

Learning Gaps? Attitudes and Beliefs Under Described and Experienced Signals

DAVID GONZALEZ-JIMENEZ^a

^a*University of St. Gallen* *

October 27, 2026

Abstract

We show that ambiguity attitudes vary systematically with the informativeness of signals. In an experiment varying the strength of signal-event associations, we separate ambiguity attitudes from beliefs within the same choices. When associations must be inferred from sampling (experience), aversion decreases as signals become informative; when associations are fully known (description), aversion shows no response. Insensitivity declines under the strongest signal in both settings, though the evidence is weaker when the association is known. For the analysis of beliefs, the separation proves consequential: raw measures indicate underreaction to information of similar size in both settings, whereas corrected beliefs, unless the assumed priors favor extreme distributions, respond more strongly to the signal under experience than under description. In contrast with the common finding of conservatism in the belief-updating literature, evidence for underreaction is weak and confined to description.

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. I also thank seminar and conference participants at FUR 2026, D-TEA 2025, LACEA-LAMES BRAIN workshop 2025, RISLAB Summer School 2024 for very helpful comments. This study was approved by the Erasmus School of Economics Internal Review Board (IRB-E, reference ETH2324-0286). Preregistered in OSF: <https://doi.org/10.17605/OSF.IO/MUZFH>

1 Introduction

Ambiguity attitudes are usually studied as characteristics of the decision-maker, with the informational environment held fixed: the decision-maker faces an unknown distribution over payoff-relevant events and must choose under uncertainty about their probabilities (Ellsberg, 1961; Schmeidler, 1989; Gilboa and Schmeidler, 1989; Klibanoff et al., 2005). A notable exception is source preference, which shows that attitudes vary across sources of uncertainty and relates this variation to the decision-maker’s knowledge or competence about the source (Tversky and Fox, 1995; Abdellaoui et al., 2011; Wakker, 2010). Knowledge and competence are not given, however: they are acquired in an informational environment in which the decision-maker learns about the source of uncertainty, and one feature of that environment is the informativeness of its signals—observable cues that are informative about, but do not determine, payoff-relevant events (Milgrom, 1981). If attitudes depend on what a decision-maker knows about a source, a question follows: does signal informativeness shape ambiguity attitudes? This paper examines it experimentally. The question has implications beyond the ambiguity literature, for information design, mechanism design, and any applied setting in which a principal chooses how informative a signal to disclose to an ambiguity-sensitive agent: the signal may shape the agent’s attitudes, not only her beliefs.

The difficulty in answering the question is that informativeness moves beliefs by definition, and an incentivized elicitation reads attitudes and beliefs off the same choices: unless the method separates them, a response to a more informative signal cannot be attributed to updating, to a change in attitude, or to both. The belief-updating literature has settled this by making elicitation robust to attitudes, recovering beliefs by neutralizing them (Healy and Leo, 2025). When attitudes are not the main object, that is the right trade; its cost is that the attitude information is discarded along the way. The trade is not available when attitudes are part of the main object. Recent work has begun to examine how uncertain signals affect decision-making, studying signal ambiguity (Epstein and Halevy, 2024), dilation effects (Shishkin and Ortoleva, 2023), and negative value of information (Kops and Pasichnichenko, 2023); because attitudes are what this literature studies, and signals affect beliefs, it must keep attitudes in an elicitation in which beliefs are moving too, and the confound returns.

Epstein and Halevy (2024) find that ambiguous signals increase deviations from Bayesian updating, but acknowledge that the deviation may reflect attitudes toward signal ambiguity and non-Bayesian updating at once, and Liang (2025) addresses the confound indirectly, without quantifying the degree of under- or overreaction to information. We resolve the confound with the belief hedge method of Baillon et al. (2021). From six matching probabilities elicited after the signal is realized, the method recovers two components of ambiguity attitudes—aversion, the tendency to avoid uncertain prospects, and insensitivity, the failure to discriminate between levels of likelihood—without requiring knowledge of beliefs, and it recovers additive beliefs corrected for those attitudes (Li et al., 2019). To our knowledge, this is the first study to tell apart how attitudes and beliefs respond as the informativeness of an uncertain signal varies.

To study this question, we design an experiment in which subjects face three urns that share the same marginal distributions over payoff-relevant events and signals but differ in the strength of the signal-event association, from no association to a pronounced one. By fixing the marginal distributions and varying only the association, we isolate the effect of informativeness on attitudes and beliefs. We implement this design under two information structures: one in which subjects are told the composition of each urn, so the signal-event association is known; and one in which the composition is unknown and subjects may learn it by sampling with replacement, so the association itself remains uncertain. This implementation follows the Description-Experience paradigm (Hertwig, 2015), and we refer to the two arms as Description (DfD) and Experience (DfE). Within each arm, only the association varies across urns and urn order is randomized; between arms, the environments differ in more than whether the association is known, so we treat that comparison as descriptive.

We find that when the association is unknown, 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. When the association is known, we find no evidence that aversion responds to informativeness, while insensitivity shows a similar but weaker pattern, declining only for the strongest signal. This contrast suggests that the informativeness effects observed

when the association is unknown reflect responses to uncertainty about the signal-event association rather than to the compound structure alone. Comparing the two information structures descriptively, aversion is lower and insensitivity higher when the association is unknown and must be learned from experience – a pattern that would be masked by any measure that conflates the two components.

The joint measurement also changes what we conclude about updating. Using raw matching probabilities yields a weight on information of 0.84 when the association is known and 0.87 when it is unknown, suggesting conservatism of similar size in both environments. After correcting for attitudes, this similarity disappears: the weights rise to 0.96 and 1.15, with subjects responding more strongly to the signal when the association is unknown and must be learned from experience. Corrected updating is close to the Bayesian benchmark when the association is known and, under the preregistered uniform prior, exceeds it when the association must be learned from experience. Failing to separate attitudes from beliefs therefore not only makes both environments appear conservative, but also masks a difference in responsiveness to information.

This paper contributes to three 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 informativeness. The result is in the spirit of source preference theory (Tversky and Fox, 1995; Abdellaoui et al., 2011; Wakker, 2010), which relates variation in ambiguity attitudes across sources to the decision-maker’s knowledge or competence about them (Heath and Tversky, 1991). To the literature on belief updating (Benjamin, 2019; Ambuehl and Li, 2018; Augenblick et al., 2024), we show that failing to separate attitudes from beliefs can reverse the qualitative diagnosis of updating. Raw matching probabilities suggest conservatism in both environments. Attitude-corrected beliefs reveal greater responsiveness to the signal when the association is unknown and learned by experience than when it is known, a difference absent in the raw measures; corrected beliefs are approximately Bayesian when the association is known and react more strongly than the Bayesian benchmark when it is unknown and learned by experience. Finally, to the Description-Experience gap literature (Hertwig, 2015; Wulff et al., 2018; Bohren et al., 2024; Oprea and Vieider, 2024), we

extend the paradigm from learning about payoff distributions to learning about signal–event associations. The decomposition reveals that the two environments differ along several dimensions that aggregate measures would obscure: aversion and insensitivity move in opposite directions across environments, while attitude-corrected beliefs respond more strongly to the signal in Experience. The opposite movement of the attitude components is difficult to reconcile with accounts in which a common underlying distortion drives differences across learning modes. More broadly, the component-specific responses to signal informativeness may help explain aggregate patterns previously documented in this literature.

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

Subjects faced three urns that differed in the informativeness of a signal about a payoff-relevant event. Each urn contained 36 balls: 12 blue, 12 yellow, and 12 red, so that each color occurred with probability $1/3$; 18 balls were marked A and 18 were marked B, so that each letter occurred with probability $1/2$. Subjects were also told that each color had at least one ball marked with each letter. We held these two marginal distributions fixed and varied only their joint distribution. Following Shishkin and Ortoleva (2023) and Epstein and Halevy (2024), we fixed the prior distribution over payoff-relevant events; unlike them, we also fixed the marginal distribution of signals, so that the only variation across urns is in the signal-event association. The resulting association was absent, weak, or strong (Table 1); we refer to these as the No, Weak, and Strong urns.

In the two informative urns, Weak and Strong, A favored red and disfavored blue, while B did the reverse; neither letter was informative about yellow. Subjects encountered all three urns in randomized order.

<i>Urn</i>	<i>Blue</i>		<i>Yellow</i>		<i>Red</i>	
	$P(A B\ell)$	$P(B B\ell)$	$P(A Y)$	$P(B Y)$	$P(A R)$	$P(B R)$
No	1/2	1/2	1/2	1/2	1/2	1/2
Weak	1/3	2/3	1/2	1/2	2/3	1/3
Strong	1/6	5/6	1/2	1/2	5/6	1/6

Table 1: Signal likelihood conditional on color, by urn

For each urn, subjects first received information about its composition according to their assigned information structure, described below. The computer then drew a ball. Its color was kept hidden and its letter was revealed. Subjects subsequently chose between a bet paying \$20 if the hidden ball had a specified color and a lottery paying \$20 with probability p . For example, a subject might choose between receiving \$20 if the hidden ball was red and a lottery paying \$20 with probability p , as illustrated in Figure 1. The matching probability is the value of p at which the subject is indifferent between the two prospects. We elicited six matching probabilities using a bisection procedure: three for single-color events and three for their complements. Monotonicity was enforced in the choices, so that the matching probability of a composite event could not be lower than that of any of its constituent single events. Further details of the bisection procedure are provided in Appendix D.

