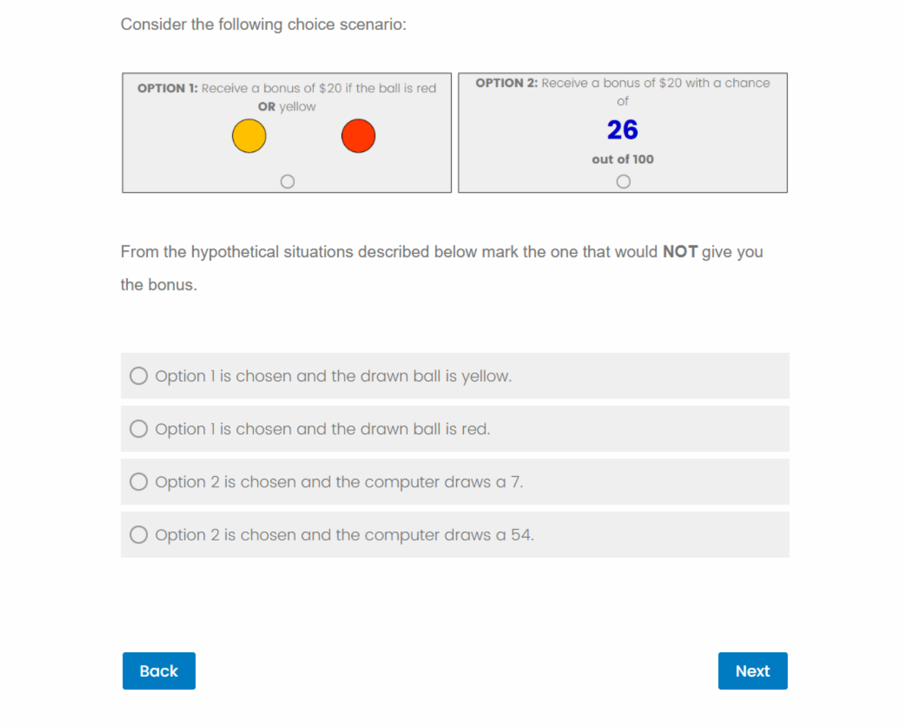


Figure 10: Comprehension question about payment mechanism

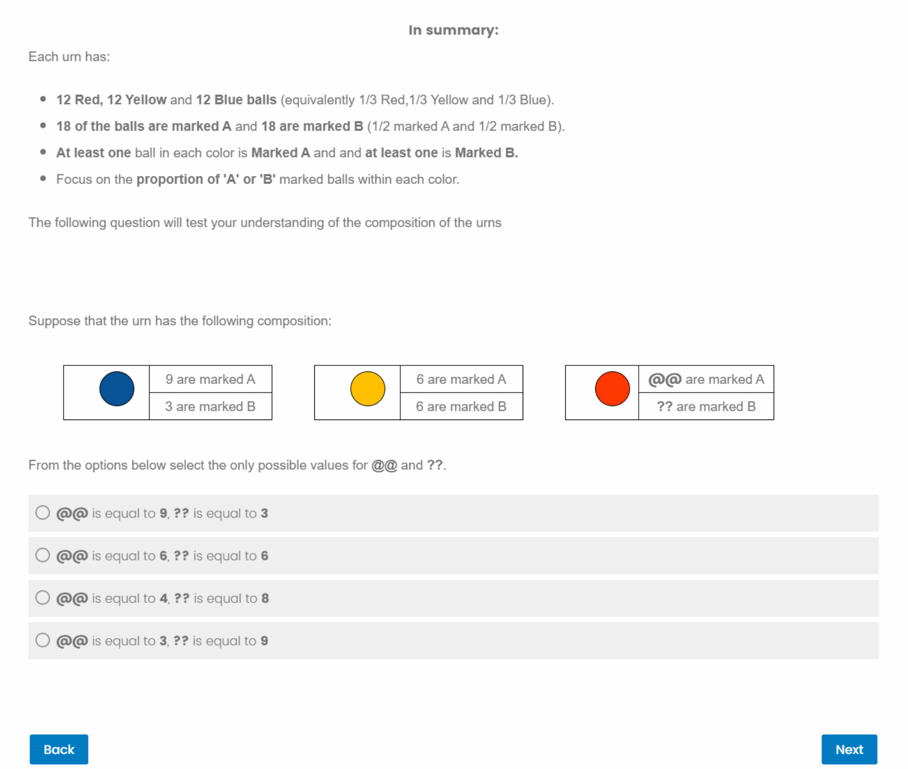


Figure 11: Comprehension question about content of the urn

C Post-Hoc analysis

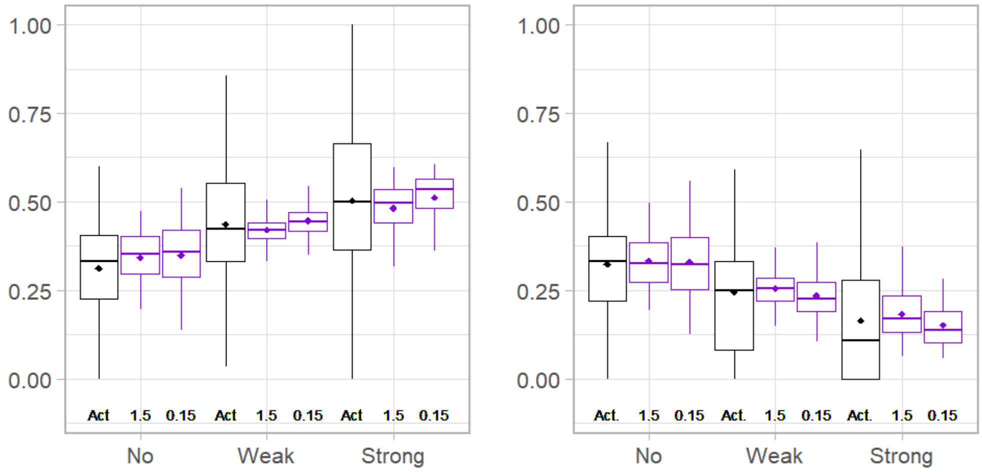
	Difference	t	p	p-corrected
Weak-N.inf	-0.04	-1.89	0.03	0.09
Strong-N.inf	-0.05	-2.40	0.01	0.03
Strong-Weak	-0.01	-0.45	0.33	0.98

