Belief-Based Utility and Signal Interpretation*

Marta Kozakiewicz†

August 7, 2025‡

Abstract

Do people update their beliefs differently after positive versus negative feedback? The existing literature disagrees on the magnitude and direction of the bias. In this paper, I propose a new experiment guided by a simple model of belief choice. One novelty lies in designing a control condition that allows a clear identification of belief *manipulation* and provides robust evidence on the underlying mechanism. Second, the experimental design allows me to examine learning after "unexpected" signals (indicating states to which subjects assigned a prior probability of zero). The analysis sheds light on heterogeneity in beliefs, in particular, why we observe not only people with overconfident or unbiased beliefs but also underconfident individuals. The results demonstrate that the process of belief formation is more complex than previously thought, and the manifestations of belief-based utility—more diverse.

Keywords: overconfidence, belief formation, learning, experiment
JEL classification: C91, D83

*The author gratefully acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (project A01). Financial support from the German Research Foundation (DFG project 462020252) is gratefully acknowledged. The experiment was pre-registered in AEA RCT Registry (Registration No. AEARCTR-0006233) and certified by German Association for Experimental Economic Research (Certificate No. vke5RZr7).

†Frankfurt School of Finance and Management; email: m.kozakiewicz@fs.de

‡The first version of the paper was presented in the academic job market 2020/2021.

1 Introduction

Overconfidence, a tendency to overestimate one's abilities, has received a lot of attention from economists, as it affects a wide range of economic behaviors resulting in significant costs.[1] The literature put forward several explanations for its cause, one of which concerns *asymmetric updating*.[2] It hypothesizes that people might be updating their beliefs differently after "good" and "bad" news, responding more to positive information. While this explanation is consistent with theoretical models that assume direct utility from beliefs (Brunnermeier and Parker, 2005; Caplin and Leahy, 2019), the empirical evidence on asymmetric updating is not conclusive. The experimental data brought mixed results, revealing optimistic or pessimistic updating, or no effect (Benjamin, 2019).[3] Relatively less attention in the literature has been devoted to underconfidence, mostly studied in the context of a gender gap (Bandiera et al., 2022). Yet, a considerable fraction of subjects in experimental studies is underconfident: analysis of the data from recent papers shows that between 31% and 45% of subjects underestimate their ability (see Appendix C.6). Persistent underconfidence suggests an overreaction to negative information, a pattern at odds with the current theoretical and empirical work. This raises a question: Under what circumstances do people respond more strongly to "bad" news? How can it be reconciled with a framework in which agents derive direct utility from their beliefs?

In this paper, I take the next step to fully understand the process of belief formation. Guided by a model of belief choice, I design an experiment that exposes the conditions necessary for asymmetry to arise. The results provide evidence of why some work did not capture a differential response to "good" and "bad" news and point toward a more comprehensive theory of learning about parameters important to the agent.

The experiment has the following structure. In the treatment condition, participants solve an IQ test and receive a noisy signal about their relative performance. Then, they report a subjective probability that the received signal corresponds to their performance. I designed the task in a way that, according to the model, would enhance the chances of capturing the effect. I also use a state and signal space that enables distinguishing "unexpected" signals, that is, signals indicative of states to which subjects assigned zero prior probability. Furthermore, I introduce a new control condition. Participants in the

[1] The behaviors in question include, among many others, excessive selection into competition (Camerer and Lovallo, 1999), excessive trading (Barber and Odean, 2001), suboptimal managerial and investment decisions (Malmendier and Tate, 2005, 2008), and political polarization (Ortoleva and Snowberg, 2015).

[2] Other explanations that consider motivated reasoning rather than cognitive processes fall into two categories: information avoidance (see Golman et al., 2017, for a review of the literature) and selective recall (Chew et al., 2020; Huffman et al., 2022; Zimmermann, 2020).

[3] Inconsistent findings can be attributed to methodological limitations—the lack of a control condition based on the same primitives suggested by Barron (2021) and Drobner and Goerg (2024), or ignoring the timing of the resolution of uncertainty (Drobner, 2022).

control group solve the same IQ test and consider the same signal structure but report their beliefs about *hypothetical* signal realizations. By comparing the two conditions, one can isolate belief manipulation triggered by receiving "good" or "bad" news. The collected data reveal a substantial asymmetry in the treatment condition: there is a significant difference in subjects' responses to "good" versus "bad" news. Subjects tend to interpret positive signals as 11 pp more likely to be informative (20% increase in relative terms) compared to the control condition. The result provides evidence of utility-driven belief manipulation that operates in the direction of the preferred state. There are, however, limits to the effect: it is diminished after signals indicating states to which subjects assigned zero prior probability. I show that this result does not contradict the workings of belief-based utility and can be easily incorporated into the existing framework.

The paper makes several contributions. First, the upgraded design of the treatment condition allows me to capture the effect that, while suggested by the theory, was elusive for previous studies. It also allows to examine a new dimension: expected vs unexpected news, implications of which go beyond the lab. The insights are applicable for research on responsiveness to information, including the growing field of information provision experiments (Haaland et al., 2023). Second, the new control condition offers a benchmark based on the same information over the same unknown state, solving the problem prevalent in the literature on asymmetric updating. Leveraging the controlled lab environment, I show that the method brings results that are consistent with both the theoretical predictions and the Bayesian benchmark. The results validate the use of hypothetical thinking as a way to dampen the emotional responses to news, providing researchers with a tool for studying learning about polarizing issues. Third, together with the model, the experiment allows for a clear identification of belief *manipulation* and provides robust evidence on the mechanism behind asymmetric updating. It advances the research that, so far, provided only indirect evidence on the underlying process. Lastly, I complement the findings with the data on achievement emotions and emotion regulation strategies, casting light on the behavioral underpinnings of the bias.

In Section 2, I describe the design in detail. Subjects first solved an IQ test and then reported a subjective probability of their score falling into the 1^{st}, 2^{nd}, ..., 10^{th} decile of the test score distribution. Thus, I elicited a prior belief distribution over deciles, which I referred to as "ranks". In the treatment condition, participants received a noisy signal about their performance. They were shown a number between 1 and 10 that could be equal to their rank. The framework was described as follows: *There are two boxes. Box 1 contains ten balls with numbers 1 to 10 written on them (each number occurs exactly once). Box 2 contains ten balls with the same number written on every one of them. That number is equal to your rank.* For example, if a subject's rank was 4, Box 2 contained ten balls with the number "4" written on them. Subjects were informed that one ball would be randomly drawn from one of the boxes (either box can be selected with equal probability 0.5) and displayed on their computer screen. Their task was to tell us what they thought: Which box did the ball come from? Using an incentive-compatible mechanism, I elicited beliefs that the ball came from Box 2.[4] The design with two boxes is equivalent to the design used in the literature extended to 10 states of the world.[5] I adopt a richer state and signal space to generate a stronger effect—it is more painful to learn that your score was among the worst 10% than to learn that it was below the median. However, the main advantage of eliciting beliefs about the box instead of rank is that it minimizes confounds arising from people's desire to be consistent (Falk and Zimmermann, 2017). By reframing the question, I avoid asking about the rank multiple times, whereas the composition of Box 2 ensures that I obtain the relevant probability. As a robustness check, I again elicited the entire distribution of beliefs at the end of the study. I confirm that beliefs about the box are consistent with the posterior distribution.

In the control condition, subjects reported their beliefs for every possible signal realization in a procedure akin to the Strategy Method (Brandts and Charness, 2009). I took several steps to alleviate concerns about the non-comparability of the two conditions. For example, I also required the treatment group to consider, one by one, every possible number before they received an actual signal. Participants in the control condition evaluated each number separately, using the same interface as subjects in the treatment condition. Additional tests address issues of anchoring or time effects.

The hypothetical control has an important advantage over alternatives used in the literature: it is based on the same subjective beliefs over the same unknown state. Previous work used to compare how people update beliefs about an "ego-relevant" outcome (e.g., one's performance in an IQ test) and a neutral parameter (e.g., the performance of a robot).[6] However, this comparison involves not only learning about different objects but also updating subjective beliefs, possibly multiple priors, and updating objective

[4] I present subjects with only one signal, in contrast to previous work that introduced multiple signals and reported the average effect (see design comparison in Appendix F). I aim at disentangling the effect of signal valence from the way people aggregate signals. The later task is more complex, hence susceptible to behavioral attenuation (Enke et al., 2024)—a potential explanation for inconsistent findings.

[5] One cannot use the same design as in the literature, because the signal structure becomes too complicated when extended beyond the binary case (see Appendix F). The two-box design introduces 10 states in a way that is easy to explain to participants and allows for a simple elicitation of conditional beliefs.

[6] See, for instance, Coutts (2019), Eil and Rao (2011), and Möbius et al. (2022). A different method was proposed by Drobner and Goerg (2024) who introduce an exogenous variation in subjects' perception of the IQ test validity. Their results are consistent with my findings, however, this treatment manipulation can result in asymmetry even in absence of utility from beliefs. When learning about a parameter that is a *combination* of multiple uncertain factors (in this case, individual test performance depends on 1) one's cognitive ability, and 2) the validity of the IQ test as a measure of cognitive ability, both unknown), beliefs about the underlying factors are all relevant for updating. The proposed manipulation generates a difference in these beliefs without acknowledging its effect on learning.

probabilities given by the experimenter. The treatment manipulation changes more than one feature of the experimental design, undermining causal inference.

In Section 3, I describe a standard model of belief formation in the manner of Brunnermeier and Parker (2005). In the model, an agent forms beliefs about his unknown ability. He starts with a prior and receives a noisy signal with known precision. Then, he chooses a posterior belief facing a trade-off between the utility from the new belief and the cost of belief manipulation, which is increasing in the distance between the chosen belief and an "unmanipulated" posterior. The later assumption, commonly used in the theoretical work, implies that belief formation is partly driven by a rational process. The costs and benefits from belief manipulation are all the agent cares for, as the uncertainty is not resolved at the end of the first period.[7] Crucially, the utility is derived from beliefs that the agent holds *at the moment*. A hypothetical signal does not change the agent's beliefs, and thus, it does not bring additional utility. In absence of utility from beliefs, there is no incentive for belief manipulation. This reasoning, formalized in Section 3.1, serves as the basis for the experimental design and is central for identification.

In Section 4, I present the results. The main outcome variable is the posterior belief revealed through the decision about the box. It corresponds to the posterior chosen by the agent in the model. The "unmanipulated" posterior is approximated using either 1) the Bayesian benchmark, or 2) the decisions in the control condition.[8] I test for asymmetry by comparing the difference between the revealed belief and the unmanipulated posterior after "good" and "bad" signals. First, I show that subjects' reports tend to be 9.6 pp higher after a signal "1", "2", "3", or "4" (the best signals) compared to the reports made after worse signals (p-value of two-tailed t-test = 0.004). The result is robust to changes in the definition of a "good" signal and controlling for observables. Importantly, it is not driven by selection—the coefficient is nearly identical if I estimate the effect on a sub-sample of subjects who observed a random number. In contrast to the treatment, the reports in the control condition do *not* depend on the signal value. Moreover, the estimated weight placed on the Bayesian benchmark is no different than in the treatment condition—subjects seem to assess signals in the same way but without the overreaction to positive information. It provides suggestive evidence that the treatment effect is not due to the different structure of the two conditions (hypothetical vs not) but stems from belief-based utility. Finally, the good-news effect is present across

[7] I incorporated this feature by informing participants that their test results and the details of their payoff will be available to them only one week after the session.

[8] The update is based on the prior probability assigned to the rank indicated by a signal. However, the Bayesian benchmark is less adequate when a prior is equal to zero (Ortoleva, 2012, 2022). For this reason, the primary sample is restricted to subjects who assigned a positive prior probability to the relevant rank. All results are robust to including participants who received a signal to which they assigned zero prior probability but was close to their individual belief distribution.

the experimental conditions. The analysis reveals that people tend to report between 10.8 and 11.9 pp higher beliefs after receiving "good" news—a result consistent with the one obtained by using the Bayesian benchmark.

At the same time, subjects who received an unexpected signal tend to respond more strongly to "bad" news. This result contradicts the assumption that the rational process follows the Bayes' rule after a low-probability event. I address this issue by replacing the Bayesian benchmark with Hypothesis Testing Model by Ortoleva (2012)—an alternative updating rule that circumvents the known limitation of the Bayesian model. The model prescribes the Bayes' rule after a signal expected with a sufficiently high prior probability (a prior probability of zero is a natural choice for the threshold), and deviates from it otherwise. I show that, in this setting, the overreaction to unexpected "bad" news does not contradict the workings of belief-based utility. The conjecture that expected and unexpected news are treated differently finds support in supplementary data: after unexpected "bad" news subjects' decisions are strongly correlated with negative anticipatory emotions such as anxiety. The results suggest that underconfidence can arise due to an affect-driven "paradigm shift"—adopting a new model of the world, as described in Ortoleva (2012)—after unexpected news.