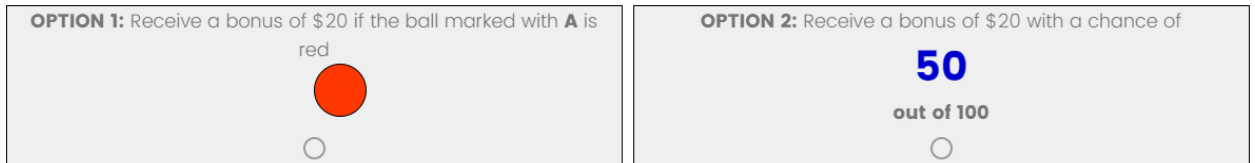


Figure 1: Example of a matching probability

Subjects received a \$5 participation payment and could earn a \$20 bonus based on one randomly selected choice. The choice to be implemented for payment was selected before any decisions were made, but kept hidden from subjects, so that uncertainty attitudes did not affect incentive compatibility.

2.2 Information arms

We implemented the same three urns under two information structures, following the Description-Experience paradigm (Hertwig, 2015).

In the *Description* arm (DfD), subjects were told the exact composition of the urn: how many balls of each color carried each letter. The signal-event association was therefore known. Choices in this arm involve compound risk but no uncertainty about the association itself. Figure 2 shows the information presented to subjects.

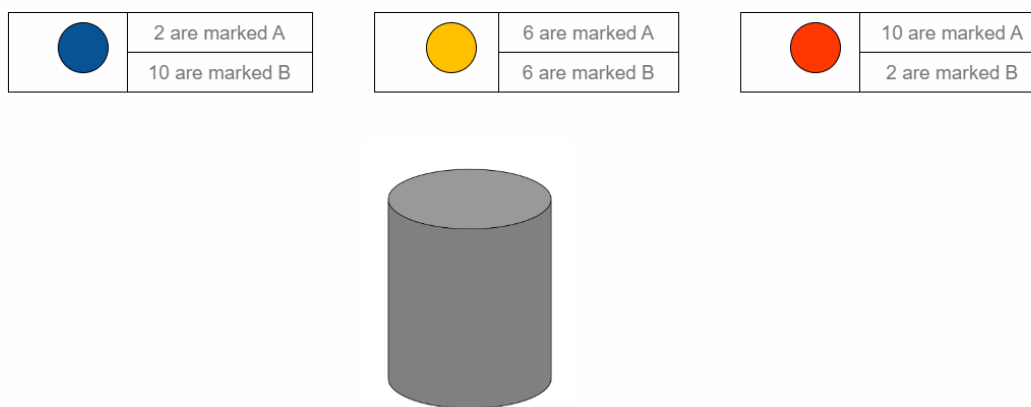


Figure 2: Interface of the information in the urn in the DfD arm

In the *Experience* arm (DfE), subjects were not told the urn’s exact composition. Instead, they learned about it by sampling 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). The signal-event association was therefore unknown and had to be inferred from observed draws. To limit memory demands, the complete sampling history remained visible throughout sampling and during the subsequent choices (Cubitt et al., 2022). Figure 3 shows the sampling interface. Sampling behavior is reported in Appendix I.

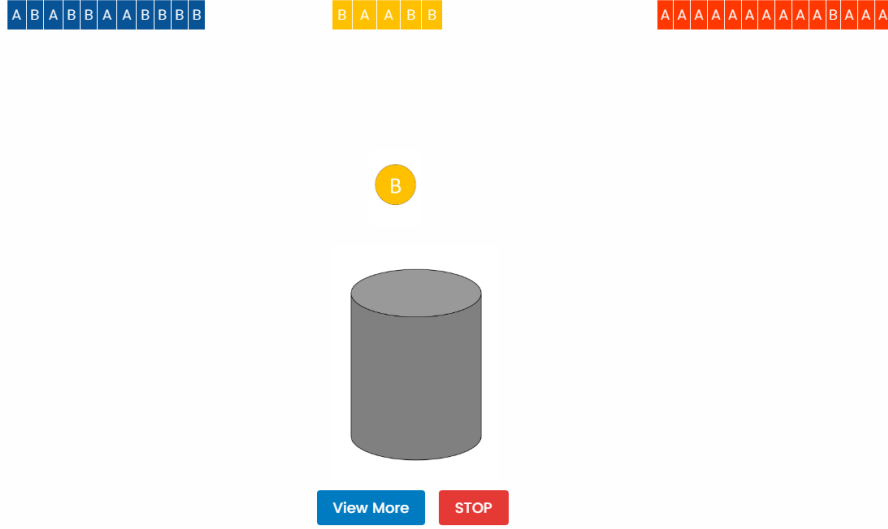


Figure 3: Interface for the Sampling stage of DfE arm

3 Analysis

Eliciting the six matching probabilities allows us to adopt the belief hedge method of Baillon et al. (2021), using the specification of Baillon et al. (2018). The method measures ambiguity attitudes without knowledge of subjects' beliefs, and recovers those beliefs in turn. Both directions matter here. Variation in matching probabilities across urns can speak to how informativeness shapes attitudes only if it reflects attitudes rather than beliefs; and the belief measures can speak to updating only if they are free of attitudes. The two need not move together: after a strong signal, subjects may place greater probability on the color the signal favors while at the same time becoming less uncertain about the signal-event association. Separating these responses allows us to study how informativeness affects attitudes and belief updating independently.

3.1 Ambiguity attitudes

Individuals in both arms face three urns $u \in \{N, W, S\}$, where N , W , and S denote the No, Weak, and Strong urns. 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 $m_u(Col) + m_u(Col^c) < 1$. Measuring how beliefs are contaminated used to require eliciting the uncontaminated beliefs beforehand; with the belief hedge method (Baillon et al., 2021) this is no longer necessary.

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

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

For Subjective Expected Utility maximizers we have $\bar{m}_{S_u} = 1/3$ and $\bar{m}_{C_u} = 2/3$ 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 compound urn's risk in favor of the simple 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-insensitivity¹ index (Dimmock et al., 2016).

¹a-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 literature 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. The null hypothesis is that the indices are equal across the No, Weak, and Strong urns within the same treatment arm,

$$\bar{b}_N^j = \bar{b}_W^j = \bar{b}_S^j, \quad \bar{a}_N^j = \bar{a}_W^j = \bar{a}_S^j,$$

with $j \in \{DfD, DfE\}$, tested with repeated-measures ANOVA. Because a stronger signal-event association should make the event more appealing, the alternative we consider is directional: the indices decrease as the signal becomes more informative,

$$\bar{b}_N^j \geq \bar{b}_W^j \geq \bar{b}_S^j, \quad \bar{a}_N^j \geq \bar{a}_W^j \geq \bar{a}_S^j,$$

with at least one strict inequality. Accordingly, following the preregistration, post-hoc pairwise comparisons test each ordered contrast one-sided, with Bonferroni correction. 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:

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

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

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

$$\bar{a}_u^{DfD} \neq \bar{a}_u^{DfE}$$

$$\bar{b}_u^{DfD} \neq \bar{b}_u^{DfE}$$

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 same six matching probabilities also identify beliefs. Once the attitude indices are known, their effect on the elicited probabilities can be removed, and what remains are the a-neutral probabilities: additive probabilities that capture beliefs once corrected for ambiguity attitudes² (Li et al., 2019). The a-neutral probability of color i conditional on the signal is

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

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

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 complement³, 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

²Li 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.

³Ba et al. (2024) argue against using the Grether equations for more than three states 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. By estimating the log-odds

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.

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

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 question⁴ was added about the contents of the urn and subjects who did not correctly answer it were removed from the analysis.

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

This equation will be estimated for both experience and description independently. Our preregistered hypotheses test whether each arm underweights information compared to the Bayesian benchmark ($H_0 : \rho = 1$). Beyond the preregistered per-arm tests, we additionally estimate a pooled specification (Eq. 4) that allows us to infer whether the two arms weight information differently and, through its sign, the direction of that difference.

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

Where DfE in Eq. 4 is a binary variable that indicates whether subject i was assigned to the DfE arm. Thus, ρ_1 captures whether base rate 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 is OLS of each state and its complement, each regression retains the binary logit structure of the standard Grether framework avoiding the multiple pairwise log-odds ratios that their critique targets, while reducing the number of necessary log-odds combinations to equal the number of states. Moreover, the discrete prior over admissible compositions assumed for the DfE arm (Section 3.2.2) yields posterior predictive probabilities $P_i(\sigma_u|Col)$ and $P_i(\sigma_u|Col^c)$ directly, so the log-likelihood ratio for each color versus its complement is well-defined.

⁴The question can be found in Figure 7 of Appendix A

with standard errors clustered at the subject level. As a robustness check, we also report a fixed-effects specification in Appendix K. In DfE the information term is derived from each subject’s sampling history, so traits that affect both sampling and belief reports may correlate the individual component with the regressor; clustering absorbs the within-subject dependence, and fixed effects remove the individual component.

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 Description arm, since individuals know the contents of the urn, this probability corresponds to the actual proportion of letter σ_u in color Col . In the Experience arm, individuals do not observe the urn composition directly, so we model their inference problem over 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 marked B, and that each color has at least one ball of each letter, there are 91 possible urn compositions. We place a symmetric Dirichlet prior over these compositions. This symmetry of the priors entails two assumptions: that no color-letter combination is favored a priori (equal α parameters), reflecting the principle of insufficient reason (Savage, 1954); 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: the concentration parameter α controls prior dispersion — low values favor extreme compositions, high values favor moderate ones, and $\alpha = 1$ yields a uniform distribution over all 91 compositions. After each sampled draw, posterior weights are updated via Bayes’ rule. The signal probability $P_i(\sigma_u|Col)$ is then 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 over compositions given subject i ’s sampling history. By deriving signal probabilities from each subject’s observations rather than the true urn composition, this approach addresses concerns that comparisons in experience-based settings should reflect the information subjects actually faced (Aydogan, 2021; Abdellaoui et al., 2011). The primary analysis uses the preregistered uniform prior $\alpha = 1$. The full derivation of the prior specification, the updating mechanics, and the

implications of alternative α values are provided in Appendix E; analyses with $\alpha = 1.5$ and $\alpha = 0.15$ are reported in Appendix C.