Table 9: Mean comparison by signal accuracy in aversion for the experience condition

	Experience				**Description**			
	Difference	t	p	p-corrected	Difference	t	p	p-corrected
Weak-N.inf	-0.01	-0.46	0.32	0.97	0.00	0.15	0.56	1.00
Strong-N.inf	-0.11	-4.31	0.00	0.00	-0.04	-1.80	0.04	0.11
Strong-Weak	-0.10	-3.85	0.00	0.00	-0.05	-2.30	0.01	0.03

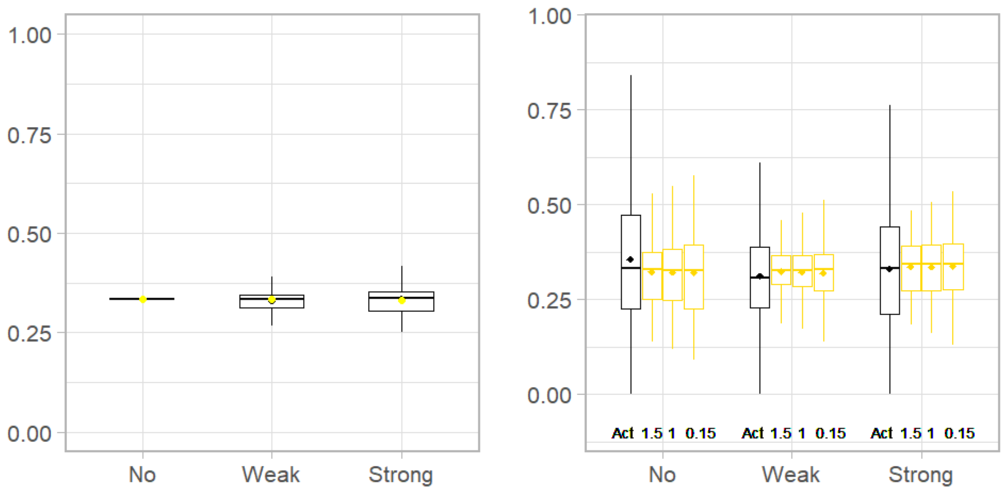
Table 10: Mean comparison by signal accuracy in a-insensitivity