This paper contributes to the literature on motivated reasoning, in particular, the literature concerned with the consumption value of beliefs (Bénabou and Tirole, 2016). In theoretical work, an agent is assumed to derive utility from his beliefs, which creates an incentive to adopt overly optimistic views (Brunnermeier and Parker, 2005; Caplin and Leahy, 2019). On the applied side, behavioral economists have been studying belief updating about ego-relevant traits, such as cognitive ability, with a focus on estimating weights placed on "good" and "bad" signals in various updating tasks. This approach brought inconsistent results. Some authors documented asymmetry in the direction of the preferred state (Drobner, 2022; Drobner and Goerg, 2024; Eil and Rao, 2011; Möbius et al., 2022), others found no asymmetry (Buser et al., 2018; Schwardmann and Van der Weele, 2019; Zimmermann, 2020), or asymmetry in the opposite direction (Coutts, 2019; Ertac, 2011). As these studies differ along several dimensions (described in Appendix F), the findings are difficult to reconcile.[9] The current paper advances a more structured way of thinking about asymmetric updating, connecting experimental work to the theoretical literature on belief formation. In particular, the new control condition enables a direct test of a model of belief choice informing the underlying mechanisms. It also opens the way for studying the utility from beliefs, its functional form and properties, in different

[9]A recent study by Drobner (2022) shows that the differences in findings can be driven by the differences in the expected timing of the resolution of uncertainty, as people expecting to learn the state sooner might face incentives to form more accurate beliefs. My paper complements this work by providing direct evidence of the role of instantaneous utility from beliefs in asymmetric updating.

contexts and across different samples.[10] The subsequent analysis provides new insights into the nature of the bias, revealing differential response to expected and unexpected signals, and its connection to anticipatory emotions and emotion regulation strategies. The final contribution lies in the experimental design that extends the state and signal space, which makes it possible to generate a quantitatively large shift in beliefs. On account of that, my paper hints at why some of the previous work did not capture the effect. If the increase in utility is not large enough, due to a coarse signal structure or a particular functional form, the manipulation might be hard to detect in the data.[11]

2 Experimental Design

The experiment consisted of two parts. In the first part, subjects completed an IQ test. The second part included the elicitation of prior and posterior beliefs and a stage in which subjects received signals (or considered every possible signal realization in the control condition). The outline of the experiment is presented in Figure 1.

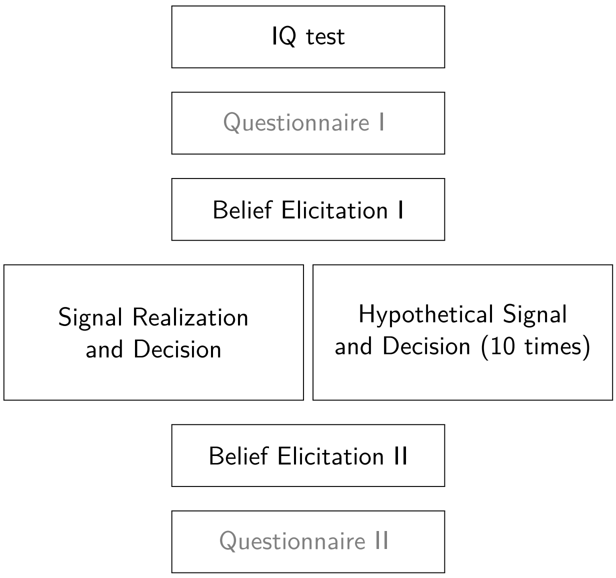


Figure 1: The outline of the experiment.

[10]The method can be applied to other domains—as long as people derive utility from their beliefs, their learning will be guided by similar principles as learning about cognitive ability. Possible applications include the formation of beliefs on politically contentious issues such as climate change or vaccination, learning about others (e.g., their trustworthiness or cooperativeness), or updating beliefs about social norms when confronted with changes in socially acceptable behavior.

[11]For example, the experiments studying updating in the financial domain found little asymmetry (see Barron, 2021). The utility from a "good" signal about a small monetary gain is likely to be lower than the utility from a "good" signal about one's IQ, which is a known predictor of *all* future earnings.

2.1 IQ Test

In the first part of the experiment, I evaluated subjects' cognitive ability using an IQ test.[12] The test consisted of 29 standard logic questions and participants were asked to solve as many of them as possible in 10 minutes. Individual scores were calculated based on the number of correctly answered questions minus the number of incorrect answers, and subjects were paid 0.75 euro for every point they obtained.

Subjects were informed that their earnings from the IQ test will be added to their earnings from the remaining parts of the experiment and paid at the end of the session. They were also informed that, although they will receive the entire sum of money at the end of the study, they will not learn their IQ test score nor how much money they earned in each part. Their test results and the details of their payoffs will be available to them one week after the session. Every participant received a personal link to a website where he could check his (and only his) IQ test result and payment details. This procedure enabled me to minimize the dynamic concerns (e.g., subjects might adopt overly pessimistic beliefs to "prepare" themselves for a disappointing outcome) and focus on the trade-off described in the model.

2.2 Belief Elicitation

At the beginning of the second part, participants were told that they will complete three tasks, for which they can earn up to 12 euro. They were informed that *one task* will be drawn at random at the end of the session, and they will be paid only for that task. In the first task, I elicited subjects' beliefs about their test scores being in the $1^{st}, 2^{nd}, ..., 9^{th}$ and 10^{th} deciles of the distribution of the test scores of 300 participants who took the same test in the BonnEconLab in the past. I introduced 10 "ranks", with Rank 1 denoting the highest rank (assigned to participants whose test scores were higher or equal to the test scores of 90–100% of former participants), and Rank 10 denoting the lowest rank (defined analogously). The first task was to allocate 100 points among the ten ranks in a way that reflects one's beliefs about their relative performance.

[12] Cognitive ability is known to correlate with educational achievement, success in the labor market, and income. For the researcher, it has several advantages over other sources of utility from beliefs, e.g., political ideology. First, there are established methods to assess the parameter of interest, providing us with a measure that is valid and easy to obtain. Second, the utility is almost certainly increasing in the parameter value. Eliciting subjects' beliefs is also relatively straightforward, as we can use incentive-compatible methods developed in the literature. Lastly, one do not need additional information on subjects' political affiliations or preferred policies to formulate the model predictions.

Figure 2: The interface used in the first task (the prior belief elicitation).

A screenshot of the computer interface used by subjects is presented in Figure 2. Participants allocated points by dragging blue arrows on ten scales corresponding to Rank 1 to 10. Subjects were informed they can move the arrows back and forth to correct their choices. The text below each scale informed a participant how many points he allocated to a given rank, and the allocation immediately appeared on the graph to the right. The number above the graph indicated how many points the participant still has to allocate before he can proceed to the next task.

To incentivize truthful reports, I used the binarized scoring rule (Hossain and Okui, 2013) as follows. A random variable X takes one of 10 values: (1,0,...,0,0), (0,1,...,0,0), ..., (0,0,...,1,0), (0,0,...,0,1); the position of 1 indicates the decile into which a subject's IQ test score falls. The agent makes a report $x = (x_1, ..., x_{10})$, where x_i denotes the share of points allocated to the decile $i \in \{1, ..., 10\}$. The researcher observes the IQ test score in the k^{th} decile, the agent wins the prize if the QSR for multiple events,

$$s(x, k) = 2x_k - \sum_i x_i^2 + 1,$$

exceeds a uniformly drawn random variable with the support $[0, 2]$.

The procedure was explained to the subjects in a simple way. More importantly, the instructions directly spelled out the main implication of the method: the probability of getting a large prize (12 euro) is maximized when a person allocates the points in a way that reflects her beliefs about her rank. I followed the same procedure during the second belief elicitation at the end of the study. It is worth noting that, during the first belief elicitation, subjects were not aware that they will be asked to state their beliefs one more time.

2.3 The Signal Stage

After the first belief elicitation, participants received instructions for the second task. The task was explained in simple language, using pictures and illustrative examples. It was framed in a neutral way and described as follows. *There are two boxes. Each box contains 10 balls with numbers written on them. Box 1 contains balls with numbers from 1 to 10, and every number appears exactly once. The composition of the second box depends on your rank in the IQ test. Box 2 contains 10 balls that all have one number written on them, and this number is equal to your rank.*

The composition of the boxes of a person assigned Rank 2 is presented in Figure 3. For every participant, a computer program randomly selected one of the two boxes. Next, one ball was drawn from the selected box and displayed on the participant's screen. The participant did not know which box the ball was drawn from, but he knew that either box can be selected with equal probability $\frac{1}{2}$. After seeing the ball, he had to state his beliefs about the box selected by the computer. I used the same incentive-compatible elicitation method as for the prior belief elicitation. Subjects had to allocate 100 points between Box 1 and Box 2 in proportions that reflected their beliefs about the source

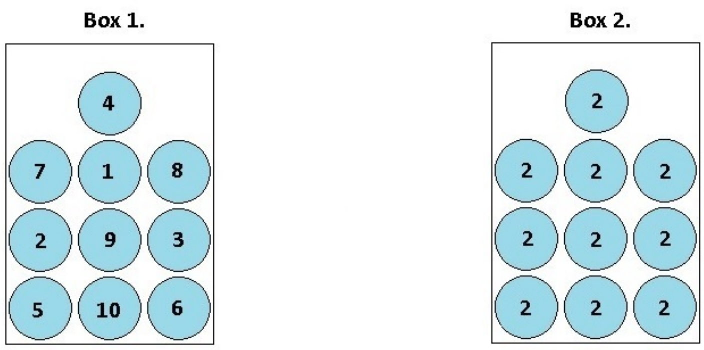


Figure 3: The composition of the boxes of a person whose rank was 2.

of the signal. The probability of getting a large prize (12 euro) was maximize when a person allocated her points in a way that reflected her true beliefs.

Participants were instructed on how to arrive at the Bayesian posterior given one's belief.[13] I explained it in two steps with a simple example. First, I demonstrated how a person should allocate her points after different signal realizations if she knew her rank. Then, I showed how a person should allocate her points if she was not sure about her rank, but was assigning a certain probability to it.

Step 1: How should a person ranked 2 allocate her points if she knew for sure that her rank is 2, and saw a ball with the number "2" on it? There are 10 times as many balls with "2" in Box 2 as there are in Box 1, hence it is 10 times as likely that the ball came from the second box. Therefore, the person should allocate 9 points to Box 1, and 10 times as many, 90 points, to Box 2 (the remaining point should be allocated to the box with a higher probability).

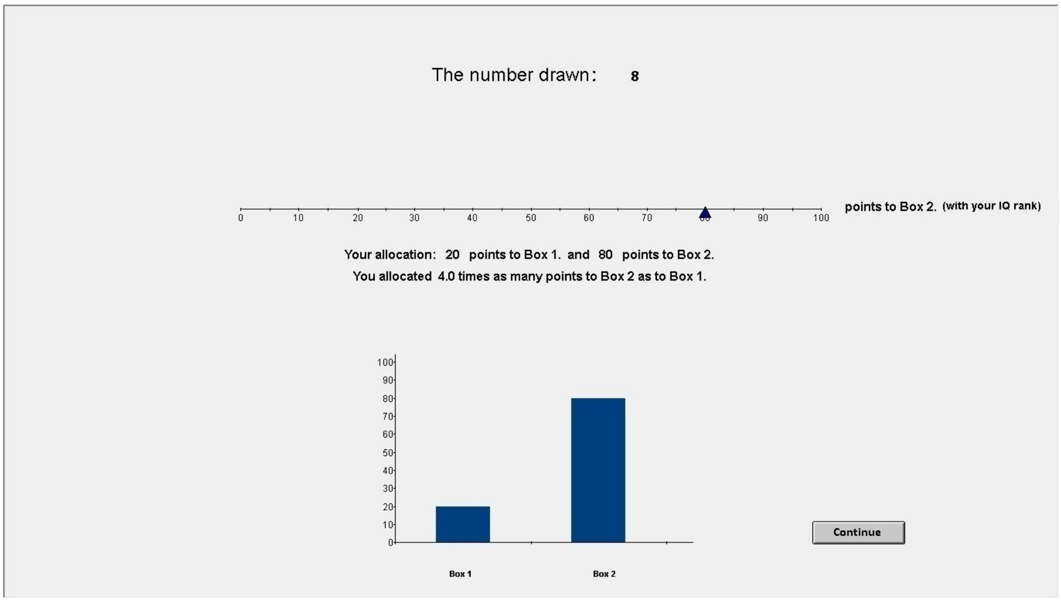
Step 2: What if a person did not know her true rank, but she believed that there is 30% chance that her rank is 2? The same logic applies to this case. One can visualize 30% chance as 3 out of 10 balls in Box 2 having a number "2" on them. In this imaginary case, there are 3 times as many balls with the number "2" on them in Box 2 as in Box 1, implying an allocation of 25 points to Box 1 and 3 times as many (75 points) to Box 2.