4 Results

4.1 Data

We recruited 400 subjects from prolific, comprising 194 males and 197 females⁵. 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.

	<i>Full sample</i>			<i>After exclusion</i>		
	All	DfD	DfE	All	DfD	DfE
<i>Age</i>						
Mean	38.98	38.35	39.61	37.83	37.54	38.15
SD	12.76	13.17	12.32	12.36	12.89	11.79
<i>Education (%)</i>						
High school	33.08	32.18	34.17	34.58	33.33	35.92
Bachelor	49.00	48.02	50.25	48.47	47.06	50.00
Post-graduate	17.66	19.80	15.58	16.95	19.61	14.08
<i>Gender (%)</i>						
Female	49.00	48.02	50.25	50.17	50.98	49.30
Male	48.26	50.50	46.23	47.12	48.37	45.77
Non-binary	1.99	0.99	3.02	2.37	0.65	4.23

Table 2: Descriptive statistics. Full sample: $N = 400$. After exclusion: $N = 295$ (DfD: 153, DfE: 142), including only subjects who passed both comprehension checks.

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 A. The

⁵The rest of the subjects reported they were non-binary or preferred not to reveal their gender

comprehension question about urn composition (Figure 7) requires understanding that the marginal totals are fixed, so that the allocation of letters to one color constrains the allocation to the remaining colors. This zero-sum constraint is precisely the feature that generates differences in signal-event associations across urns. A subject who does not grasp it faces a task in which the three urns need not differ in any systematic way, making their responses uninformative about the effect of signal informativeness on attitudes or beliefs. The exclusion criterion thus maps directly onto treatment-relevant knowledge rather than onto outcomes.

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. The right side of Table 2 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 Appendix J; analyses that keep these individuals with their probabilities as recovered, without forcing them into the unit interval, do not differ in magnitude from those presented here.

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 ($M_{W-N} = -0.04, t = -1.89, p_{bonf} =$

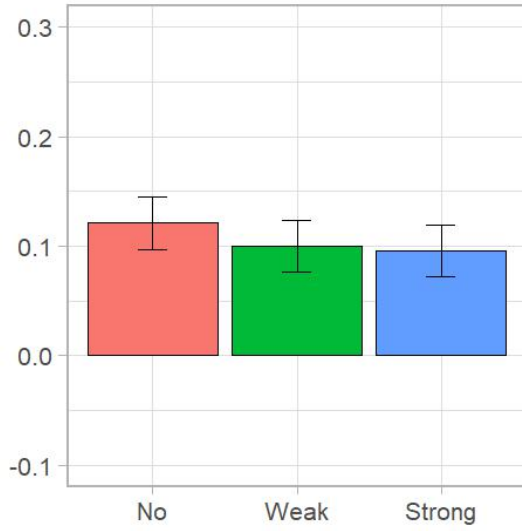
0.09). In the DfD arm, by contrast, aversion does not differ across signal strengths ($F(2, 304) = 1.84, p = 0.16$). When the signal-event association is fully known, varying its strength does not change how much individuals avoid the event; as we show next, it affects only how well they discriminate 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 ($M_{S-N} = -0.11, t = -4.31, p_{\text{bonf}} < 0.01$) and the weak signal ($M_{S-W} = -0.10, t = -3.85, p_{\text{bonf}} < 0.01$), while the no-information and weak signals do not differ ($p_{\text{bonf}} = 0.97$). 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 ($M_{S-W} = -0.05, t = -2.30, p_{\text{bonf}} = 0.03$). The complete pairwise post-hoc comparisons are in Table 9 and Table 10 in Appendix F.

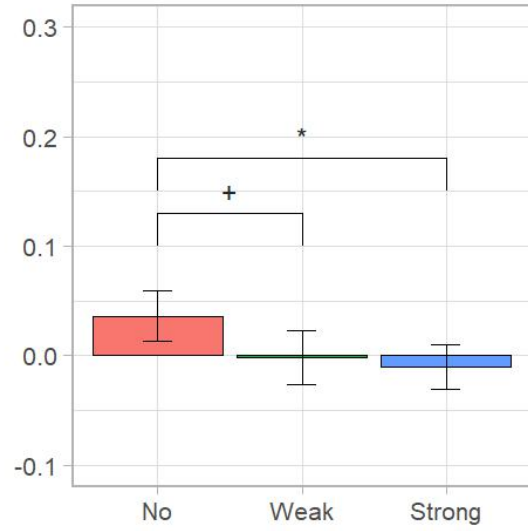
The effects of varying signal informativeness on aversion and insensitivity differ in two respects. First, the two components respond differently to the degree of informativeness: for aversion, it is sufficient for the signal to carry some information; for insensitivity, only the strongest signal produces 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 4 and Figure 5.

A natural concern is that sampling is endogenous: subjects choose when to stop, so the information they carry into their decisions is self-selected. Since subjects sample significantly less from informative urns (Table 12), a quantity-based explanation would predict the opposite pattern: increasing the quantity of information reduces ambiguity aversion (Ert and Trautmann, 2014; Abdellaoui et al., 2025), so more sampling should if anything produce even lower aversion, not higher⁶.

⁶Theoretical conditions exist under which more information can increase ambiguity (Epstein and Schneider, 2007), though these typically involve irreducible ambiguity about the signal-generating process. In our setting, the urn composition is fixed within each urn so the signal-event association is in principle learnable from sampling.

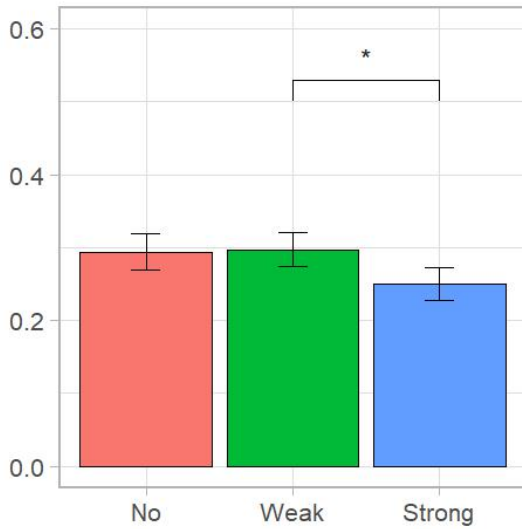


(a) Decision from Description arm

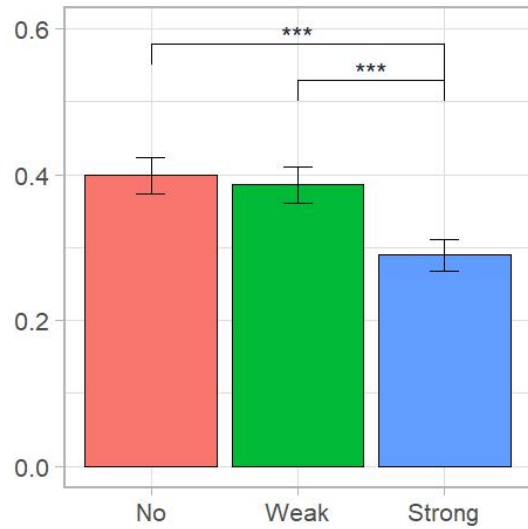


(b) Decision from Experience arm

Figure 4: Mean aversion index by urn. Symbols indicate Bonferroni-adjusted post-hoc pairwise differences in mean aversion across urns *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$



(a) Decision from Description arm



(b) Decision from Experience arm

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

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 3 reports summary statistics for the ambiguity attitude indices by treatment arm. Averaged across urns, 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 across urns is close to neutrality. 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 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. A-insensitivity moves the other way: it is lower in DfD than in DfE both on average and within the No and Weak signal urns, while for the Strong signal urn we find no statistically significant difference between arms (Table 3, $t = 1.26, p = 0.21$) — insensitivity converges across arms, consistent with the within-arm finding that strong signals reduce insensitivity in both conditions. The two arms thus differ on both components, and in opposite directions: aversion is lower in Experience, while insensitivity is higher.

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. Running the Grether regression (Eq. 3) with raw matching probabilities as the dependent variable yields a weight on information of $\hat{\rho} = 0.84$ in the Description arm and $\hat{\rho} = 0.87$ in the Experience arm (Table 6 in Appendix B). Both lie below the Bayesian benchmark ($\rho = 1$), the conservatism typically found in the belief-updating literature (Benjamin, 2019), and they are similar across arms. Using a-neutral probabilities instead yields $\hat{\rho} = 0.96$ for Description and $\hat{\rho} = 1.15$ for Experience (Table 4). The correction changes the reading in both arms: Description moves close to the Bayesian benchmark,

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

Experience moves above it, and the similarity between arms disappears. Failing to separate attitudes from beliefs would therefore do more than overstate underreaction: it would make the two arms appear similar when they differ in how strongly subjects respond to the signal. The remainder of this section uses a-neutral probabilities throughout. Distributions of the measured posteriors and their Bayesian benchmarks for each arm and signal group are reported in Appendix G.

4.3.1 Uniform Prior Analysis

Table 4 shows the parameter that captures the weight individuals give to the information in the signal (weight of information) for both arms and for the pooled models. Estimated separately for each arm, in columns (1) and (2), it is $\hat{\rho} = 0.96$ in Description and $\hat{\rho} = 1.15$ in Experience. The gap between these point estimates is the difference the correction exposes, but by itself does not establish that responsiveness differs across arms. Models (3) and (4) pool both conditions and test this difference directly. The key term is $W.Info \times DfE$, which tells us whether people in the Experience arm react differently to the information in the signal than people in the Description arm: it is 0.19 ($p = 0.03$). The coefficient

on *DfE* itself is not statistically significant, so there is no overall shift in beliefs between arms once we account for the information in the signal; the difference lies in how strongly participants respond to that information. Because in the Experience arm the information term is derived from each subject’s sampling history, it is natural to ask whether a different way of accounting for individual differences changes the estimate (Section 3.2.1). It does not: under subject fixed effects the interaction is 0.18 (Table 20 in Appendix K) and the results are qualitatively similar when we take out the five subjects whose a-neutral probabilities lay outside the unit interval (Appendix J). Thus we can conclude that subjects respond more strongly to the signal when the association must be learned from experience than when it is known.