D Beliefs



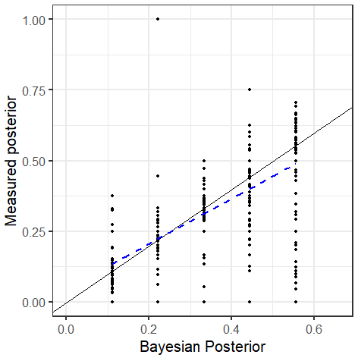
(a) Positive relation signals (b) Negative relation signals

Figure 12: Comparison actual posterior (Black) vs Bayesian benchmarks (Purple) with $\alpha = 0.15$ and $\alpha = 1.5$ in DfE arm

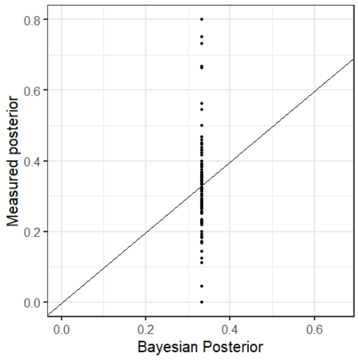


(a) Description (b) Experience

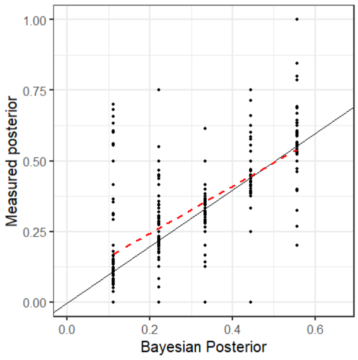
Figure 13: Comparison Actual Posterior and Bayesian Posterior Yellow Signal



(a) Blue

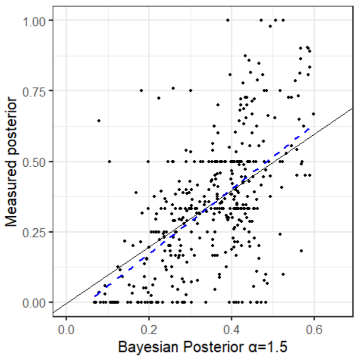


(b) yellow

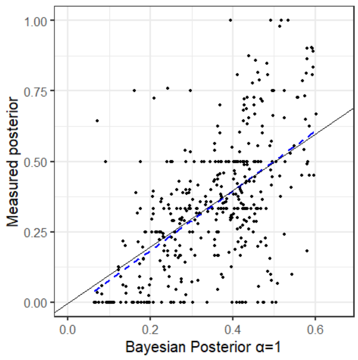


(c) Red

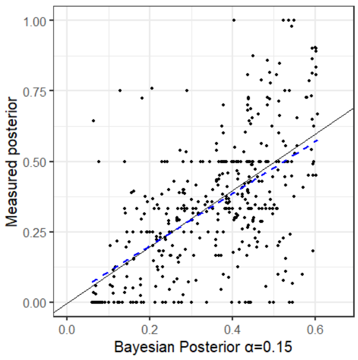
Figure 14: Plot points with different prior parameters Description



(a) $\alpha = 1.5$

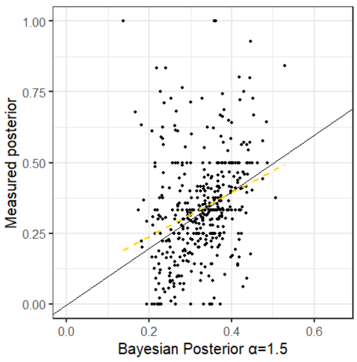


(b) $\alpha = 1$

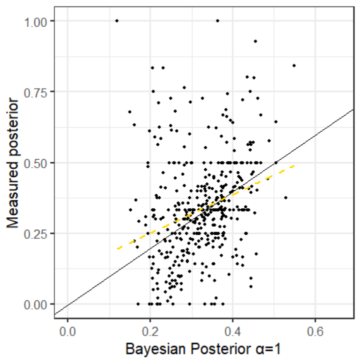


(c) $\alpha = 0.15$

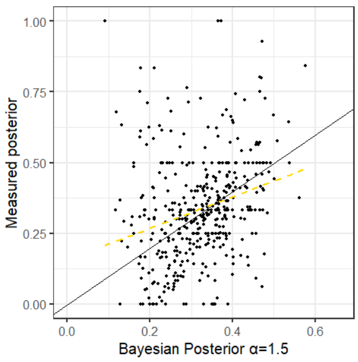
Figure 15: Plot points with different prior parameters color Blue Experience