A screenshot of the interface used in this task is presented in Figure 4. Importantly, the interface enabled subjects to split the points in the desired proportions without calculating the ratios. This feature was added to minimize computational errors. The text below the scale informed participants about the current allocation and the ratio between the points allocated to the two boxes. By moving the cursor, a subject could choose the number of points corresponding to allocating x times as many points to one of the boxes (with $x \in \{1, 1.1, ..., 99\}$). The graph below showed the current allocation.

Before proceeding to the signal stage, subjects were required to answer a set of control questions. The questions were designed to check participants' understanding of the task including the steps necessary to arrive at the Bayesian posterior.

[13] Detailed explanations allow to minimize mistakes and ensure subjects' understanding of the task. They are particularly relevant in the lab environment, which grants a tight control over the belief formation process but is far from natural (in everyday life, it is uncommon for people to know the precise signal structure and use it to form probabilistic beliefs). Still, there is a risk of framing or creating expectations of what "should" be done in the experiment. Fortunately, I find little evidence of people blindly following the rule: the share of decisions equal to the Bayes' rule is 12% and slightly higher in the treatment condition (3.5 percentage points difference, significant at the 10%-level).

Figure 4: The interface used in the second task (the signal stage).



2.4 Experimental Conditions

I introduced two experimental conditions: treatment and control. In the control condition, subjects stated their beliefs ex-ante, conditional on a number being drawn. They were informed that, although they will see every number, their choices are not entirely hypothetical. Later, one box will be selected at random and one ball will be drawn from the selected box. They will be paid for the decision that corresponds to the number drawn. This procedure is incentive-compatible as the probability of drawing any number is at least 5%. However, if subjects weigh the cost of effort against the expected payoffs, they might exert lower effort in the control condition. I address this issue with additional robustness checks (e.g., comparing the variances of the reports, the share of people choosing the default option, and testing for time effects) in Section.

To alleviate concerns about the non-comparability of the two conditions, I adopted special procedures targeting the issues raised in the literature. One concern is that using the Strategy Method might lead to a better understanding of the game—a consequence of considering the problem from different points of view. For this reason, I also asked participants in the treatment condition to consider every possible signal *before* they saw the actual draw. Subjects were required to go through 10 slides with screenshots of the interface used in the control condition. They were asked to contemplate a hypothetical decision before clicking on the button "Continue" which appeared on the screen after

15 seconds. While only the control group was allowed to enter their choices, both groups were required to think about every signal. Another problem of the Strategy Method concerns framing subjects with the order of options. I addressed this issue by randomizing the order of numbers displayed in the control condition and the order of slides presented in the treatment. Moreover, participants in the two conditions used the same interface—the only thing that differed was the headline, which said: "The number drawn" in the treatment condition, and "Consider the number" in the control condition.

2.5 Questionnaires

Although the hypotheses stated in the paper do not follow from a particular psychological theory (but rather a close inspection of the economic concept of belief-based utility), the idea is related to psychological research on emotions and decision-making (see Lerner et al., 2015, for a comprehensive literature review). One conclusion from this literature is that emotions may influence decisions via changes in the content of thought, and vice versa.[14] I included two questionnaire to shed light on the psychological forces behind the results. The exploratory analysis, delegated to Appendix G, is meant to raise a question about the role of emotion regulation in asymmetric updating, which can be further examined with economic tools. I consider it a promising avenue for future research.

The first questionnaire, displayed after the IQ test, included a short version of the Big-5 personality test (Gerlitz and Schupp, 2005) and the state-trait anxiety inventory STAI (Spielberger, 1983). The second set of questions, displayed after the three tasks, comprised the Emotion Regulation Questionnaire (ERQ) by Gross and John (2003) and a subset of questions from the Achievement Emotions Questionnaire (AEQ) taken from Pekrun et al. (2011). While Big-5 and STAI are often used in behavioral economics, the last two questionnaires require some explanation. The ERQ was designed to assess the habitual use of two strategies commonly used to alter emotions. First, one can alleviate the emotional impact of a situation by reinterpreting it in a different way. This emotion regulation strategy, known as *reappraisal*, relies on "applying mental models to the often ambiguous and incomplete information" (Uusberg et al., 2019). The second strategy, *suppression*, involves "inhibiting ongoing emotion-expressive behavior" (Gross and John, 1998, cited in Uusberg et al., 2019). People differ in their use of reappraisal and suppression, and these differences have implications for their experiences of emotions, their behavior in response to those emotions, and general well-being (Gross and John, 2003). To measure the use of these strategies, I administered a 10-item questionnaire devel-

[14] A similar hypothesis about anxiety has been recently tested by Engelmann et al. (2019).

oped by Gross and John (2003).[15] The AEQ measures *achievement emotions* (emotions directly linked to achievement activities and outcomes) experienced by students in academic settings (Pekrun et al., 2011). I adopted part of the questionnaire to measure the following test-related emotions: enjoyment, hope, pride, relief, anger, anxiety, shame, and hopelessness. I discuss potential links between experienced emotions, the habitual use of emotion regulation strategies, and belief updating in Appendix G.

3 Theoretical Framework

In this section, I present a one-period model that underlies the experimental design. I formulate testable predictions and describe the empirical strategy used in the analysis.

3.1 The Model

An agent is learning about the state of the world ω (e.g., his cognitive ability) that can be high or low, $\omega \in \{H, L\}$. He has a prior belief about his ability being high p_0 and receives a signal $s \in \{H, L\}$ that induces a posterior belief $p_{1,s}$. The agent derives utility from his beliefs $u(p)$. The utility function $u(\cdot)$ is increasing in the subjective probability of the high state, concave and twice continuously differentiable. Because of it, the agent has an incentive to *manipulate* his beliefs by choosing a different posterior $\tilde{p}_{1,s}$, which then enters his utility function. I assume a quadratic cost of belief manipulation that depends on the distance from the unmanipulated posterior $p_{1,s}$. The agent's utility after a signal s has the following form:

$$U(\tilde{p}_{1,s}) = u(\tilde{p}_{1,s}) - \frac{1}{2\gamma}(p_{1,s} - \tilde{p}_{1,s})^2, \tag{1}$$

where $\gamma > 0$ is the cost parameter. The choice of $\tilde{p}_{1,s}$ that maximizes (1) describes the process of forming a belief about one's ability. Because the process is outside of conscious awareness, (1) does not include the monetary reward for truthful reporting—a conscious decision resolved with the elicitation mechanism. The first-order condition gives us:

$$\gamma u'(\tilde{p}_{1,s}) = \tilde{p}_{1,s} - p_{1,s}. \tag{2}$$

The corner solution $\tilde{p}_{1,s} = 1$ is attained if $1 - \gamma u'(1) \leq \tilde{p}_{1,s}$. Note that if u satisfies $\lim_{p \to 0} u'(0) = \infty$, then $\tilde{p}_{1,s} = 0$ is never optimal. The equation (2) describes the

[15] The habitual use of the two strategies is measured by the degree to which subjects agree with particular statements, e.g. "I keep my emotions to myself" or "When I want to feel less negative emotion, I change the way I'm thinking about the situation".

solution to the agent's problem *after* receiving a signal, that is, the problem faced by participants in the treatment condition.

Asymmetry in belief updating

Asymmetric updating emerges when the decision-maker puts higher weight on signals indicating the preferred state compared to the weight he places on the remaining signals (Benjamin, 2019). Overreaction to negative signals corresponds to belief manipulation in the direction of a less preferred state: $\tilde{p}_{1,L} < p_{1,L}$.[16] In my setting, overweighting positive signals is reflected in the manipulation after positive signals being larger than the manipulation after negative signals:

$$\tilde{p}_{1,H} - p_{1,H} > p_{1,L} - \tilde{p}_{1,L}. \tag{3}$$

In the case of no asymmetry, the left-hand side of (3) is equal to the right-hand side.

Unmanipulated belief

Ideally, one would measure the unmanipulated belief held by an agent. In reality, not only the researcher cannot observe it, but the agent himself might not have access to this belief—the process of belief manipulation might operate beyond his consciousness. As a second best, I use two different approaches to approximate its unknown value.

In the first approach, I assume the unmanipulated belief $p_{1,s}$ to be a linear function of the belief formed by an impartial observer—the rational belief $p^B_{1,s}$ (the superscript B denotes the Bayesian update):

$$p_{1,s} = \beta p^B_{1,s}, \tag{4}$$

where $\beta > 0$ captures the degree to which the process of belief formation follows Bayesian updating. I assume this process to be independent of the directional effect arising from belief-based utility. The special case of a prior belief equal to zero is described below.

In the second approach, I use the data from the control condition. A hypothetical signal does not change the agent's belief and does not affect the utility function $u(\cdot)$. The agent keeps deriving utility from his prior belief p_0. He makes a report $\bar{p}_{1,s}$ about the posterior he would form after a signal s. The agent's utility is:

$$U(\bar{p}_{1,s}) = u(p_0) - \frac{1}{2\gamma}(p_{1,s} - \bar{p}_{1,s})^2, \tag{5}$$

[16] In this case, the asymmetry means that the agent manipulates his beliefs by choosing the probability of the low state, $(1 - \tilde{p}_{1,L})$, higher than the unmanipulated probability $(1 - p_{1,L})$. Rearranging gives us the inequality $\tilde{p}_{1,L} < p_{1,L}$, and the extent of manipulation after a negative signal is: $p_{1,L} - \tilde{p}_{1,L}$.

where the second term denotes the quadratic cost of manipulation of beliefs about the conditional posterior. As previously, the cost depends on the distance from the unmanipulated posterior $p_{1,s}$. The agent's problem is to maximize (5) by choosing $\bar{p}_{1,s}$. The first-order condition gives us:

$$\bar{p}_{1,s} = p_{1,s}. \tag{6}$$

Since the agent has no incentive to manipulate his beliefs, he reports the probability equal to the unmanipulated posterior.

Prior probability of zero

It is important to consider a special case: a signal indicating a state to which the agent assigned a prior probability of zero. After such signal, the Bayesian update is not defined, and the agent should keep her initial belief. There are reasons to doubt that it is an appropriate description of human behavior (Ortoleva, 2012, 2022). In light of these findings, I adopt the Hypothesis Testing Model of Ortoleva (2012) for signals outside of subjects' prior belief distribution (± 1 rank to account for inaccurate reporting).

One advantage of the model is that it can be easily combined with the belief-based utility framework. Moreover, for expected signals, it prescribes Bayesian updating, which is widely recognized as a rational benchmark. Only after an unexpected signal—a signal with with an ex-ante probability below a certain threshold—the agent revises her model of the world by choosing a different prior, one that better explains the evidence at hand. The new prior serves as a basis for updating (again, using the Bayes' rule). I describe the model and its implications in more detail in Appendix A.

3.2 Testable Predictions

The overarching question of the paper is whether there is asymmetry in belief formation. To answer this question, I adopt two different approaches. In the first approach, I test the prediction (3) using the model's rational benchmark as a proxy for the unmanipulated belief $p_{1,s}$. The benchmark can be easily calculated for signals with non-zero priors, as the Hypothesis Testing Model boils down to the Bayesian update. For unexpected signals, the effect depends on a set of models under consideration and their probabilities. The direction of the effect is not clear, but one cannot rule out the presence belief-based utility: any observed behavior can be rationalized by some hierarchical structure and belief manipulation (see Appendix A for an illustrative example).

In what follows, I formulate testable predictions for the non-zero priors. I postpone the discussion of unexpected signals until the end of the section. I consider the following regression:

$$Y_i = \alpha_0 \ + \ \alpha_1 \ Y_i^{Bayes} \ + \ \alpha_2 \ X_i^{signal} \ + \ \epsilon_i. \tag{7}$$

The dependent variable Y_i denotes the decision in the main task—how many points a subject allocated to Box 2 after a signal s. This decision reveals the manipulated belief that the state is high after $s = H$ or that the state is low after a signal $s = L$. I regress it on an independent variable Y_i^{Bayes}, which denotes the number of points that *should be* allocated according to Bayes' rule (the Bayesian benchmark), and an indicator variable X_i^{signal}, which takes value 1 if the subject received a "high" signal. If people place a higher weight on "good" signals, it will be reflected in a positive coefficient α_2.

Hypothesis 1.1
Subjects tend to manipulate their beliefs to a larger extent after "good" signals. The coefficient α_2 in (7) is positive.

The equation (2) makes clear that asymmetry stems from the belief-based utility function $u(\cdot)$. According to (6), no asymmetry is expected in decisions about hypothetical signal realizations. When considering a signal, the decision should be guided solely by the rational process.[17] I test this prediction by estimating the regression (7) using the data from the control condition.