Having established that correcting for attitudes reveals a difference in responsiveness across arms, we next ask how the estimated weights compare with the Bayesian benchmark. In Description, $\hat{\rho} = 0.96$, and the test of $\rho < 1$ is significant at the 10% level ($t = -1.42$, $p = 0.079$ with subject-clustered standard errors; $p < 0.01$ under fixed effects, Appendix K). Subjects therefore place a weight on information close to the Bayesian benchmark, with only weak evidence of conservatism. In Experience, $\hat{\rho} = 1.15$, and the same test provides no evidence of conservatism: the point estimate lies above the Bayesian benchmark.⁷ Thus, while subjects respond more strongly to the signal in Experience than in Description, only Description shows some evidence of conservatism relative to the Bayesian benchmark. Additionally, in both arms the constant is close to the log-odds of the prior, $\ln(\frac{1}{2/3}) \approx -0.69$ ($t = -0.47$, $p = 0.64$ in DfD; $t = 0.46$, $p = 0.65$ in DfE). This result should not be overread: 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.

⁷Some a-neutral probabilities equal zero or one, for which log-odds are undefined; the preregistered specification excludes them. They are more frequent in Experience than in Description (9.5% against 2.1% of belief observations) and concentrated in the Strong urn (Appendix H). These are the observations in which the signal moved beliefs furthest, so their exclusion attenuates $\hat{\rho}$ in both arms, and more in Experience. Retaining them by winsorizing leaves Description at the benchmark and moves Experience further above it (Appendix L).

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	0.96*** (0.03)	1.15*** (0.08)	0.96*** (0.03)	0.96*** (0.03)
W. Info x DfE			0.19* (0.09)	0.19* (0.09)
DfE			0.01 (0.01)	0.01 (0.01)
Male				0.02 (0.01)
Age				0.01* (0.00)
Bachelor				-0.00 (0.01)
Post-Graduate				0.00 (0.02)
Constant	-0.70*** (0.00)	-0.69*** (0.01)	-0.70*** (0.00)	-0.70*** (0.01)
R ²	0.67	0.34	0.46	0.46
Adj. R ²	0.67	0.34	0.46	0.45
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. Clustered standard errors at the subject level

Table 4: Reactiveness to information for the DfD and DfE arms and pooling both conditions

Comparison with different priors As discussed in Section 3.2.2, the choice of the concentration parameter α affects what is assumed about subjects’ prior expectations over urn compositions, and thus the implied informativeness of each draw. Within the Experience arm, the estimated weight of information ranges from $\hat{\rho} = 0.97$ under $\alpha = 0.15$ (one-sided test of $\rho < 1$: $t = -0.46$, $p = 0.32$), which cannot be distinguished from the Bayesian benchmark, to $\hat{\rho} = 1.27$ under $\alpha = 1.5$, above it; under none of the priors is there evidence of conservatism. In the pooled regressions comparing arms, the interaction $W.Info \times DfE$ varies with the assumed prior: it is 0.32 ($p < 0.01$) under $\alpha = 1.5$ but becomes small and insignificant under $\alpha = 0.15$. Thus, differences are found unless priors that favor extreme urn compositions are assumed. The full results for each prior specification are reported in Appendix C.

4.4 Summary of results

Table 5 summarizes the main findings.

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 respond more strongly to the signal in DfE than in DfD, unless extreme priors are assumed, a difference the correction for attitudes exposes. Relative to the Bayesian benchmark, only DfD shows weak evidence of conservatism; in DfE there is no evidence of conservatism under any of the priors.

5 Discussion

The central finding of this paper is that ambiguity attitudes respond to the informativeness of available signals when the signal–event association must be inferred from experience, and that the two components respond differently: aversion declines as signals become more in-

Panel A: Attitudes across informativeness levels within each arm				
Comparison	Aversion		Insensitivity	
	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
$\alpha = 1$	0.96	1.15	0.19*
$\alpha = 1.5$	—	1.27	0.32**
$\alpha = 0.15$	—	0.97	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 8). Stars denote significance vs zero; tests of $\rho = 1$ are reported in Section 4.3. DfD is invariant to α .

Table 5: Summary of results

formative, whereas insensitivity declines only at the strongest signal. The two responses also differ in scope. Insensitivity declines at the strongest signal whether or not the association is known, though the evidence is weaker when it is known. For aversion, by contrast, we find no evidence of a response when the association is known, so its response is specific to uncertainty about the association. It is worth being precise about what these attitudes are. What we capture is the posterior ambiguity attitude after the realization of an uncertain signal: subjects know the letter on the drawn ball when they choose, and since the color composition of the urn is known, the uncertainty in the Experience arm lies in how much the letter tells about the color, because the relationship between letters and colors is unknown — a setting with parallels to those explored by Kops and Pasichnichenko (2023) and Liang (2025). These are different attitudes from the one studied by Epstein and Halevy (2024). Their attitude toward signal ambiguity is defined by violations of the martingale property, and therefore requires eliciting both unconditional and conditional matching probabilities; ours is the attitude toward the event once a signal has been realized.

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. Since a common source implies a common weighting function (Tversky and Fox, 1995; Abdellaoui et al., 2011), the variation in attitudes we observe is more naturally read as evidence that different informativeness levels define distinct sources, though a formal test of this implication is beyond the scope of the present design.

The attitudes observed in the Description arm deserve separate consideration. In DfD, subjects know the exact composition of each urn, so there is no uncertainty about the signal-event association. The positive aversion observed in DfD therefore reflects avoidance

of compound risk — betting on an event whose probability arises from a two-stage 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). The distinction also matters for measuring beliefs: matching probabilities net out risk attitudes only if the bet on the event and the matching lottery are distorted alike (Ambuehl and Li, 2018), but in our case one is compound and the other is a simple risk. Correcting for attitudes in our Description arm accordingly moves the weight on information closer to the Bayesian benchmark (Section 4.3).

The correction does more than move Description toward the benchmark; it changes what the belief data say about the two arms together. Measured with raw matching probabilities, the two arms look alike: both underweight the signal relative to the Bayesian benchmark, and to a similar degree (0.84 in Description and 0.87 in Experience). This similarity is the product of forces that offset each other. In Description, the correction mainly removes the compression that attitudes impose on the elicited probabilities. In Experience, attitudes compress the elicited probabilities as well, but the corrected beliefs react more strongly to the signal than in Description, so relative to the Bayesian benchmark the two push in opposite directions and the raw estimate lands close to the Description one. A measure that conflates attitudes and beliefs therefore mistakes a common pattern for a common object: the difference in responsiveness across arms is hidden in the raw data, not absent.

Once the conflation is removed and the difference emerges, it can be examined on its own. It depends on the prior assumed over urn compositions in Experience — present under the uniform prior and under priors that favor moderate compositions, consistent with the priors estimated by Aydogan et al. (2024), but absent under priors that favor extreme ones. Where present, it lies in responsiveness rather than in the level of beliefs: with the information in the signal accounted for, beliefs do not shift between arms. Nor does it reflect limited memory of what was sampled, since the full sampling history remained visible during choice. This is plausibly where the distinction between the two learning environments, description and experience, matters most. That the same information can yield different beliefs depending on how it arrives is not new: Hoffman et al. (2011) find that updating differs when two signals are received sequentially rather than simultaneously. Controlling for the information

seen, how evidence is acquired remains a natural candidate for the stronger response in Experience.

Relative to the Bayesian benchmark, the corrected weight on information in Description is close to it, with only weak evidence of conservatism. This contrasts with Ba et al. (2024), who document overreaction once the state space exceeds two states; our specification evaluates each color against its complement and so retains the binary structure of the standard Grether framework (footnote 3). In Experience, there is no evidence of conservatism under any of the priors we consider, and under the preregistered uniform prior the point estimate lies above the benchmark. This runs against the prediction of Augenblick et al. (2024) for signals whose direction is uncertain: people then shade toward no reaction and underreact. In Experience, the meaning of the letter for the color must be inferred from the sample, so that uncertainty is present by design, yet corrected beliefs show no conservatism.

Comparing the two arms, which differ in more than whether the association is known, Experience shows lower aversion but higher insensitivity, so the two attitude components move in opposite directions across environments; corrected beliefs also respond more strongly to the signal. Of these three differences, the one in insensitivity deserves particular attention. Even when the signal-event association is fully known, insensitivity declines for the strong signal. Because insensitivity moves in the same direction in both arms, a shared cognitive mechanism may underlie both declines (Bohren et al., 2024): unless the signal-event association is sufficiently pronounced, subjects may struggle to discriminate between probability levels regardless of whether the association is known or inferred. The same comparison, however, also reveals a difference in levels: insensitivity is higher in Experience, where learning is generally regarded as more natural and less cognitively demanding (Hertwig, 2015), not lower. This suggests that cognitive ease alone does not account for the pattern, and that uncertainty about the signal-event association may also play a role.