(a) $\alpha = 1.5$

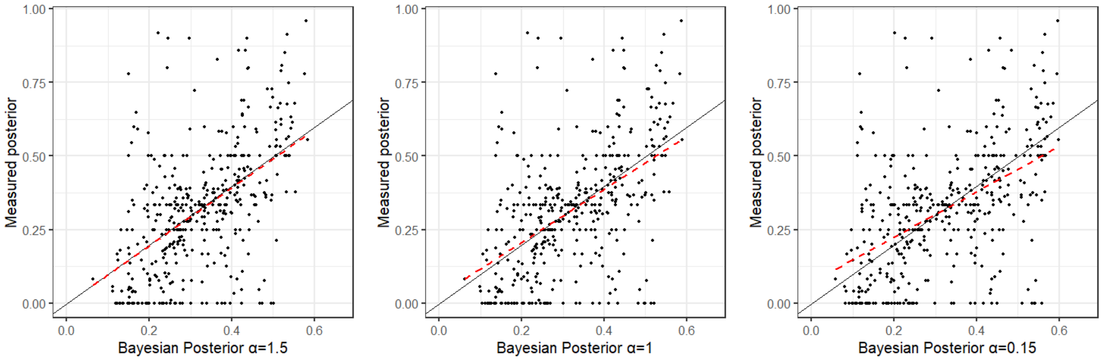


(b) $\alpha = 1$



(c) $\alpha = 0.15$

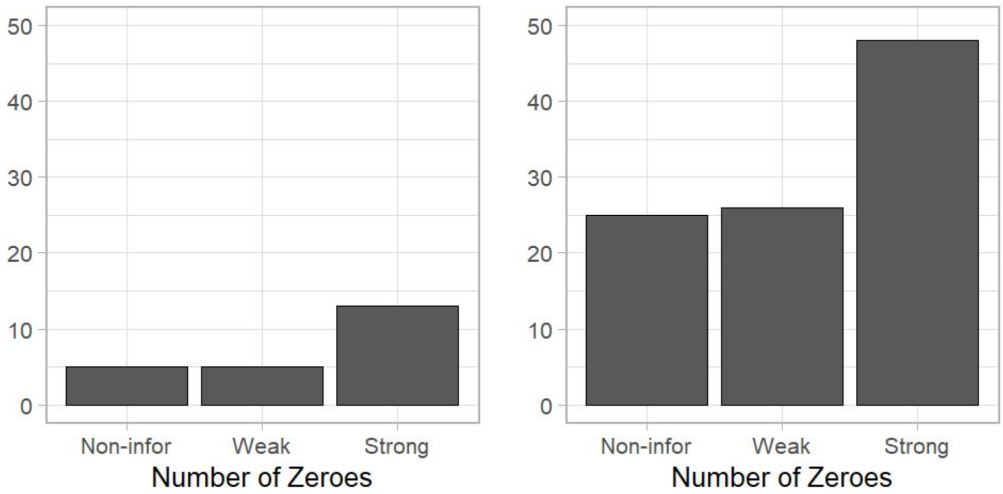
Figure 16: Plot points with different prior parameters color yellow



(a) $\alpha = 1.5$ (b) $\alpha = 1$ (c) $\alpha = 0.15$

Figure 17: Plot points with different prior parameters color Red Experience

E zeros in probabilities and Na's in logodds



(a) Description

(b) Experience

Figure 18: Frequency of subjects with one zero in the probability estimation by Informativeness of the urn

F Sampling

	N	Mean	SD	Median	q25	q75
No	142	31.97	27.83	29.00	11.00	44.00
Weak	142	27.66	24.19	22.00	11.00	38.00
Strong	142	28.08	28.50	21.00	10.25	36.00

Table 11: Summary statistics of sampling by urn

	Cluster	RE	FE
Weak	-4.31^{*}	-4.31^{**}	-4.31^{**}
	(1.71)	(1.63)	(1.63)
Strong	-3.89^{**}	-3.89^{*}	-3.89^{*}
	(1.42)	(1.63)	(1.63)
Constant	31.97^{***}	31.97^{***}	
	(2.34)	(2.26)	
R^2	0.01	0.02	0.03
Num. obs.	426	426	426

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 12: Regression of sampling by urn