Hypothesis 1.2
There is no asymmetry in how participants respond to "good" and "bad" signals in the control condition. The coefficient α_2 in (7) is not significantly different from zero.

In the model, there is no difference in the underlying rational process in the two conditions, as (2) and (6) differ only with respect to the utility component $\gamma u'(\tilde{p}_{1,s})$. In both cases, the unmanipulated belief can be approximated with the Bayesian benchmark. That being so, the weight placed on the Bayesian benchmark should be the same regardless of the experimental condition. The coefficient at the Y_i^{Bayes} should be the same when (7) is estimated using the data from the treatment or the control condition.

[17] For lack of a better term, I use the term "rational process" to denote the process driving the formation of unmanipulated beliefs (beliefs that would be formed by an impartial human observer). While this process might be subject to learning biases, I assume that they are independent of the signal value.

Hypothesis 1.3
The coefficient α_1 in (7) estimated on the data from the treatment condition should be no different from the same coefficient estimated on the data from the control condition.

The causal effect of receiving a "good" signal on belief manipulation can be established in a difference-in-difference analysis. In this approach, the unmanipulated beliefs are approximated with the control data. Pooling the data from the two conditions, I estimate the following regression:

$$Y_i = \beta_0 \; + \; \beta_1 \, Treat_i \; + \; \beta_2 \, X_i^{signal} \; + \; \beta_3 \, Treat_i \times X_i^{signal} \; + \; \epsilon_i, \tag{8}$$

where $Treat_i$ is an indicator variable taking value 1 if a subject was assigned to the treatment condition. The coefficient at the interaction term informs us about the effect of receiving "good" news on beliefs in the treatment compared to the control condition.[18]

Hypothesis 2.1
Subjects tend to manipulate their beliefs after "good" signals to a larger extent in the treatment compared to the control condition. The coefficient β_3 in (8) is positive.

One can also use the data from the control condition to predict, for every participant in the treatment group, a counterfactual outcome: what the subject would have decided, had the signal not affected his belief-based utility. I use the nearest neighbor matching model to predict the counterfactual outcome $\hat{Y}_i$ for every participant in the treatment group. Then, I estimate the following regression:

$$Y_i - \hat{Y}_i = \gamma_0 \; + \; \gamma_1 \, X_i^{signal} \; + \; \zeta_i. \tag{9}$$

The dependent variable is the deviation from the counterfactual outcome, and I regress it on a dummy indicating a "good" signal. If there is asymmetry in belief formation in the direction of the preferred state, the deviations should be greater after positive signals. This effect will be captured by the coefficient at the X^{signal} variable.

Hypothesis 2.2
Subjects tend to deviate more from the counterfactual outcome after "good" signals. The coefficient γ_1 in (9) is positive.

[18] I also estimate (8) controlling for the subject's rank, belief distribution, and the signal value.

Predictions for unexpected signals

The decisions after unexpected signals enable us to test whether Ortoleva (2012) provides a more adequate description of belief updating than the Bayesian model. If an agent updates his unmanipulated belief like a Bayesian, he should place zero probability on the state indicated by the unexpected signal. With additional manipulation due to belief-based utility, the overall effect should be positive. However, if the agent first revises his model of the world, both a positive and a negative effect is possible, depending on a set of models under consideration. Consequently, while a positive effect would not help us distinguish between the two theories (it can be generated by either theory), a negative effect after an unexpected signal would imply that the agent is not Bayesian, lending credibility to the Hypothesis Testing Model. Unfortunately, not knowing the set of models considered by the agent, one cannot calculate the rational benchmark, so we cannot say whether the belief manipulation occurs (see Appendix A).[19]

Predictions for the entire sample

The model prediction for the entire sample, which includes the decisions after expected and unexpected signals, is ambiguous. The overall effect will depend on the direction of the effect for unexpected signals and their proportion in the sample. Importantly, the absence of the effect in the entire sample does not prove the theory of belief-based utility wrong, unless it is driven by signals about states with positive prior probability.

3.3 Identification

The identification strategy relies on the model and two sources of exogenous variation: random assignment to the treatment condition, and random assignment of signals drawn from Box 1. The model provides testable predictions, guides the empirical analysis, and allows us to interpret the identified difference as belief manipulation. In Section 4.1, I test the Hypotheses 2.1 and 2.2, and in Section 4.2, the Hypotheses 1.1-1.3. The results in the two sections should be evaluated jointly, through the lens of the model. As I further discuss in Section 5, together, they provide strong evidence for utility-driven process of belief formation.

[19] In principle, the control group should be able to arrive at the rational benchmark after an unexpected signal *unless* re-considering the model happens *only after* the signal is realized. It might be that agents do not have access to all possible models and priors beforehand—a postulate that cannot be verified in my dataset. Without the rational benchmark, it is not possible to test whether subjects' beliefs after unexpected signals show any sign of manipulation, nor whether the results of the treatment-control comparison are consistent with the benchmark-based results. That is, I cannot perform the same checks to validate the method for unexpected signals—a validation much needed, as considering different models *in addition to* different signal realizations is likely to be much more effortful and complex.

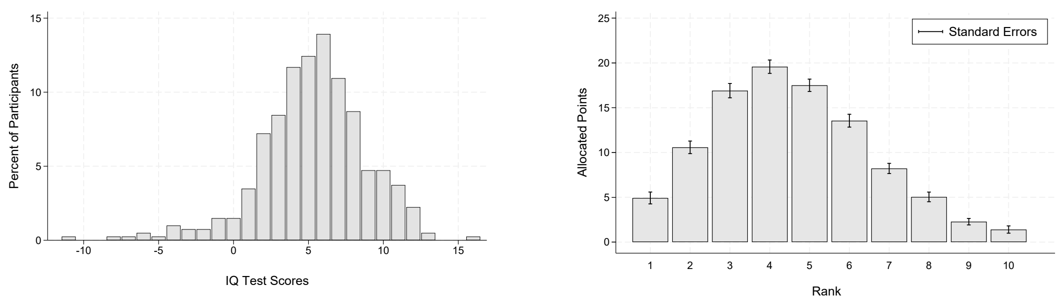
4 Results

The experiment was conducted in two waves in summer 2020 and 2023 in BonnEconLab at the University of Bonn. I collected data from 322 participants in the treatment condition and 106 participants in the control condition.[20] The experimental sessions lasted around 80 minutes and participants earned 21.4 euro on average. I report the analysis based on the data from 402 participants who made less than three mistakes in five control questions (I excluded 26 participants, that is, 6% of the sample).[21]

Rank and Prior Beliefs

First, I briefly describe the raw data. For a more detailed description, additional figures and tests, see Appendix B. Panel a) in Figure 5 presents the distribution of the IQ test scores. The distribution has a mean of 5.30 and a standard deviation of 3.58. The average rank is equal to 5.54, with a standard deviation of 2.75 (the average rank does not equal 5.5 because ranks were defined based on performance of a different group, as described in Section 2.2). Importantly, there is no significant difference between the two groups in the average IQ test score or the average rank (see Appendix B.1).

Figure 5: IQ Test Results and Prior Beliefs.



(a) Distribution of the IQ test scores.

(b) Average number of allocated points.

In order to describe subjects' prior beliefs, first, I look at the aggregate belief distribution. For every rank, I calculate the average number of points allocated by the participants. I present the averages in Panel b) in Figure 5. The distribution is skewed to the right, with the mean of 4.56 and the median of 4. On aggregate, subjects appear

[20]Women constituted 26% of the sample, with the same share of women in the treatment and in the control condition. The gender differences are gathered in Appendix D.9. While there is little difference in prior beliefs and decisions about signals with non-zero prior probability, women are more likely to believe that an unexpected signal (a signal to which they assigned zero prior probability) is their rank.

[21]The sample size was estimated based on the number of participants excluded from the analysis in the first wave (13 people, 5.8% of the sample) to match the pre-register number of 400. The results are similar if I include observations from mistaken individuals in the analysis (see Appendix D.8).

to be *overconfident*, as they put a higher probability mass on lower (better) ranks. Then, I examine individual belief distributions. The average mean equals 4.56 and is not different from the average median belief of 4.55. However, only 45 participants revealed a symmetric belief distribution. Almost half of subjects revealed a positively skewed belief distribution, and around 40%—a negatively skewed distribution. The average difference between the mean and the median in the two groups was 0.24 and -0.23, respectively. It should be noted that there is a small difference in the average beliefs in the two conditions (0.3 rank, significant at the 5%-level). In the analysis, I always control for prior beliefs or the rational benchmark that is based on subjects' priors.

Signals Received or Considered

In the treatment condition, 60% of participants (173 subjects) received a signal corresponding to a state to which they assigned non-zero prior probability. In the control condition, every participant made at least one decision about a signal of this type. I refer to the decisions about signals to which participants assigned non-zero prior probability as "the baseline sample". In an enhanced sample, I also included subjects who received a signal *adjacent* to their prior belief distribution, thus increasing the fraction of participants from the treatment group to 70% (212 subjects). In the baseline sample, 44% of subjects from the treatment received a signal between 1 and 4. This fraction is very similar in the control condition. If we defined a signal "5" as a "good" signal, the ratios would increase to 65% and 61%, respectively. The exact number of observations for each signal in the two conditions can be found in Figures 8-11 in Appendix C.1.

Decisions in the Main Task

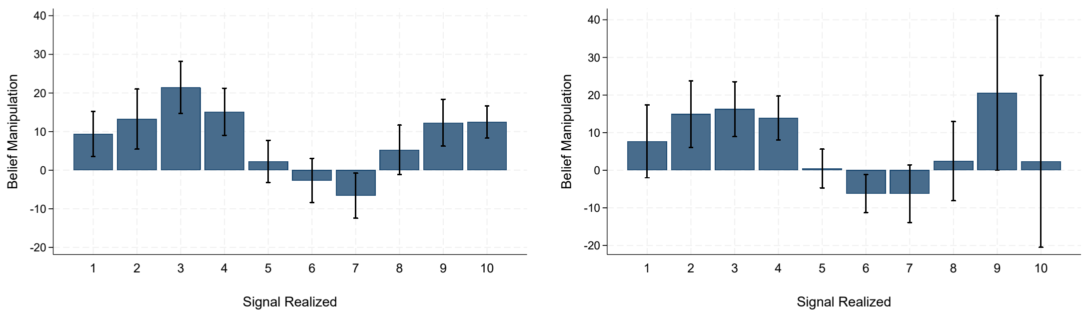
The average numbers of points allocated to Box 2 in the treatment condition is 38.5 and is 8.5 points higher than the average in the control (p-value of two-tailed t-test = 0.001). In the baseline sample, the average amounts to 56 points, 5.9 points higher than in the control condition (p-value of two-tailed t-test = 0.018). In both samples, participants tend to allocate more points after signals "1", "2", "3", or "4" in the treatment compared to the control condition (graphs presenting raw data on decisions in the two conditions are delegated to Appendix C.1). A possible explanation for the observed differences is an upward belief manipulation after a positive signal. However, to establish utility-driven belief manipulation, it is necessary to disentangle the factors driving the effect.

4.1 Data Analysis: Treatment Effect

In this section, I compare the decisions in the two conditions. Figure 6 shows how the treatment effect depends on a signal realization. The graph in Panel a) is based on a sample of all participants, and the graph in Panel b), on a sample of subjects who received a signal to which they assigned a non-zero prior probability. The bars show the difference between the average number of points allocated in the treatment and in the control condition. The whiskers represent the standard error of the difference between two means. I interpret these differences as belief manipulation after a signal. One can notice that subjects tend to manipulate their beliefs upwards after better signals: the effect is positive and significant for signals “1”, “2”, “3”, and “4”. In Panel a), there is also a positive difference after the worst signals “9” and “10”. This effect is mostly driven by participants who received a signal far from their prior belief distribution.

The observations are confirmed in a regression analysis in Table 1. The dependent variable is the number of points allocated to Box 2. In Specification (1), I regress it on a treatment dummy, a “Good Signal” variable indicating one of the best four signals, and their interaction.[22] In Specification (2), I include controls for the subject’s rank and the prior probability assigned to the rank equivalent to the signal received. The results in the first two columns (“All”) are based on a sample of all participants. In the remaining

Figure 6: Belief manipulation after a signal.



(a) All subjects. (b) Expected signals.

Note: The bars show differences between the reports in the treatment and the control condition, which I interpret as the average belief manipulation after a signal. The whiskers denote standard errors for a difference between two means. The graph in Panel a) is based on the data from all subjects, and the graph in Panel b) is based on the data from subjects who assigned non-zero prior probability to the rank corresponding to the signal received or considered.

[22]The literature provides little guidance on what constitutes a “good” signal when there are more than two states—previous studies mostly looked at beliefs about one’s performance being above or below the median. However, receiving a signal “better than median” is not equivalent to receiving a signal “5”, which rules out the possibility of being Rank 4 or better. The model predicts positive manipulation only after “good” signals—discernible in first four signals in my dataset (see Figure 6).