A natural concern with our DfE design is that sampling is endogenous: subjects choose when to stop, so the information they carry into their decisions is self-selected. The sampling data speak against this being what produces the attitude patterns. For aversion, subjects sampled less from the two informative urns than from the uninformative one (Appendix I), so an account based on the quantity of information predicts the opposite of what we find: more

sampling reduces ambiguity aversion (Ert and Trautmann, 2014; Abdellaoui et al., 2025), and the urns sampled least should then be the most averse, not the least.⁸ For insensitivity, the pattern is similar in the Description arm, where no sampling takes place: insensitivity differs little between the No and Weak urns and is lower at the Strong urn than at the Weak urn, as in Experience, though the evidence is weaker. That a similar pattern appears without any sampling suggests that it is not sampling that produces it. Finally, within the Experience arm the number of draws is unrelated to aversion and at most weakly related to insensitivity (Tables 13 and 14).

It is not clear, however, that fixing the amount of information exogenously would resolve the identification problem. With both marginal distributions known, the only unknown in our design is the joint distribution of signals and events, so information about the joint determines both how informative the signal appears and how much uncertainty remains about the association. Providing many observations may leave little residual uncertainty, while providing few may make the urns appear similar in informativeness simply because subjects have learned little about their underlying associations. In both situations, ambiguity would be artificially similar across urns. Thus, the amount of information chosen by the experimenter may itself determine how much attitudes can differ across urns. Whether this entanglement confounds identification is an open question, but it is one reason we chose free sampling, where the residual uncertainty at the point of decision is a natural consequence of subjects' own information gathering.

Our findings speak to recent work in behavioral economics accounting for the description–experience gap. Bohren et al. (2024) attribute the gap to attention and memory constraints rather than to preferences or to the mode of learning itself. In our design, however, the non-additive component of attitudes is removed and a memory aid is provided, yet beliefs still differ between learning modes, which suggests that memory constraints alone do not account for the difference we observe. Oprea and Vieider (2024) instead attribute the gap to noisy cognition, with the mode of learning determining the structure of the noise. In

⁸An account in which ambiguity-averse subjects avoid information (Kops and Pasichnichenko, 2023) would predict the least sampling in the urn where aversion is highest. In our data that urn, the No urn, is the one sampled most.

their framework, the influence of prior expectations and the compression of evidence are jointly governed by confidence in the encoded information, so differences in noise move these distortions together. In our data, by contrast, aversion and insensitivity move in opposite directions across arms, while corrected beliefs respond more strongly to the signal in Experience. This places restrictions on accounts in which a common underlying distortion generates the gap.

The patterns we document may also help organize the heterogeneity in the classical description–experience gap, where some studies find underweighting of rare events and others overweighting (Wulff et al., 2018). These studies fit a weighting function with beliefs fixed at sampled or objective probabilities, so any difference in how beliefs respond has nowhere to go but the weighting function: beliefs that respond more strongly than the benchmark push the fitted function toward an S-shape, while insensitivity pushes it toward an inverse-S shape, and which of the two dominates in a given design determines the fitted shape. Within a single mode, moreover, the weighting function can shift with the informational environment, as it does across our urns, and designs differ in how much their samples reveal. Separating beliefs from attitudes in an environment where informativeness varies therefore allows us to hypothesize why the description–experience gap is heterogeneous across studies.

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. Apply-

ing the method to the comparison of known and learned associations reveals that the two conditions differ along both dimensions, with opposite signs: aversion is lower when the association must be learned, while insensitivity is higher. For beliefs, raw measures make the two conditions look alike, both conservative to a similar degree; once attitudes are removed, subjects respond more strongly to the signal when the association must be learned by experience than when it is known, and only the known-association condition retains any sign of conservatism. This difference 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 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. (2025). 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.
- 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 mns.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.
- Epstein, L. G. and Schneider, M. (2007). Learning under Ambiguity. *The Review of Economic Studies*, 74(4):1275–1303. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].
- Ert, E. and Trautmann, S. T. (2014). Sampling experience reverses preferences for ambiguity. *Journal of Risk and Uncertainty*, 49(1):31–42.
- 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.
- Healy, P. J. and Leo, G. (2025). Belief elicitation: A user’s guide. In Snowberg, E. and Yariv, L., editors, *Handbook of Experimental Methodology*, volume 1, pages 81–162. North-Holland.
- Heath, C. and Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1):5–28.
- Hertwig, R. (2015). Decisions from Experience. In *The Wiley Blackwell Handbook of Judgment and Decision Making*, pages 239–267. John Wiley & Sons, Ltd.

- Hoffman, R. M., Kagel, J. H., and Levin, D. (2011). Simultaneous versus sequential information processing. *Economics Letters*, 112(1):16–18.
- 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. (2025). Learning from Unknown Information Sources. *Management Science*, 71(5):3873–3890. Publisher: INFORMS.
- 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].
- 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.
- Oprea, R. and Vieider, F. M. (2024). Minding the gap: On the origins of probability weighting and the description-experience gap. *Working Paper*.

- 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 Comprehension Questions

Consider the following choice scenario:

OPTION 1: Receive a bonus of \$20 if the ball is red OR yellow   ☐	OPTION 2: Receive a bonus of \$20 with a chance of 26 out of 100 ☐
--	--

From the hypothetical situations described below mark the one that would **NOT** give you the bonus.

- ☐ Option 1 is chosen and the drawn ball is yellow.
- ☐ Option 1 is chosen and the drawn ball is red.
- ☐ Option 2 is chosen and the computer draws a 7.
- ☐ Option 2 is chosen and the computer draws a 54.

Back

Next

Figure 6: Comprehension question about payment mechanism

In summary:

Each urn has:

- **12 Red, 12 Yellow and 12 Blue balls** (equivalently $\frac{1}{3}$ Red, $\frac{1}{3}$ Yellow and $\frac{1}{3}$ Blue).
- **18 of the balls are marked A** and **18 are marked B** ($\frac{1}{2}$ marked A and $\frac{1}{2}$ marked B).
- **At least one ball** in each color is **Marked A** and **at least one** is **Marked B**.
- Focus on the **proportion of 'A' or 'B'** marked balls within each color.

The following question will test your understanding of the composition of the urns

Suppose that the urn has the following composition:

	9 are marked A
	3 are marked B

	6 are marked A
	6 are marked B

	@@ are marked A
	?? are marked B

From the options below select the only possible values for @@ and ??.

- ☐ @@ is equal to 9, ?? is equal to 3
- ☐ @@ is equal to 6, ?? is equal to 6
- ☐ @@ is equal to 4, ?? is equal to 8
- ☐ @@ is equal to 3, ?? is equal to 9

Back

Next

Figure 7: Comprehension question about content of the urn

B Analysis of information weight using raw matching probabilities

	DfD	DfE
Weight of Info. (ρ)	0.84*** (0.03)	0.87*** (0.05)
Constant	-0.59*** (0.01)	-0.50*** (0.02)
R ²	0.62	0.41
Adj. R ²	0.62	0.41
Num. obs.	1377	1278

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$. DfE model estimated with $\alpha = 1$. The dependent variable is the log-odds of the elicited matching probability of the event relative to its elicited complement, $\ln(m_u(Col)/m_u(Col^c))$. Standard errors clustered at the subject level.

Table 6: Weight of information estimated using raw matching probabilities.

C 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. Exploring how

the assumed prior affects the estimated responsiveness to information is therefore 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 7 shows the estimate of the average weight of information for each assumed concentration parameter α . With $\alpha = 0.15$, the estimate is close to the bayesian benchmark and we find no evidence of underreaction ($t = -0.46$, $p = 0.32$). In contrast, with $\alpha = 1.5$ the estimate is greater than one. Thus there is no evidence of conservatism with any of the assumed priors in the experience arm.

	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 0.15$
Weight of Info. (ρ)	1.15*** (0.08)	1.27*** (0.09)	0.97*** (0.07)
Constant	-0.69*** (0.01)	-0.69*** (0.01)	-0.68*** (0.01)
R ²	0.34	0.34	0.33
Adj. R ²	0.34	0.34	0.32
Num. obs.	1146	1146	1146

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$ Standard errors clustered at the subject level

Table 7: Reactiveness to information for different initial parameters α for the Experience arm

Table 8 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 and significant, implying stronger reaction 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.

Similar analyses with alternative model specifications are reported in Appendix K of the Supplemental Appendix. The patterns are very similar: results for $\alpha = 1.5$ do not change,

	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 0.15$
Weight of Info. (ρ)	0.96*** (0.03)	0.96*** (0.03)	0.96*** (0.03)
W. Info x DfE	0.19* (0.09)	0.32** (0.10)	0.01 (0.08)
DfE	0.01 (0.01)	0.00 (0.01)	0.02 (0.01)
Constant	-0.70*** (0.00)	-0.70*** (0.00)	-0.70*** (0.00)
R ²	0.46	0.46	0.45
Adj. R ²	0.46	0.46	0.45
Num. obs.	2470	2470	2470

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$ Standard errors clustered at the subject level

Table 8: Pooled models of reactivity to information for different parameters α for the Experience arm

while for $\alpha = 0.15$ the pooled interaction remains close to zero.