	RE	FE	Cluster
N. Samples	0.00	−0.00	0.00
	(0.00)	(0.00)	(0.00)
Constant	0.03		0.01
	(0.03)		(0.04)
s_idios	0.16		
s_id	0.21		
R^2	0.03	0.04	0.01
Num. obs.	426	426	426

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 13: Relationship between ambiguity aversion and Number of samples

	RE	FE	Cluster
N. Samples	-0.00^{+}	-0.00	-0.00^{+}
	(0.00)	(0.00)	(0.00)
Constant	0.39^{***}		0.39^{***}
	(0.02)		(0.03)
s_idios	0.23		
s_id	0.17		
R^2	0.01	0.01	0.01
Num. obs.	426	426	426

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 14: Regression of a-intensitivity against sampling

Weight of Info. (ρ)	0.72^{***}
	(0.02)
Constant	-0.60^{***}
	(0.02)
s_idios	0.46
s_id	0.11
R^2	0.47
Adj. R^2	0.46
Num. obs.	1377

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 15: Weight of information estimated using raw matching probabilities. The dependent variable is the log-odds of the elicited matching probability of the event relative to its elicited complement,$ln(m(Col)/m(Colc))$. DfD arm only. Random effects estimation.

G Analysis of information weight using raw matching probabilities

H Analysis without Individuals with probabilities outside the unit interval

Taking out the individuals whose probabilities lay outside the unit interval, we find that the differences in magnitudes are very small, and the comparisons between arms are still significative in the same way as they were in Section 4.2.

	Experience		Description			
	Mean	SD	Mean	SD	t	p
Aversion						
Average	0.02	0.22	0.11	0.28	-4.54	0.00
Non-informative	0.05	0.27	0.12	0.30	-2.24	0.03
Weak	0.01	0.28	0.10	0.29	-2.53	0.01
Strong	-0.00	0.23	0.10	0.29	-3.14	0.00
Insensitivity						
Average	0.36	0.22	0.28	0.25	4.01	0.00
Non-informative	0.40	0.30	0.29	0.31	3.04	0.00
Weak	0.39	0.30	0.30	0.29	2.55	0.01
Strong	0.29	0.25	0.25	0.29	1.31	0.19

Table 16: Comparison of means between urns in the Experience and Description arms

As there was no subject in the DfD arm whose probabilities were outside of the unit interval, within-treatment comparisons do not change for them. The within-treatment anova for the DfE arm does change in the case of aversion ($F(2, 274) = 2.81, p = 0.06$), while significant at the 5% level before it is only significant at the 10% level. However, for insensitivity, the results do not change as much ($F(2, 274) = 10.06, p < 0.01$). The post hoc comparison of the aversion index show that

	Difference	t	p	p-corrected
Weak-N.inf	-0.03	-1.56	0.06	0.18
Strong-N.inf	-0.05	-2.37	0.01	0.03
Strong-Weak	-0.01	-0.74	0.23	0.69

Table 17: Mean comparison by signal accuracy in aversion for the experience condition

the difference between the strong signal and the noninformative signal is still significant, however there is no difference between the weak and noninformative signal at the 10% level. However, for the a-insensitivity the differences are still significant for the comparison between the strong signal urn, and the non-informative and weak signals.

	Difference	t	p	p-corrected
Weak-N.inf	-0.02	-0.59	0.28	0.83
Strong-N.inf	-0.11	-4.26	0.00	0.00
Strong-Weak	-0.09	-3.64	0.00	0.00

Table 18: Mean comparison on insensitivity by signal accuracy

Excluding these subjects results in a lower estimate of the weight of information in the Experience arm, and the pooled interaction coefficient is no longer significant at the 5% level. Table 20 shows a similar pattern across prior specifications: the interaction coefficient is lower but remains significant for $\alpha = 1.5$

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	0.81^{***}	0.90^{***}	0.81^{***}	0.81^{***}
	(0.02)	(0.05)	(0.03)	(0.03)
W. Info x DfE			0.10^{+}	0.11^{*}
			(0.05)	(0.05)
DfE			0.02	0.02
			(0.03)	(0.03)
Male				0.01
				(0.03)
Bachelor				-0.00
				(0.03)
Post-Graduate				-0.00
				(0.04)
Constant	-0.70^{***}	-0.68^{***}	-0.70^{***}	-0.70^{***}
	(0.01)	(0.03)	(0.02)	(0.03)
s_idios	0.52	0.93	0.74	0.74
s_id	0.00	0.00	0.00	0.00
R^2	0.48	0.21	0.31	0.32
Adj. R^2	0.48	0.21	0.31	0.31
Num. obs.	1324	1119	2443	2434