Table 1: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	5.778**	3.935*	0.409	0.622	1.899	0.231	8.682**	9.174***
	(2.87)	(2.11)	(3.23)	(2.73)	(3.12)	(2.47)	(3.46)	(3.36)
Good Signal	5.851*	3.067	-0.417	-1.254	-0.238	-0.830	7.979**	4.983
	(3.53)	(2.41)	(3.16)	(2.89)	(3.09)	(2.51)	(3.60)	(3.86)
Treat × Good	7.642	4.434	12.450**	11.812**	12.868**	11.565***	-14.056**	-14.102**
	(5.27)	(3.81)	(5.10)	(4.88)	(5.12)	(4.30)	(6.58)	(6.33)
Controls		✓		✓		✓		✓
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. For specifications that include different measures of the prior belief distribution, see Table 11 in Appendix D.1. The results based on alternative specifications can be found in Appendix D.2.

columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp") and a sample that also includes participants who received a signal adjacent to their prior belief distribution ("Exp+").[23] The later takes into account that individual reports might be inaccurate; in the case of a very low but positive belief, participants might be rounding their reports to zero. The results based on a sample of participants who received an "unexpected" signal—a signal indicating a rank assigned a prior of zero that is more than 1 rank away from the rank(s) assigned positive prior probability—are gathered in the last two columns.

The coefficient at the interaction term estimated using the data from all participants is positive but not significant. The remaining columns in Table 1 reveal substantial heterogeneity in how participants treat expected versus unexpected signals. In the "Exp" group, the coefficient at the interaction term is positive and significant at the 5%-level (p-value of two-tailed t-test = 0.015). Its size is substantial: 12.450 constitutes 22% of the average decision made in the treatment condition. The coefficient does not change much if we control for the subject's rank or his prior beliefs. The effects are very similar in the enhanced sample ("Exp+"). At the same time, subjects who received unexpected news

[23] A signal is considered adjacent if it is one rank lower (higher) than the first (the last) rank assigned a non-zero prior probability. For example, if a subject with prior beliefs $p = [0, 0, 0.3, 0.5, 0.2, 0, 0, 0, 0, 0]$, where i-th element denotes the probability placed on Rank i, received a signal 3, 4, or 5, he would be included in the "Exp" sample. If the signal was 2, 3, 4, 5, or 6, he would be included in "Exp+". If he received a signal between 7, 8, 9, 10, or 1, he would be included in "Unexp". The exact number of participants in each group who received a signal 1, 2, ..., 10 can be found in graphs in Appendix C.3.

allocated significantly less points after a "good" signal. This result is not at odds with belief-based utility if unexpected signals make subjects re-consider their priors—I discuss this point in more detail in Section 5. One can notice that, in the first specification, there is also a positive effect of a "good" signal in the control condition. However, the effect disappears when I add relevant controls. For a discussion of unexpected signals in the control condition, see Footnote 19 and Appendix A.

The results are robust to controlling for other measures of prior belief distribution (Appendix D.1) or choosing different regression specifications (Appendix D.2). The effect is still present, albeit weaker, if I consider the signal "5" to be a "good" signal, or use the best three signals instead of four (Appendix D.3). Moreover, the results are not driven by time effects or learning in the control condition. The treatment effect is of similar magnitude when I use either 1) the first five decisions made in the control condition or 2) the last five decisions (Appendix D.5). I also show that the results are not driven by subjects choosing the default option of 50 pp, confusing the highest and the lowest rank, or anchoring on the rational benchmark (Appendix D.6 and D.7).[24]

Crucially, the effects are not due to selection bias. In a setting with informative signals, selection might be a problem because high-ability (low-ability) subjects are more likely to get better (worse) signals. To exclude this possibility, I ran the same regressions on a subsample of subjects who received a signal from Box 1. Those subjects observed a random number—a signal unrelated to their ability. The results are gathered in Appendix D.4. The estimates are very similar, supporting the claim that the effect is due to the differential treatment of "good" versus "bad" news and does not stem from the differences in updating between more and less cognitively able individuals. An additional figure in Appendix D.4 shows that decisions after random numbers are no different from the decisions after numbers equal to one's rank regardless of signal realization.

Lastly, I address the problem of selection, as well as potential non-linear relationships between variables, using nearest neighbor matching. The results are gathered in Table 2. The dependent variable is the difference between the report in the treatment and a counterfactual constructed using the data from the control condition. The counterfactual is obtained with the nearest neighbor matching (the number of neighbors is set to 3) based on the following characteristics: the signal received (for this variable, I require the match to be exact), the prior assigned to it, the subject's rank, and their mean belief (in Appendix D.1, I consider matching participants on other measures of prior belief distributions, with similar results). I regress individual deviations from the coun-

[24]A potential issue is that people might be blindly following the rule described in the instructions on how to translate one's prior beliefs into an inference about the signal (see Section 2.3), and thus anchoring on a rational benchmark.

Table 2: Matching.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.447	11.922**	10.841**	-10.085
	(4.24)	(5.37)	(4.94)	(7.89)
Observations	301	173	212	89

The reported errors are bootstrap standard errors based on 500 replications.

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

Note: The dependent variable is the difference between the actual decision and a counterfactual constructed using the nearest neighbor matching (the number of neighbors is set to 3). To predict the counterfactual outcome, I use all observations from the control condition and the following characteristics: the subject's rank, the signal received (exact match), the prior assigned to it, and the mean of the prior belief distribution.

"Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further from their priors ("Unexp").

terfactual on a variable indicating a "good" signal, defined as previously. The coefficients at the "Good Signal" variable are qualitatively similar to the ones presented in Table 1. After an expected good signal, the effect is positive and significant—participants tend to deviate in the direction of the preferred state. I conclude that

Result 1

Subjects tend to adopt higher (more optimistic) beliefs after receiving a "good" signal, if the signal was close to their prior belief distribution. The effect is reversed after a "good" signal far from the initial beliefs.

The results provide strong support for the model for people who received a signal close to their prior beliefs. Subjects tend to interpret expected "good" signals as more informative compared to what they would say about the same signals ex-ante. At the same time, participants in the experiment tend to be skeptical of positive unexpected signals. As explained in Section 3.2 and Appendix A, this behavior does not contradict the theory. From empirical perspective, the result suggest another force limiting motivated reasoning (in addition to the costs of belief manipulation). The results also show the importance of distinguishing possible (or expected) versus unexpected information. Exploring this feature of individual feedback is a promising direction for future research.

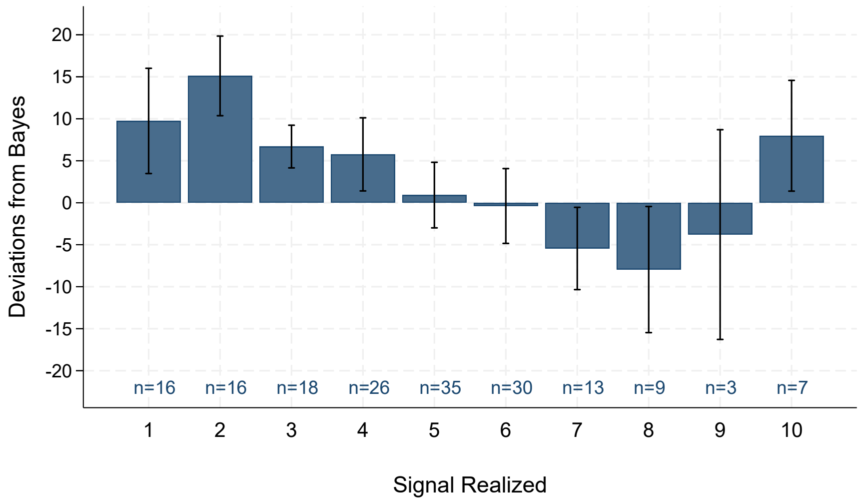
4.2 Data Analysis: The Second Approach

In this section, I adopt the second approach to examine belief manipulation after a signal. I approximate the unmanipulated belief with the Bayesian benchmark and test the model predictions separately in the treatment and in the control condition.

4.2.1 Manipulation in the treatment

Figure 7 shows how belief manipulation, defied by equations (2) and (4), depends on the signal realization. To generate the graph, I use $\beta = \alpha_1$ obtained by estimating (7). The numbers above the x-axis show how many participants received a given signal. One can notice that after better (worse) signals subjects tend to manipulate their beliefs upwards (downwards). The pattern is confirmed in a regression analysis presented in Table 3. The results in the first column is based on the entire sample ("All"), and the second column includes participants with non-zero priors assigned to the signal ("Exp"). The groups are defined as previously. For participants in the "Exp+" group who received a signal adjacent to their prior belief distribution, I replaced the null prior with the smallest positive probability feasible in the experiment (1%) and calculate the Bayesian benchmark accordingly. In every regression in Table 3, the dependent variable is the number of points allocated to Box 2. I regress it on the number of points that should be

Figure 7: Belief manipulation after a signal in the "Exp" group.



Note: The bars show the average belief manipulation after every signal realization. The averages are taken over $Y_i - \beta Y_i^{Bayes}$, where Y_i is the number of points allocated to Box 2, Y_i^{Bayes} is the number that should be allocated according to the Bayes' rule, $\beta = \alpha_1$ obtained by estimating (7) in a given group with a "good" signal defined as one of the best four signals. The graph is based on the data from subjects in the treatment condition who received a signal to which they assigned non-zero prior probability. The numbers above x-axis show their frequencies; the whiskers denote standard errors.

Table 3: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.224*	9.586***	8.797***	-6.077
	(2.78)	(3.28)	(3.00)	(6.23)
Bayes	0.697***	0.860***	0.779***	
	(0.04)	(0.09)	(0.06)	
Observations	301	173	212	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

allocated according to the Bayes' rule and a dummy indicating a "good" signal defined as previously. The coefficient at the "Good Signal" variable equals 9.586 and is significant at the 1%-level (p-value of two-tailed t-test = 0.004). The estimate does not change much when I include subjects who received a signal that was outside of their prior distribution but no further than 1 rank away (the third column in Table 3). The effect is robust to manipulations described in the previous section (see Appendix C). Importantly, it is not due to selection bias. The estimates are very similar if I restrict the sample to subjects who observed a random number.

Result 2

Subjects tend to manipulate their beliefs to a larger extent after receiving a "good" signal. The coefficient at the "Good Signal" variable is positive.

It is worth noting that Result 2 is consistent with the results obtained with the first method, providing additional support for the theory and the empirical approach.

4.2.2 Manipulation in the control

The analysis of hypothetical choices is based on 1010 decisions made by participants in the control condition. In Appendix C.4, I present the average belief manipulation depending on a signal—an equivalent of Figure 7 for the control condition. The average deviation is slightly below zero and, in contrast to the treatment condition, there is

Table 4: The effect of a "good" signal on beliefs in the control condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	2.661	-0.083	-0.293	7.979**
	(2.40)	(2.95)	(2.58)	(3.60)
Bayes	0.673***	0.869***	0.720***	
	(0.03)	(0.07)	(0.04)	
Observations	1010	483	652	358

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

Standard errors clustered at the individual level. Their values in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

no downward pattern.[25] In order to test Hypothesis 1.2, I analyze the data in the same way as the data from the treatment condition. Table 4 gathers the estimates of regressions based on the data from the control condition. The estimates reveal no significant relationship between a "good" signal and the conditional choice (except for the "Unexp" group, discussed in Footnote 19 and Appendix A). I conclude that

Result 3

There is no asymmetry after "good" nor "bad" expected signals in the control condition. The coefficient at the "Good Signal" variable is not significantly different from zero.

Additionally, I examine whether the weight placed on the Bayesian benchmark is the same in the two conditions. To this end, I compare the coefficient at the "Bayes" variable in every regression in Table 3 to the corresponding coefficient in Table 4. The results can be found in Table 9 in Appendix C.5. In every case, I cannot reject the hypothesis that the two coefficients are equal. This result substantiates the assumption that β in (4) is the same in the two conditions. It provides suggestive evidence that the estimated β captures the underlying rational process (which should be the same in the two conditions), and the difference between the two conditions is *not* due to their structure (hypothetical vs not) but stems from the utility from beliefs.

[25]Similarly to the treatment condition, the relationship breaks down in the last two signals. Yet, the direction of the deviation is *opposite* to the one in the treatment. It seems that "reconsidering" a prior is not expected in the control condition: subjects strongly believe that they would reject the signal.

Result 4

The relationship between the Bayesian benchmark and the manipulated belief, captured by the parameter β, is not significantly different in the two conditions.