D Bisection Method

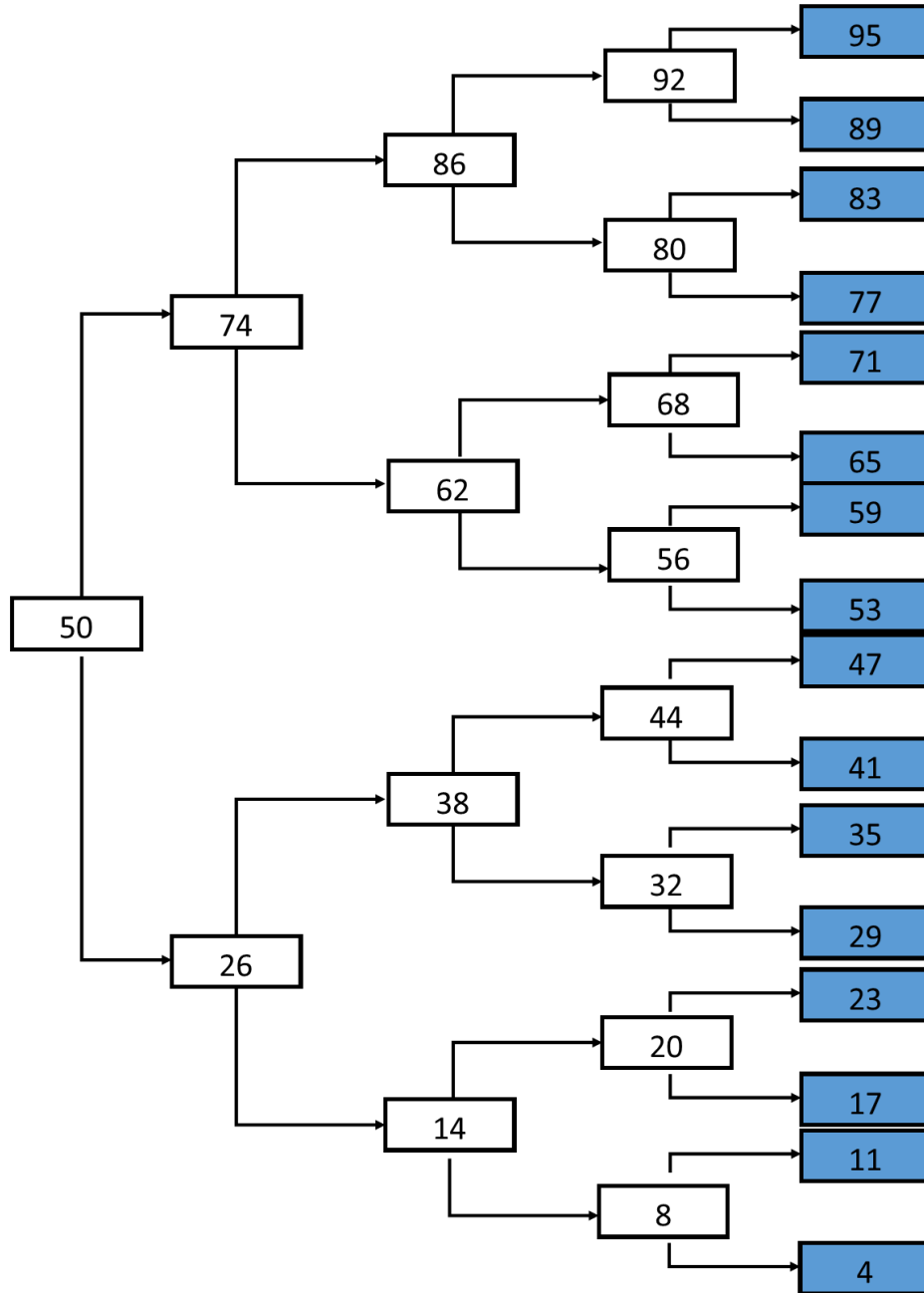


Figure 8: Sequence of lottery choices in the bisection method.

In this experiment, we implement the bisection method in the following way. Individuals first faced a choice, as seen in Figure 1. 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 8. The mechanic was as follows if subjects chose to bet on the color, then 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.

E Dirichlet Prior Specification

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 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.

F 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

G Beliefs

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 9 for DfD and Figure 10 for DfE.

For individuals in the DfD arm, the main distribution statistics are summarized in Figure 9. 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 and its distribution tells us there is not a lot of deviation. For weak and strong signals, the main pattern seems to be that the sample's mean posteriors (in black) are very close to the bayesian posterior (in purple), if anything a little bit lower when the signal increases the posterior.

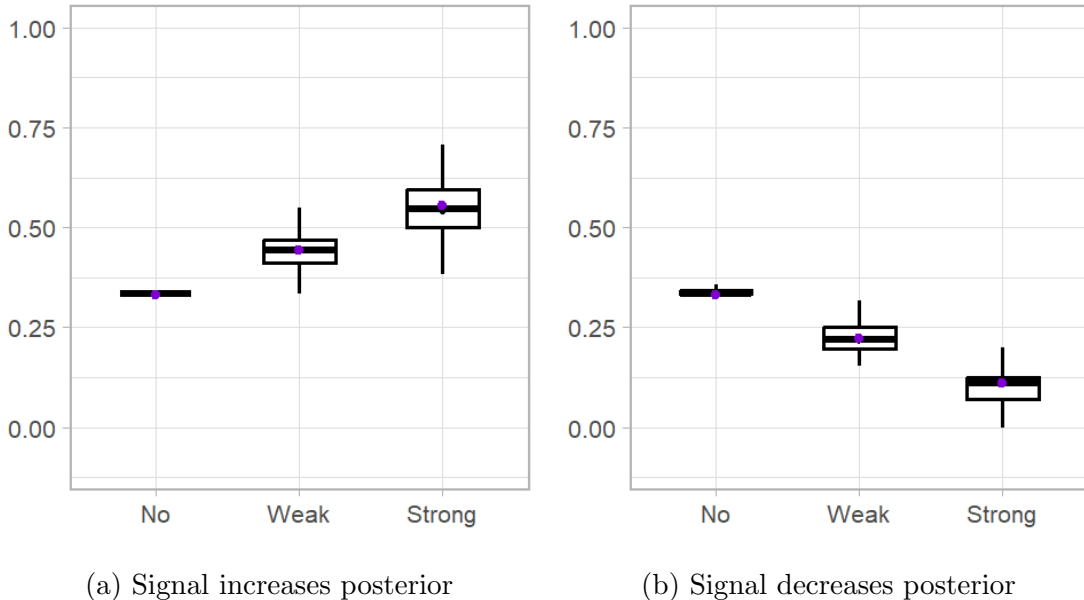


Figure 9: Comparison actual posterior (Black) vs Bayesian benchmark (Purple) in DfD arm

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 10 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 10 shows that the mean of the measured posterior when the signal increases is higher in the strong and weak urn, When the signal decreases the posterior, the mean of the strong and weak signal is lower than the benchmark. Both of these patterns indicate overreaction. This underlines the importance of the analysis of the under- and overreaction using Eq. 3 to derive inferences about the subjects' behavior. Figure 11 shows the comparison of the benchmarks with the values of $\alpha = 1.5$ and $\alpha = 0.15$.

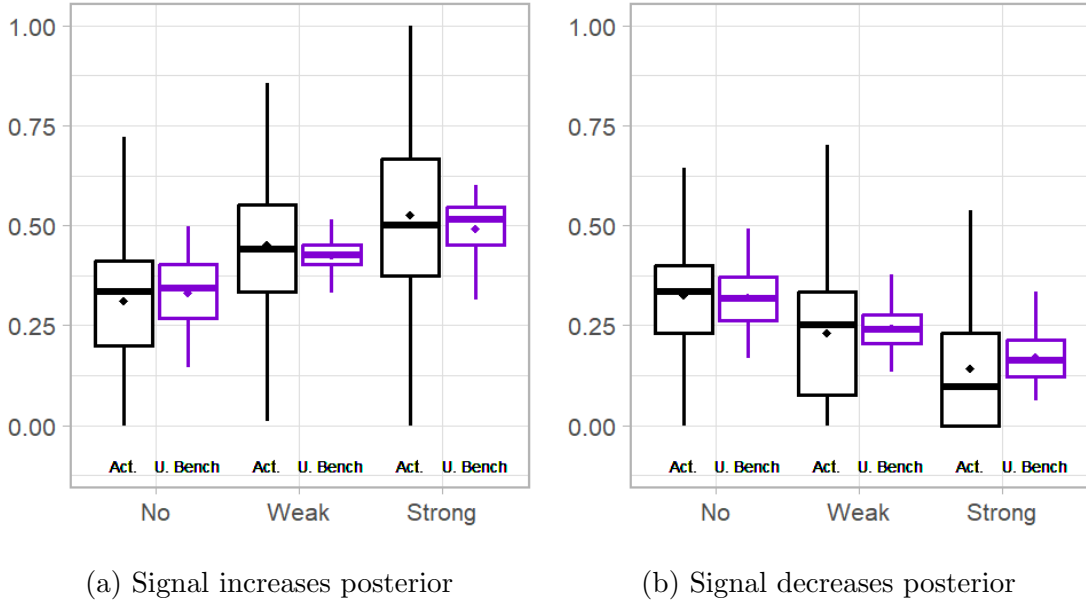
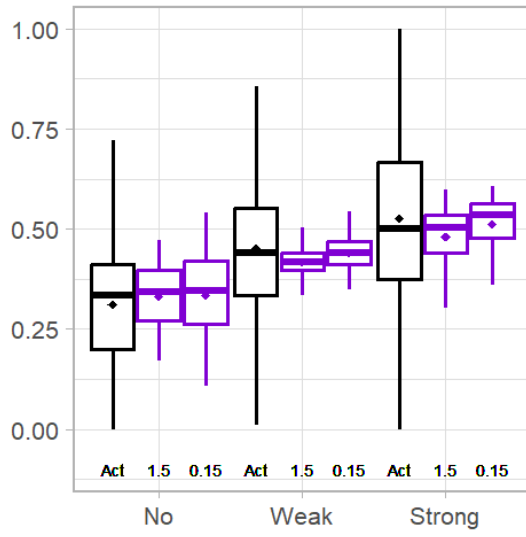
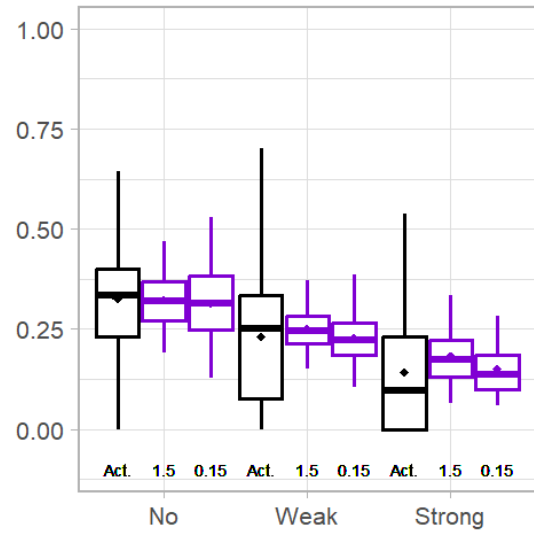