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Random Effects estimation

Table 19: Reactiveness to information for the DfD and DfE arms and pooling both conditions

	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 0.15$
Weight of Info. (ρ)	0.81^{***}	0.81^{***}	0.81^{***}
	(0.03)	(0.03)	(0.03)
W. Info x DfE	0.10^{+}	0.19^{***}	-0.06
	(0.05)	(0.06)	(0.05)
DfE	0.02	0.02	0.03
	(0.03)	(0.03)	(0.03)
Constant	-0.70^{***}	-0.70^{***}	-0.70^{***}
	(0.02)	(0.02)	(0.02)
s_idios	0.74	0.74	0.74
s_id	0.00	0.00	0.00
R^2	0.31	0.32	0.30
Adj. R^2	0.31	0.31	0.30
Num. obs.	2443	2443	2443

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Random Effects estimation

Table 20: Pooled models of reactiveness to information for different parameters α for the Experience arm

I Fixed effects estimation reaction to information

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	0.81***	0.90***	0.81***	0.81***
	(0.02)	(0.05)	(0.03)	(0.03)
W. Info x DfE			0.10^{+}	0.11^{+}
			(0.06)	(0.06)
Controls	No	No	No	Yes
R^2	0.49	0.21	0.31	0.31
Adj. R^2	0.42	0.10	0.22	0.22
Num. obs.	1324	1146	2470	2461

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$ Fixed Effects estimation

Table 21: Reactiveness to information for the DfD and DfE arms and pooling both conditions

	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 0.15$
Weight of Info. (ρ)	0.81***	0.81***	0.81***
	(0.03)	(0.03)	(0.03)
W. Info x DfE	0.10^{+}	0.20***	−0.06
	(0.06)	(0.06)	(0.05)
R^2	0.31	0.31	0.30
Adj. R^2	0.22	0.22	0.21
Num. obs.	2470	2470	2470

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 22: Pooled models of reactiveness to information for different parameters α for the Experience arm

J Estimation of reaction to information with clustered robust standard errors

	DfD	DfE	Pool	
	Model 1	Model 2	Model 3	Model 4
Weight of Info. (ρ)	0.81^{***}	0.92^{***}	0.81^{***}	0.81^{***}
	(0.04)	(0.09)	(0.04)	(0.04)
W. Info x DfE			0.11	0.12
			(0.10)	(0.10)
DfE			0.02	0.02^{+}
			(0.01)	(0.01)
Male				0.01
				(0.01)
Age				0.01^{+}
				(0.01)
Bachelor				-0.00
				(0.02)
Post-Graduate				0.01
				(0.02)
Constant	-0.70^{***}	-0.68^{***}	-0.70^{***}	-0.70^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.48	0.22	0.31	0.31
Adj. R^2	0.48	0.22	0.31	0.31
Num. obs.	1324	1146	2470	2461

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 23: Reactiveness to information for the DfD and DfE arms and pooling both conditions with clustered standard errors.

	1	1.5	0.15
Weight of Info. (ρ)	0.92^{***}	1.02^{***}	0.76^{***}
	(0.09)	(0.09)	(0.07)
Constant	-0.68^{***}	-0.68^{***}	-0.67^{***}
	(0.01)	(0.01)	(0.01)
R^2	0.22	0.22	0.20
Adj. R^2	0.22	0.22	0.20
Num. obs.	1146	1146	1146

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 24: Reactiveness to information for different initial parameters α for the Experience arm with clustered standard errors

	1	1.5	0.15
Weight of Info. (ρ)	0.81^{***}	0.81^{***}	0.81^{***}
	(0.04)	(0.04)	(0.04)
W. Info x DfE	0.11	0.21^{*}	-0.05
	(0.10)	(0.10)	(0.09)
DfE	0.02	0.02	0.03^{*}
	(0.01)	(0.01)	(0.01)
Constant	-0.70^{***}	-0.70^{***}	-0.70^{***}
	(0.01)	(0.01)	(0.01)
R^2	0.31	0.32	0.30
Adj. R^2	0.31	0.31	0.30
Num. obs.	2470	2470	2470

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 25: Pooled models of reactiveness to information for different parameters α for the Experience arm