4.3 Belief Elicitation II

The data from the additional belief elicitation is described in Appendix E.1. On average, subjects in the baseline sample allocated 32.86 points to the rank corresponding to the signal they received. The difference between the prior and the posterior belief amounts to 11.36 percentage points and is highly significant (p-value one-tailed t-test = 0.000). The change in beliefs varies with the signal value as expected: it is 80% higher after signals 1 to 4 than after signals 5 to 10 (p-value of one-tailed t-test = 0.0097). In Appendix E.2, I confirm that all results continue to hold if, instead of the belief about the box, I use the final belief about the rank. Additional analysis in Appendix E.3 reveals two insights. First, there is a positive correlation between decisions about the box and posterior beliefs about the rank, which remains highly significant even if I control for subjects' priors. This result shows that the reports in the main task are consistent with subjects' beliefs about the rank. Second, when I control for the decision in the main task, signal valence has little effect on the posterior belief. In other words, there is no additional asymmetry in how subjects translate decisions about the box into beliefs about the rank. This result confirms that the two-box design can be used to study the extent of asymmetric updating. Still, the final beliefs should be interpreted with caution. I discuss the caveats of eliciting beliefs multiple times in the introduction to Appendix E.

5 Discussion

There are several points that should be noted to complete the analysis. In Section 5.1, I describe the questionnaire data and evidence that supports distinguishing unexpected signals. In Section 5.2, I discuss the assumptions that underlie the identification strategy.

5.1 Questionnaires

The questionnaire was designed to explore heterogeneity in decisions made in the treatment condition. The idea is based on theoretical work of Caplin and Leahy (2001) who examine how anticipatory feelings affect decision-making. However, the goal is not to integrate my framework with their theory, but rather, to conduct an exploratory analysis on whether the feelings experienced prior to information arrival affect its interpretation.

This is reflected in the question asked: participants in both conditions were asked to report what they felt *after* learning the nature of the main task, but *before* they saw the number(s).[26,27] I present the exploratory analysis of the data in Appendix G. The evidence is correlational: I use regression analysis to investigate whether subjects' responses correlate with their decisions in the main task. I regress participants' decisions on the Bayesian benchmark, all achievement emotions, emotion-regulation strategies, and personality traits. Due to a large number of independent variables, I also use LASSO procedure to choose a subset of independent variables that are the best predictors of participants' decisions.

Emotions and Decisions

The analysis brings us two important insights. First, anticipatory emotions and the habitual use of reappraisal are the only variables correlated with participants' reports. The decisions in the main task are associated with anticipatory emotions: anxiety, hope, and hopelessness, as well as the ability to regulate one's emotions using reappraisal. At the same time, none of the BIG-5 personality traits correlates with subjects' decisions. Second, the relationships between anticipatory emotions and subjects' decisions are very different in the case of expected and unexpected signals. In Appendix G.1.1, I directly show how the relationships between anticipatory emotions and choices differ for expected and unexpected signals. The decisions after expected signals are correlated only with hope, the effect of which is positive and significant after "good" signals. The decisions after unexpected signals are strongly correlated with negative anticipatory emotions: anxiety and hopelessness, and the habitual use of reappraisal. These effects are driven by subjects' responses to unexpected "bad" signals. While experiencing stronger emotions is associated with a significantly higher belief that the "bad" signal is one's rank, the use of reappraisal has the opposite effect.

[26]The goal of the questionnaires was *not* to prove that a realized signal incites a higher emotional response than a hypothetical one. If one wish to check that, one should measure emotions after a signal realization (preferably using physiological reactions instead of questionnaires) and compare it, ideally within-subject, to a measure obtained when the agent only considers a signal. In the case of my study, this would need to be done after each hypothetical signal, and not before the task.

[27]However, one can use subjects' responses to questionnaires to check whether changes in beliefs are correlated with experienced emotions (regardless of their source, e.g., expecting a signal in treatment or facing a tedious task in control). I report the results in Appendix E.4. I confirm that differences between the first and the second belief elicitation depend on signals and are correlated with experienced emotions in the treatment but not in the control condition. The results show that the treatment manipulation shifted subjects' beliefs, and the shift was not independent from participants' emotional state.

What happens after unexpected signals?

One possible explanation is that, when an agent receives a "bad" signal that he did not expect, he might be using his negative emotions as an additional piece of information—an "internal" signal indicating that he might have performed worse than he expected. If the received signal is not far from the subject's priors, there is no need to "turn inward" for information. In the words of the theory, emotions such as anxiety or hopelessness might reflect the prior over prior $\rho(\cdot)$ and guide the agent's choice of the alternative model of the world. Although I cannot directly test this explanation, the results seem to be in line with Ortoleva (2012), especially since the same emotions play no role when updating after expected signals.[28]

5.2 Comparability and identification

The argument presented in the paper relies on several assumptions. The first assumption is that the model described in Section 3.1 is a good representation of the phenomenon at hand. As I use an off-the-shelf model of belief-based utility, I refer the reader to the previous literature (see, e.g., Engelmann et al., 2019; Brunnermeier and Parker, 2005; Caplin and Leahy, 2019). Similarly, the rationale behind using a different updating rule for unexpected signals is carefully discussed in Ortoleva (2012). The second assumption is that one can approximate the underlying rational process with decisions made in the control condition. The method is a straightforward implication of the model: without a shift in beliefs, there is no change in utility, and no incentive for additional belief manipulation. That is not to say that people do not experience belief-based utility in the control—they do, which is captured by $u(p_0)$ in (5). However, this utility should not influence their reports. There is evidence that this is indeed the case, as I document no asymmetry in the control condition. Moreover, the fact that the results are line with another, more established, benchmark (the Bayesian update) lends additional credibility to the approach. Still, hypothetical choices are an imperfect measure of beliefs that would be formed in absence of belief-based utility. There might be factors influencing decisions in the control condition that are absent in the treatment. Confounds such as time effects, anchoring, or gravitating towards a default option, are addressed in Appendix D.5 and D.6. Another potential concern, difficulty in hypothetical thinking (Esponda and Vespa, 2014), is partly addressed by 1) using the same interface as in the treatment condition and displaying each number individually to facilitate careful consideration, and 2) demonstrating that the standard deviations of subjects' decisions are no different

[28] One might wonder whether subjects who received expected and unexpected signals differ in levels of self-reported emotions. In the last table in Appendix G.1.1, I show that this is not the case.

in the two conditions, as reported in Appendix C.7 (one would expect larger errors if the decisions in the control condition were more difficult to make or subjects exerted less effort per decision). Although I cannot entirely exclude the possibility that hypothetical thinking affects the choices, there is little evidence of systematic distortions that would cast doubt on the validity of the measure.

6 Conclusions

There is mounting evidence that people derive utility not only from physical outcomes but also from their beliefs about the current or future state. The belief-based utility is likely to be the driving force behind overconfidence, the demand for (and the avoidance of) information, and belief polarization. Yet, the way it influences people's actions and beliefs is not fully understood. My study takes the next step toward explaining its role by revealing how belief-based utility shapes the way we interpret new information.

References

Bandiera, Oriana, Nidhi Parekh, Barbara Petrongolo, and Michelle Rao (2022). "Men are from Mars, and women too: a Bayesian meta-analysis of overconfidence experiments". In: *Economica* 89, S38–S70.

Barber, Brad M and Terrance Odean (2001). "Boys will be boys: Gender, overconfidence, and common stock investment". In: *The Quarterly Journal of Economics* 116.1, pp. 261–292.

Barron, Kai (2021). "Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains?" In: *Experimental Economics* 24.1, pp. 31–58.

Bénabou, Roland and Jean Tirole (2016). "Mindful Economics: The Production, Consumption, and Value of Beliefs". In: *Journal of Economic Perspectives* 30.3, pp. 141–164.

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

Brandts, Jordi and Gary Charness (2009). "The strategy method: A survey of experimental evidence". Working Paper.

Brunnermeier, Markus K and Jonathan A Parker (2005). "Optimal expectations". In: *American Economic Review* 95.4, pp. 1092–1118.

Buser, Thomas, Leonie Gerhards, and Joël Van Der Weele (2018). "Responsiveness to feedback as a personal trait". In: *Journal of Risk and Uncertainty* 56.2, pp. 165–192.

Camerer, Colin and Dan Lovallo (1999). "Overconfidence and Excess Entry: An Experimental Approach". In: *American Economic Review* 89.1, pp. 306–318.

Caplin, Andrew and John Leahy (2001). "Psychological expected utility theory and anticipatory feelings". In: *The Quarterly Journal of Economics* 116.1, pp. 55–79.

— (2019). "Wishful Thinking". Working Paper.

Chew, Soo Hong, Wei Huang, and Xiaojian Zhao (2020). "Motivated false memory". In: *Journal of Political Economy* 128.10, pp. 3913–3939.

Coutts, Alexander (2019). "Good news and bad news are still news: Experimental evidence on belief updating". In: *Experimental Economics* 22.2, pp. 369–395.

Drobner, Christoph (2022). "Motivated beliefs and anticipation of uncertainty resolution". In: *American Economic Review: Insights* 4.1, pp. 89–105.

Drobner, Christoph and Sebastian J Goerg (2024). "Motivated belief updating and rationalization of information". In: *Management Science.*

Eil, David and Justin M. Rao (2011). "The good news-bad news effect: Asymmetric processing of objective information about yourself". In: *American Economic Journal: Microeconomics* 3.2, pp. 114–138.

Engelmann, Jan, Maël Lebreton, Peter Schwardmann, Joel J van der Weele, and Li-Ang Chang (2019). "Anticipatory anxiety and wishful thinking". Working Paper.

Enke, Benjamin, Thomas Graeber, Ryan Oprea, and Jeffrey Yang (2024). *Behavioral attenuation.* Tech. rep. National Bureau of Economic Research.

Ertac, Seda (2011). "Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback". In: *Journal of Economic Behavior & Organization* 80.3, pp. 532–545.

Esponda, Ignacio and Emanuel Vespa (2014). "Hypothetical thinking and information extraction in the laboratory". In: *American Economic Journal: Microeconomics* 6.4, pp. 180–202.

Falk, Armin and Florian Zimmermann (2017). "Consistency as a signal of skills". In: *Management Science* 63.7, pp. 2197–2210.

Gerlitz, Jean-Yves and Jürgen Schupp (2005). "Zur Erhebung der Big-Five-basierten persoenlichkeitsmerkmale im SOEP". In: *DIW Research Notes* 4.

Golman, Russell, David Hagmann, and George Loewenstein (2017). "Information avoidance". In: *Journal of Economic Literature* 55.1, pp. 96–135.

Gross, James J and Oliver P John (2003). "Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being." In: *Journal of Personality and Social Psychology* 85.2, p. 348.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart (2023). "Designing information provision experiments". In: *Journal of economic literature* 61.1, pp. 3–40.

Hossain, Tanjim and Ryo Okui (2013). "The binarized scoring rule". In: *Review of Economic Studies* 80.3, pp. 984–1001.

Huffman, David, Collin Raymond, and Julia Shvets (2022). "Persistent overconfidence and biased memory: Evidence from managers". In: *American Economic Review* 112.10, pp. 3141–3175.

Lerner, Jennifer S, Ye Li, Piercarlo Valdesolo, and Karim S Kassam (2015). "Emotion and decision making". In: *The Annual Review of Psychology* 66.

Malmendier, Ulrike and Geoffrey Tate (2005). "Does overconfidence affect corporate investment? CEO overconfidence measures revisited". In: *European Financial Management* 11.5, pp. 649–659.

— (2008). "Who makes acquisitions? CEO overconfidence and the market's reaction". In: *Journal of Financial Economics* 89.1, pp. 20–43.

Möbius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat (2022). "Managing self-confidence: Theory and experimental evidence". In: *Management Science* 68.11, pp. 7793–7817.

Ortoleva, Pietro (2012). "Modeling the change of paradigm: Non-Bayesian reactions to unexpected news". In: *American Economic Review* 102.6, pp. 2410–2436.

— (2022). "Alternatives to bayesian updating". In: *Annual Review of Economics* 16.

Ortoleva, Pietro and Erik Snowberg (2015). "Overconfidence in political behavior". In: *American Economic Review* 105.2, pp. 504–35.

Pekrun, Reinhard, Thomas Goetz, Anne C Frenzel, Petra Barchfeld, and Raymond P Perry (2011). "Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ)". In: *Contemporary Educational Psychology* 36.1, pp. 36–48.

Schwardmann, Peter and Joel Van der Weele (2019). "Deception and self-deception". In: *Nature Human Behaviour* 3.10, pp. 1055–1061.

Spielberger, Charles D (1983). "State-Trait Anxiety Inventory for Adults". In: *PsycTESTS Dataset*.

Uusberg, Andero, Jamie L Taxer, Jennifer Yih, Helen Uusberg, and James J Gross (2019). "Reappraising reappraisal". In: *Emotion Review* 11.4, pp. 267–282.