Figure 10: Comparison actual posterior (Black) vs Bayesian benchmark for $\alpha = 1$ (Purple) in DfE arm

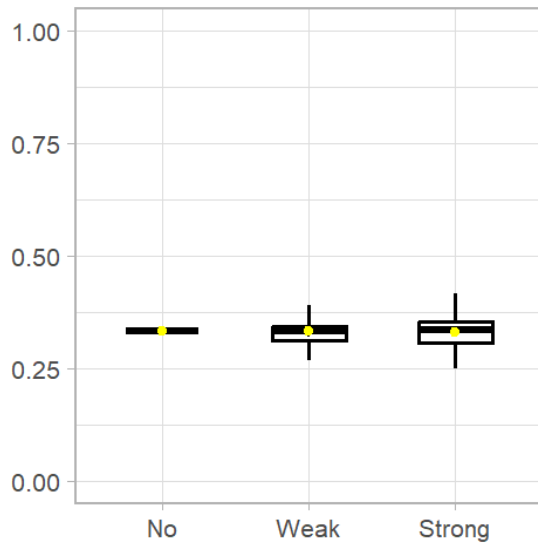


(a) Signal increases posterior

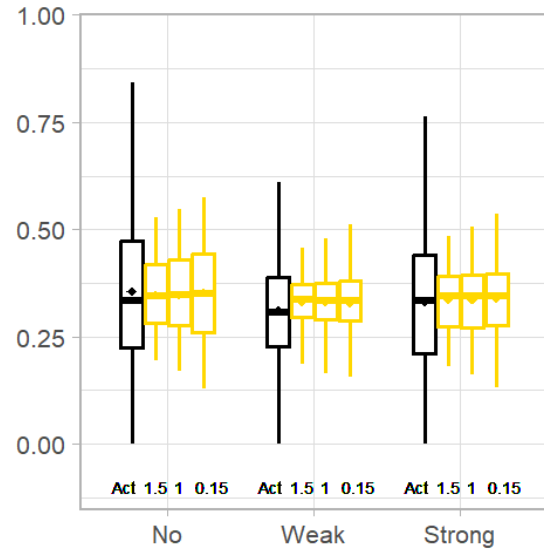


(b) Signal decreases posterior

Figure 11: Comparison actual posterior (Black) vs Bayesian benchmarks (Purple) with $\alpha = 0.15$ and $\alpha = 1.5$ in DfE arm