Zimmermann, Florian (2020). "The dynamics of motivated beliefs". In: *American Economic Review* 110.2, pp. 337–61.

A The Model (extension)

Consider an agent who is certain about his cognitive ability being high, $p_0 = 1$, but observes a signal that he is of low ability, $s = L$. Given her prior, the Bayesian update is not defined. The agent should attribute the low signal to the noise component and remain certain of his high ability. However, there are reasons to doubt that this is a good description of human behavior. First, a body of work documents non-Bayesian reactions to unexpected news (see the discussion in Ortoleva, 2012, and Ortoleva, 2022). Second, also in my dataset, subjects who received signals outside of the support of their prior belief distribution do not update their beliefs this way (see Figure 11).

To describe belief updating after unexpected signals, I adopt the Hypothesis Testing Model (henceforth HTM) by Ortoleva (2012). The theory proposes a hierarchical beliefs structure and a decision-maker who considers different "models" of the world. A model is a probability distribution π over every possible state of the world ω. In my setting, these are the probabilities assigned to one's cognitive ability being high or low: $\pi = (p_0, 1-p_0)$. After a signal, the agent might re-consider his initial model of the world. Whether he does so depends on his threshold value ϵ. The agent updates according to his initial model if and only if $\pi(\omega) > \epsilon$, where ω denotes the state indicated by the received signal. In what follows, I assume $\epsilon = 0$. It means that, for signals within the support of the agent's prior belief distribution (± 1 rank, to account for inaccurate reporting), the agent keeps her initial model and updates in line with the Bayes' rule. Only if he receives a signal outside of his prior belief distribution, i.e., a signal corresponding to the rank assigned a zero prior probability, his behavior is described by the HTM. Specifically, after an unexpected signal, the agent looks for a new model (from a set of models he considers), one that the closest to the evidence at hand. To assess the models, he uses a "meta" prior—a prior over priors—over all possible models $\rho(\pi)$. After receiving a signal indicating ω, the agent chooses π that solves:

$$\pi^*(\omega) \equiv \arg\max_{\pi} \frac{\pi(\omega)\rho(\pi)}{\int_0^1 \pi'(\omega)\rho(d\pi')}.$$

After selecting the new model, the agent takes the prior from that model and updates it using the Bayes' rule (now well-defined, as the new prior is larger than zero).

Illustrative example

An agent considers two models of the world: π^i and π^{ii}. Each model consist of a pair of probabilities: $P(\omega = H) = p_0$ and $P(\omega = L) = 1 - p_0$. In the the first model (the initial model), the agent is certain that she is of a high type: $p_0^i = 1$. She assigns a "meta" prior probability of $\rho(\pi^i) = 1 - \delta$ to this model, with $\delta \in (0, \frac{1}{2})$. In the second model, there is a non-zero probability ζ that the agent is of a low type: $p_0^{ii} = 1 - \zeta$. The agent assigns a prior probability of $\rho(\pi^{ii}) = \delta$ to this model. Let us assume that the threshold value of re-considering the initial model is $\epsilon = 0$. The agent receives a signal $s = L$. According to the first model, $P(\omega = L) = 1 - p_0 = 0$. This probability is no higher than the threshold value, so the agent looks for an alternative explanation of the evidence. Specifically, she compares the likelihood of the initial model to the alternative, given the signal $s = L$, and chooses the model that maximizes:

$$\pi^* = \arg\max \left\{ \frac{0 \times (1-\delta)}{0 \times (1-\delta) + \zeta\delta}, \frac{\zeta\delta}{0 \times (1-\delta) + \zeta\delta} \right\}. \tag{10}$$

The alternative model is more likely given the signal $s = L$, so the agent abandons her initial model and switches to $\pi^{ii} = (1 - \zeta, \zeta)$. Using the new prior $1 - \zeta$, she updates her beliefs using the Bayes' rule. Assuming the signal precision $c \in (\frac{1}{2}, 1)$, the posterior is:

$$p_{1,s=L}^{HTM} = \frac{c(1-\zeta)}{c(1-\zeta) + (1-c)\zeta}. \tag{11}$$

Without the HTM extension, the Bayesian posterior would be $p_{1,s=L}^{B} = p_0 = 1$.

Implications

It is important to note that $p_{1,s=L}^{HTM} \leq p_{1,s=L}^{B}$, meaning that agents faced with unexpected signals are always more pessimistic under the HTM. The equality holds if and only if $\zeta = 0$, that is, if the agent does not entertain any alternative model of the world. Consequently, if agents behaved in line with the Bayesian model, we should never observe $p_{1,s=L} < p_0$ after an unexpected signal. Yet, it can happen as a result of the two-step process in which the signal leads to reconsideration of the model and, only afterwards, updating the new prior (which is lower than the former prior belief).

Belief-based utility

The Hypothesis Testing Model can be augmented in a straight-forward manner with the belief-based utility model described in Section 3.1. Again, the agent derives utility from a belief that he is of a high type, $\tilde{p}_{1,s}$. Recall that the optimal chosen belief is given by the first-order condition (2) reproduced here for convenience:

$$\gamma u'(\tilde{p}_{1,s}) = \tilde{p}_{1,s} - p_{1,s},$$

where $p_{1,s}$ is the unmanipulated (or rational) belief and γ captures the cognitive costs of belief manipulation. In the main text, I approximate the rational belief with the Bayesian posterior: $p_{1,s} = \beta p^B_{1,s}$, where $\beta > 0$ captures the degree to which the process follows Bayesian updating. After an unexpected signal, however, the unmanipulated belief changes according to the HTM model. The agent revises her model of the world and forms an unmanipulated belief $p_{1,s} = p^{HTM}_{1,s}$ based on the new prior. Then, she solves the same optimization problem of choosing a new, utility-maximizing posterior. The first order condition reads: $\gamma u'(\tilde{p}_{1,s}) = \tilde{p}_{1,s} - p^{HTM}_{1,s}$.

In the experimental data, we only observe p_0 and $\tilde{p}_{1,s}$. The former allows us to calculate the Bayesian posterior $p^B_{1,s}$. Unfortunately, the reported prior p_0 is only informative about the initial model of the world. The prior used after model revision is not observed. Without knowing the set of considered models and the meta-prior, we cannot distinguish between the rational posterior, $p^{HTM}_{1,s}$, and the bite of the belief based utility, $\gamma u'(\cdot)$.

Could one use the data from the control condition? Not really, if re-considering the model happens only after a signal is realized. It is not clear that people have access to all possible models and priors beforehand, and I cannot verify it in the data. Without the rational benchmark, I cannot test whether decisions in the control condition show any sign of manipulation, nor whether the two approaches are consistent. That is, I cannot run the same tests to validate the method for unexpected signals. At the same time, the validation seems to be even more important, because considering different models *in addition to* different signal realizations is likely to be much more effortful and complex. As a result, the reported belief might not fully reflect the unmanipulated posterior.

B Descriptive Statistics

B.1 Differences between participants in the two conditions

Table 5: Differences between participants in the two conditions.

	Treatment	**Control**		Diff < 0	Diff $\neq 0$	Diff > 0
IQ test score	5.349	5.168	*p-value:*	0.669	0.661	0.331
Rank	5.442	5.832	*p-value:*	0.109	0.218	0.891
Measures of Belief Distribution:						
Mean Belief	4.463	4.817	*p-value:*	0.038	0.076	0.962
1^{st} Quartile	3.738	4.069	*p-value:*	0.045	0.090	0.955
Median Belief	4.481	4.812	*p-value:*	0.044	0.088	0.956
3^{st} Quartile	5.193	5.530	*p-value:*	0.050	0.100	0.950
Range	5.096	4.782	*p-value:*	0.041	0.083	0.958
N	301	101				

Table 6: Differences between participants in the two conditions.

	Treatment	**Control**		Diff < 0	Diff $\neq 0$	Diff > 0
BIG-5: Extr	14.10	13.90	*p-value:*	0.663	0.674	0.337
BIG-5: Cons	14.10	14.06	*p-value:*	0.544	0.913	0.456
BIG-5: Open	14.67	13.97	*p-value:*	0.956	0.087	0.044
BIG-5: Neur	12.65	13.35	*p-value:*	0.070	0.141	0.930
BIG-5: Agree	15.29	15.31	*p-value:*	0.479	0.958	0.521
STAI: State	58.98	58.10	*p-value:*	0.777	0.446	0.223
STAI: Trait	58.34	57.47	*p-value:*	0.763	0.473	0.237
N	301	101				

B.2 Decisions in the two conditions

Table 7: Beliefs and decisions about signals with non-zero prior probability.

	Treatment	**Control**		Diff < 0	Diff ≠ 0	Diff > 0
Prior Beliefs*	21.502	20.910	*p-value:*	0.669	0.661	0.331
	(1.093)	(0.641)				
*Probability placed in Belief Elicitation I on the signal.						
Main Task**	56.121	50.242	*p-value:*	0.991	0.018	0.009
	(2.043)	(1.284)				
**Number of points allocated to Box 2 in the main task.						
Bayes***	61.291	60.325	*p-value:*	0.725	0.549	0.275
	(1.362)	(0.832)				
***Points that should allocated according to Bayes' rule.						
N	173	483				

Note: Standard errors in parentheses.

Table 8: Beliefs and decisions about signals with zero prior probability.

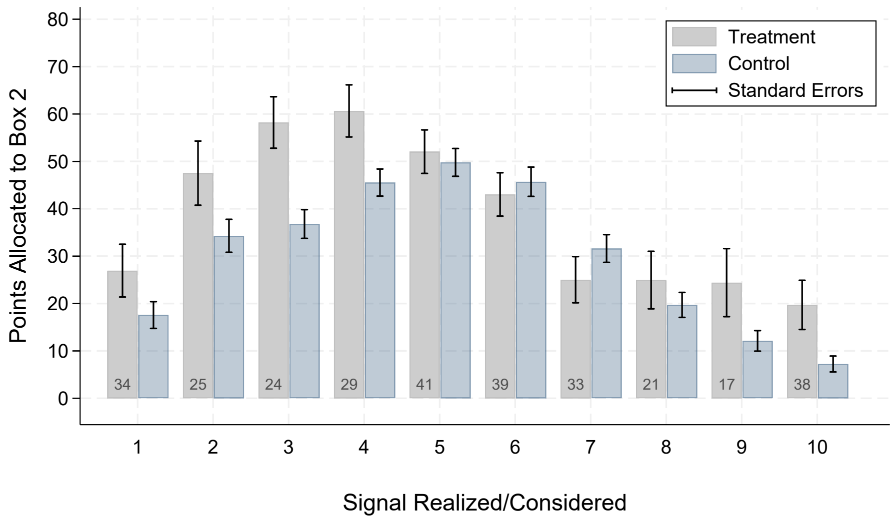
	Treatment	**Control**		Diff < 0	Diff ≠ 0	Diff > 0
Main Task*	14.664	11.512	*p-value:*	0.927	0.146	0.073
	(2.157)	(0.930)				
*Number of points allocated to Box 2 in the main task.						
N	128	527				

Note: Standard errors in parentheses.

C Additional Results

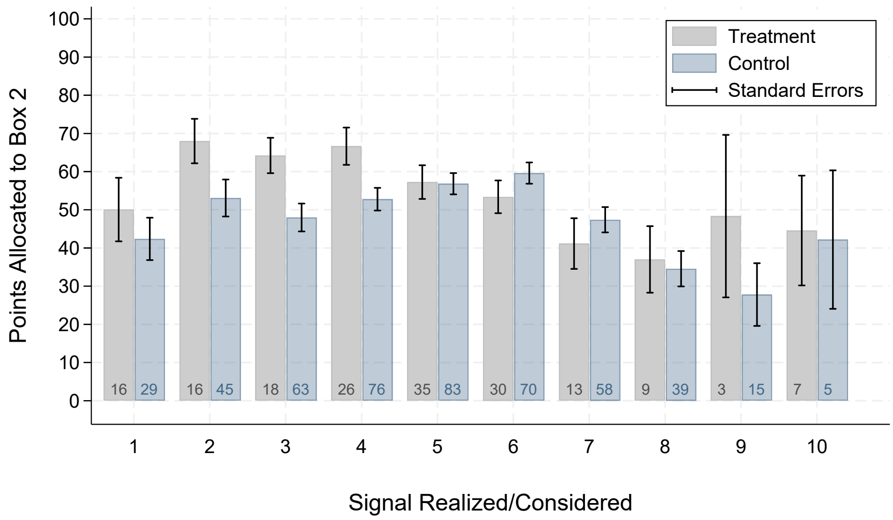
C.1 Decisions in the two conditions

Figure 8: Decisions in the main task in the two condition (all participants).