(a) Description



(b) Experience

Figure 12: Comparison Actual Posterior and Bayesian Posterior Yellow Signal

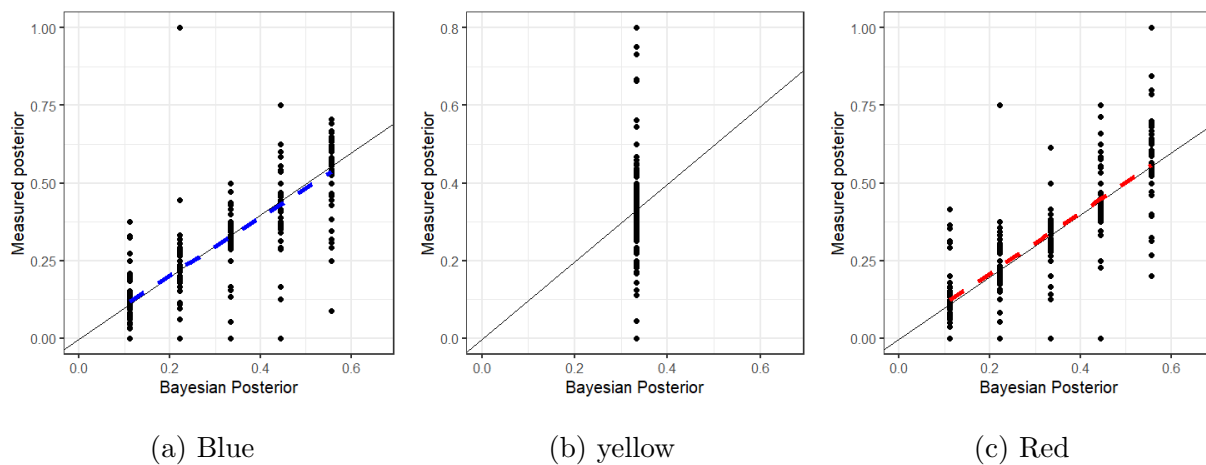


Figure 13: Plot points with different prior parameters Description

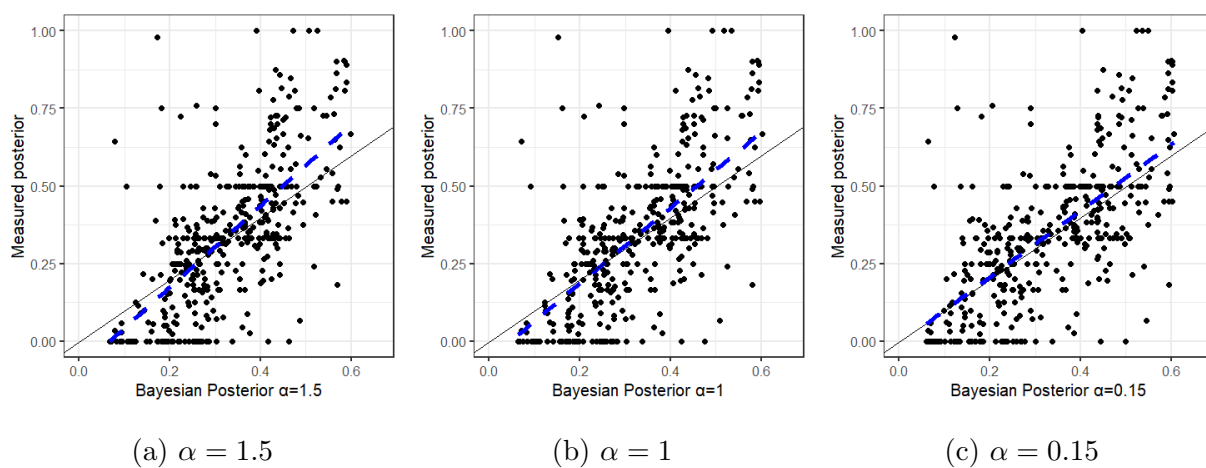


Figure 14: Plot points with different prior parameters color Blue Experience

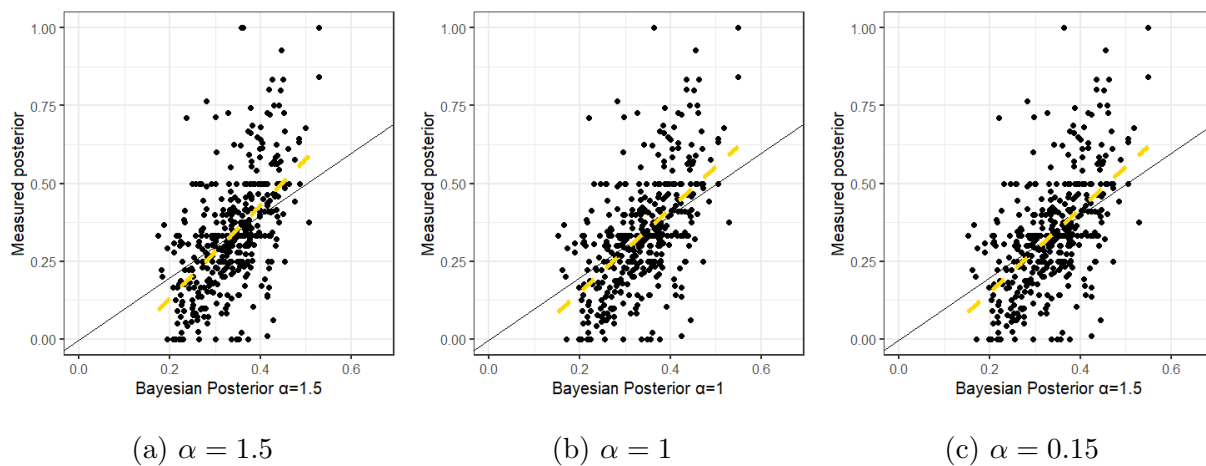


Figure 15: Plot points with different prior parameters color yellow

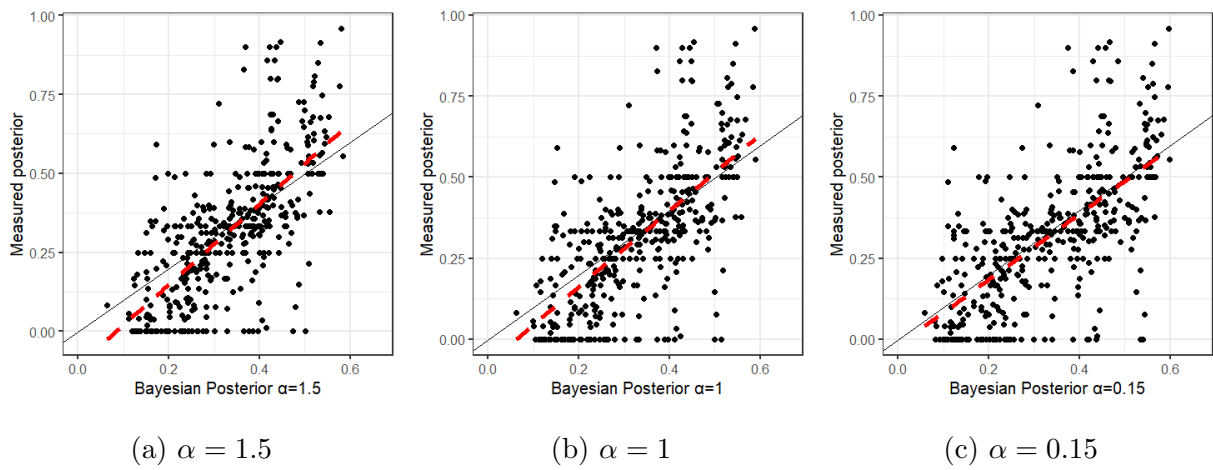


Figure 16: Plot points with different prior parameters color Red Experience

H zeros in probabilities and Na's in logodds

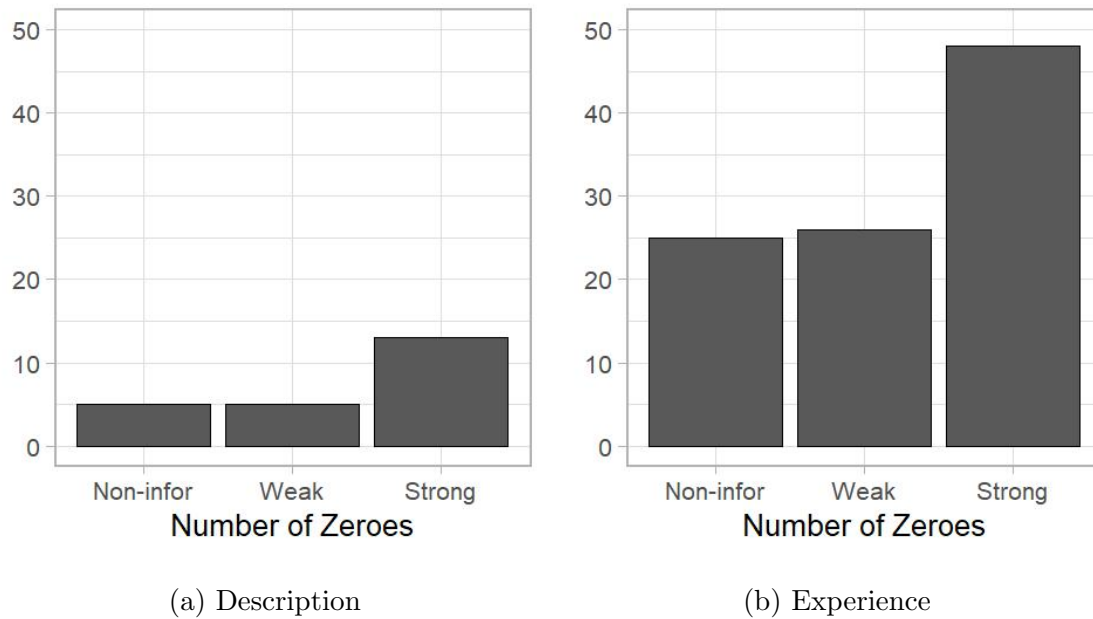


Figure 17: Frequency of subjects with one zero in the probability estimation by Informativeness of the urn

I 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 (0.03)		0.01 (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*** (0.02)		0.39*** (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-insensitivity against sampling

J 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 significant 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	-2.90	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	2.88	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 15: 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 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

	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 16: Mean comparison by signal accuracy in aversion for the experience condition 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 17: 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 19 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)	0.96*** (0.03)	1.13*** (0.08)	0.96*** (0.03)	0.96*** (0.03)
W. Info x DfE			0.18+ (0.09)	0.17+ (0.09)
DfE			0.01 (0.01)	0.01 (0.01)
Male				0.02 (0.01)
Age				0.01+ (0.00)
Bachelor				-0.01 (0.01)
Post-Graduate				0.00 (0.02)
Constant	-0.70*** (0.00)	-0.69*** (0.01)	-0.70*** (0.00)	-0.70*** (0.01)
R ²	0.67	0.33	0.46	0.45
Adj. R ²	0.67	0.33	0.45	0.45
Num. obs.	1324	1119	2443	2434

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$. Clustered standard errors at the subject level

Table 18: 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)	0.96*** (0.03)	0.96*** (0.03)	0.96*** (0.03)
W. Info x DfE	0.18 ⁺ (0.09)	0.29** (0.10)	-0.00 (0.08)
DfE	0.01 (0.01)	0.00 (0.01)	0.02 (0.01)
Constant	-0.70*** (0.00)	-0.70*** (0.00)	-0.70*** (0.00)
R ²	0.46	0.46	0.45
Adj. R ²	0.45	0.46	0.45
Num. obs.	2443	2443	2443

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$. Clustered standard errors at the subject level

Table 19: Pooled models of reactivity to information for different parameters α for the Experience arm

K Fixed effects estimation reaction to information

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	0.95*** (0.02)	1.13*** (0.05)	0.95*** (0.03)	0.95*** (0.03)
W. Info x DfE			0.18*** (0.05)	0.18*** (0.05)
Controls	No	No	No	Yes
R ²	0.68	0.33	0.45	0.45
Adj. R ²	0.63	0.24	0.38	0.38
Num. obs.	1324	1146	2470	2461

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

Table 20: 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)	0.95*** (0.03)	0.95*** (0.03)	0.95*** (0.03)
W. Info x DfE	0.18*** (0.05)	0.30*** (0.05)	-0.00 (0.05)
R ²	0.45	0.46	0.45
Adj. R ²	0.38	0.38	0.37
Num. obs.	2470	2470	2470

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$

Table 21: Pooled models of reactiveness to information for different parameters α for the Experience arm

L Robustness to the treatment of boundary a-neutral probabilities

As noted in Section H, some a-neutral probabilities equal zero or one, for which log-odds are undefined; the preregistered specifications in the main text exclude these observations. The exclusions are not random: 120 of 1266 belief observations in DfE (9.5%, spread over 71 of 142 subjects) and 29 of 1353 in DfD (2.1%) lie at the boundary. In DfE they are concentrated in the strong signal urn (54 of the 120), and their mean absolute information term is almost twice that of the retained observations (0.78 vs. 0.41). The dropped observations are therefore those where the signal had its largest effect.

As a robustness check we retain these observations by winsorizing at 0.01—boundary values are replaced by 0.01 and 0.99 before taking log-odds, with interior values untouched—following common practice in the belief-updating literature (Möbius et al., 2022); 0.01 also corresponds to the midpoint of the boundary response category of the elicitation grid. We report this treatment as a robustness check rather than as the main specification because the exclusion convention is the preregistered one, and because the point estimates under retention depend on the imputed value, whereas the direction of every comparison below does not.

Table 22 reports estimates with standard errors clustered at the subject level, and Table 23 the fixed-effects counterpart, following Appendix K. In DfD, retaining the boundary observations moves the weight of information to $\hat{\rho} = 1.06$: the small shortfall relative to the Bayesian benchmark in the main specification disappears, and the estimate cannot be distinguished from one with clustered standard errors (test of $\rho = 1$: $t = 1.23$) and lies only marginally above it under fixed effects ($t = 1.88$). In DfE, the estimate rises to $\hat{\rho} = 1.66$, and the hypothesis $\rho = 1$ is rejected from above under both specifications ($t = 6.21$ and $t = 10.39$); this point estimate should be read with the caveat that it rests on the value imputed to boundary reports. In the pooled model, the interaction $W.Info \times DfE$ rises from 0.19 to 0.61 and is significant under both estimation strategies ($p < 0.001$). Adding the demographic controls of Table 4 leaves all coefficients essentially unchanged (Model 4 of each table). Retaining the boundary observations thus moves Description to the benchmark

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	1.06*** (0.05)	1.66*** (0.11)	1.06*** (0.05)	1.06*** (0.05)
W. Info x DfE			0.61*** (0.12)	0.61*** (0.12)
DfE			-0.16*** (0.02)	-0.17*** (0.02)
Male				-0.02 (0.02)
Age				-0.00 (0.00)
Bachelor				0.03 (0.02)
Post-Graduate				0.03 (0.03)
Constant	-0.74*** (0.01)	-0.91*** (0.02)	-0.74*** (0.01)	-0.72*** (0.04)
R ²	0.51	0.37	0.41	0.41
Adj. R ²	0.51	0.37	0.41	0.41
Num. obs.	1353	1266	2619	2610

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$. Stars denote significance vs $\rho = 0$; tests of $\rho = 1$ are reported in the text. Standard errors clustered at the subject level. Boundary a-neutral probabilities winsorized at 0.01. Baseline estimates with boundary observations excluded are in Table 4.

Table 22: Reactiveness to information winsorizing boundary a-neutral probabilities at 0.01

	DfD	DfE	Pool	
	(1)	(2)	(3)	(4)
Weight of Info. (ρ)	1.06*** (0.03)	1.66*** (0.06)	1.06*** (0.04)	1.06*** (0.05)
W. Info x DfE			0.61*** (0.07)	0.61*** (0.07)
Controls	No	No	No	Yes
R ²	0.52	0.38	0.42	0.41
Adj. R ²	0.46	0.30	0.34	0.34
Num. obs.	1353	1266	2619	2610

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$. Subject fixed effects estimation. Boundary a-neutral probabilities winsorized at 0.01.

Table 23: Reactiveness to information winsorizing boundary a-neutral probabilities at 0.01, fixed effects estimation

and Experience further above it: the exclusion convention of the main text works against the difference in responsiveness between environments rather than in its favor.