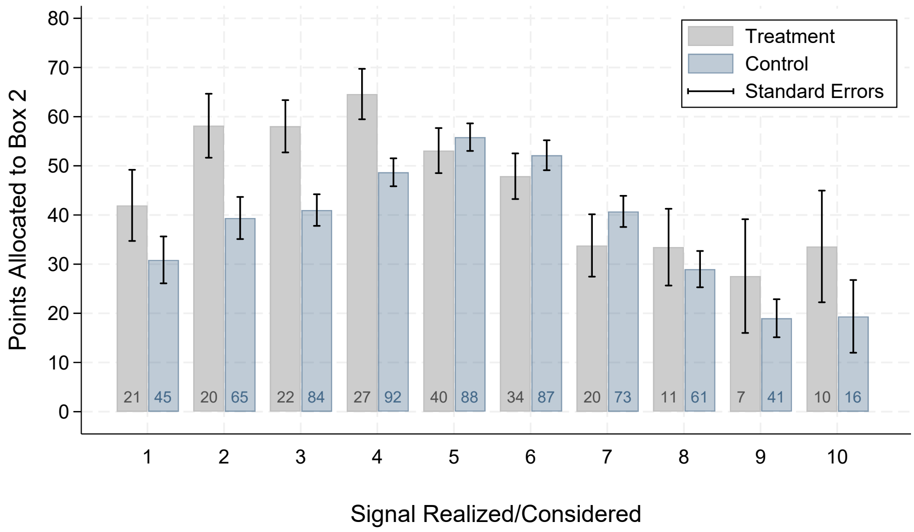
Note: Numbers within the gray bars denote the number of observations in the treatment condition. In the control condition, there are 101 observations for every signal (omitted for clarity).

Figure 9: Decisions in the main task in the "Exp" group.



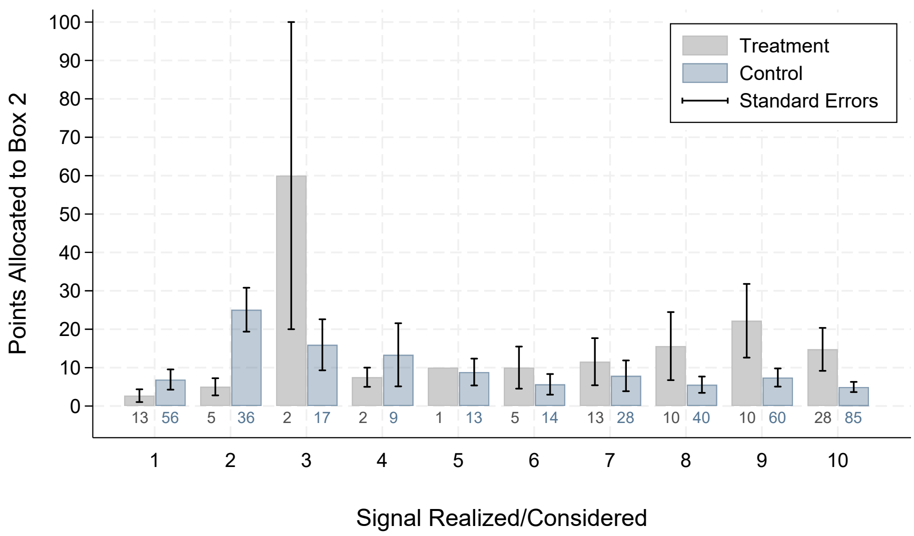
Note: Numbers within the gray bars denote the number of observations in the treatment and the control condition. The results are based on the data from subjects who received/considered a signal to which they assigned a non-zero prior probability.

Figure 10: Decisions in the main task in the "Exp+" group.



Note: Numbers within the gray bars denote the number of observations in the treatment and the control condition. The results are based on the data from subjects who received/considered a signal to which they assigned non-zero prior probability or was adjacent to their prior belief distribution.

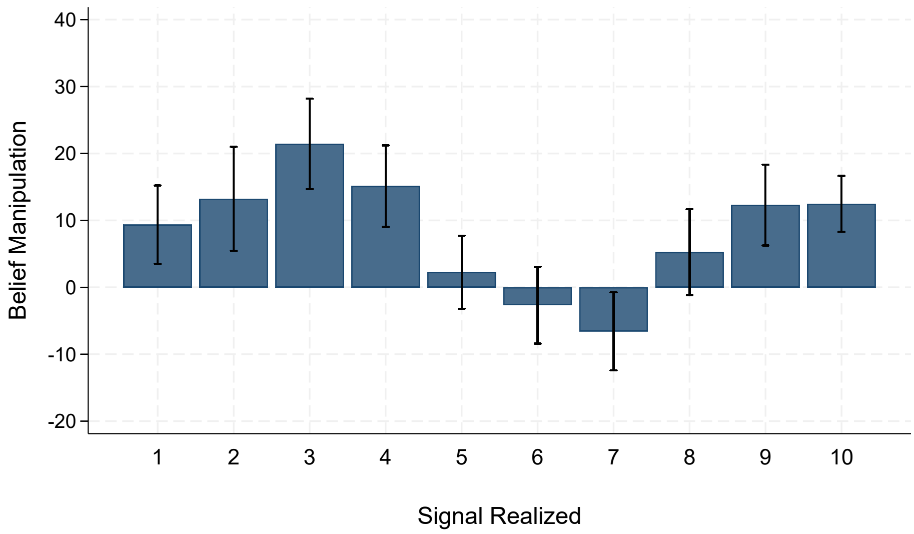
Figure 11: Decisions in the main task in the "Unexp" group.



Note: Numbers within the gray bars denote the number of observations in the treatment and the control condition. The results are based on the data from subjects who received a signal to which they assigned a zero prior probability and was further than 1 rank from their prior belief distribution.

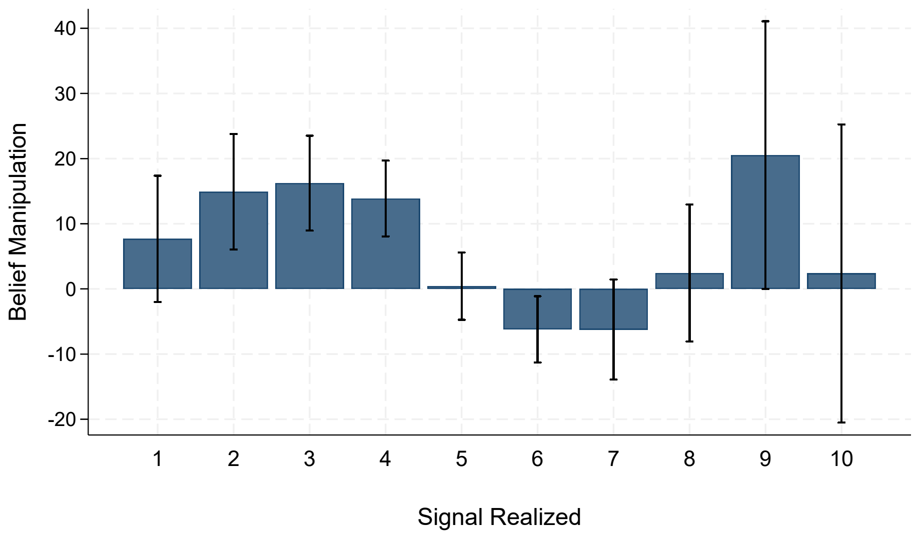
C.2 Treatment effect in each group

Figure 12: Belief manipulation after a signal (all participants).



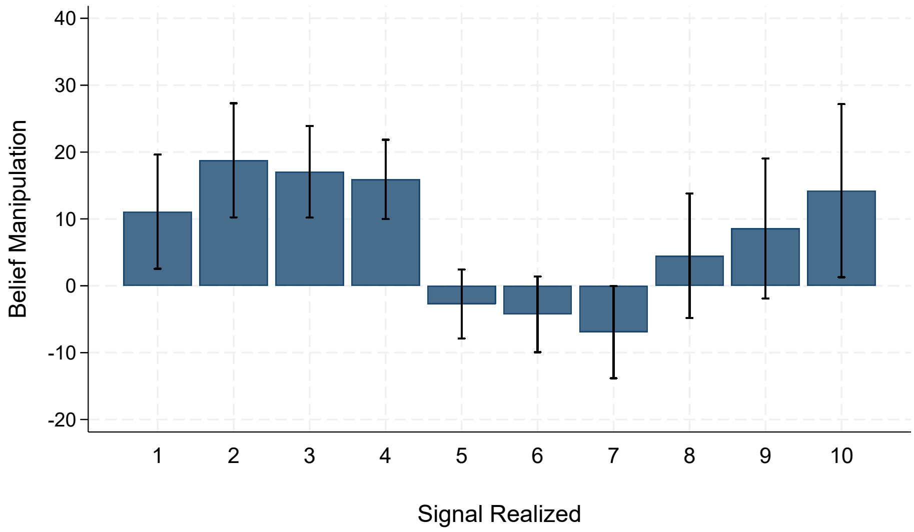
Note: The bars show differences between the reports in the treatment and the control condition, which I interpret as the average belief manipulation after a signal. The whiskers denote standard errors for a difference between two means. The graph is based on the data from all subjects.

Figure 13: Belief manipulation after a signal in the "Exp" group.



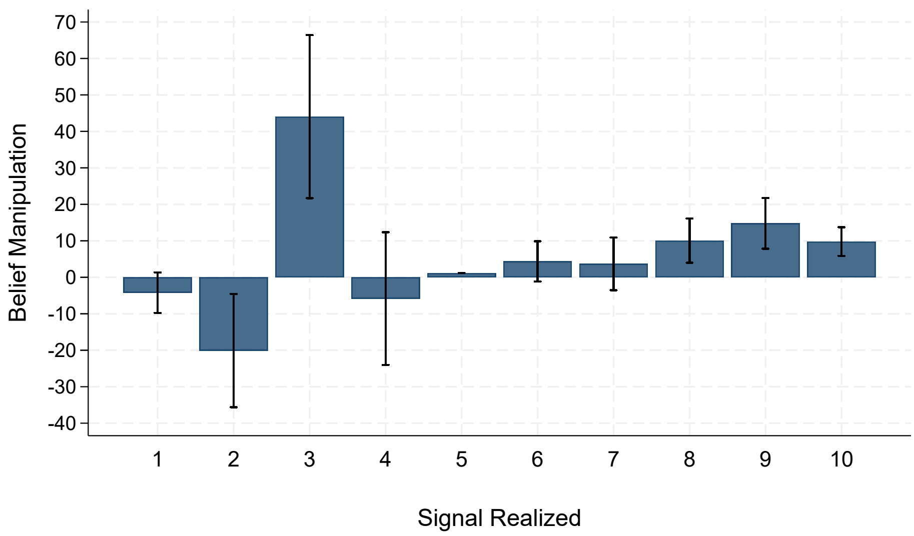
Note: The bars show differences between the reports in the treatment and the control condition, which I interpret as the average belief manipulation after a signal. The whiskers denote standard errors for a difference between two means. The graph is based on the data from subjects who assigned non-zero prior probability to the received signal.

Figure 14: Belief manipulation after a signal in the "Exp+" group.



Note: The bars show differences between the reports in the treatment and the control condition, which I interpret as the average belief manipulation after a signal. The whiskers denote standard errors for a difference between two means. The graph is based on the data from subjects who assigned non-zero prior probability to the received signal, or received a signal not far from their priors (no more than 1 rank away from the beginning or the end of individual belief distribution).

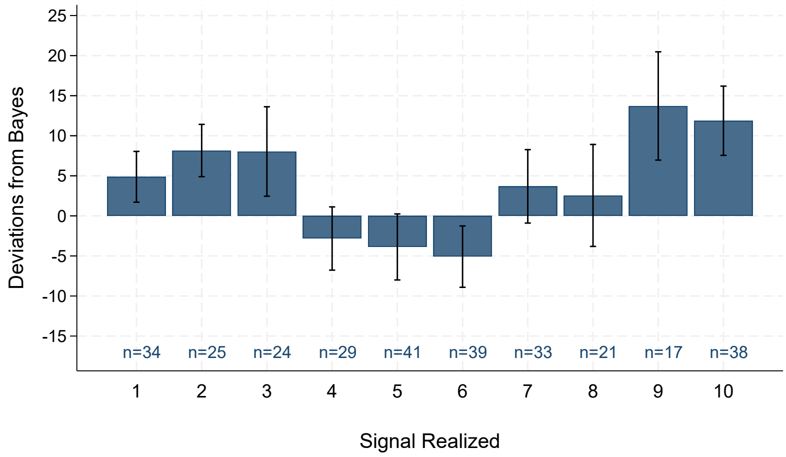
Figure 15: Belief manipulation after a signal in the "Unexp" group.



Note: The bars show differences between the reports in the treatment and the control condition, which I interpret as the average belief manipulation after a signal. The whiskers denote standard errors for a difference between two means. The graph is based on the data from subjects who received a signal far from their priors (more than 1 rank away from the beginning or the end of individual belief distribution).

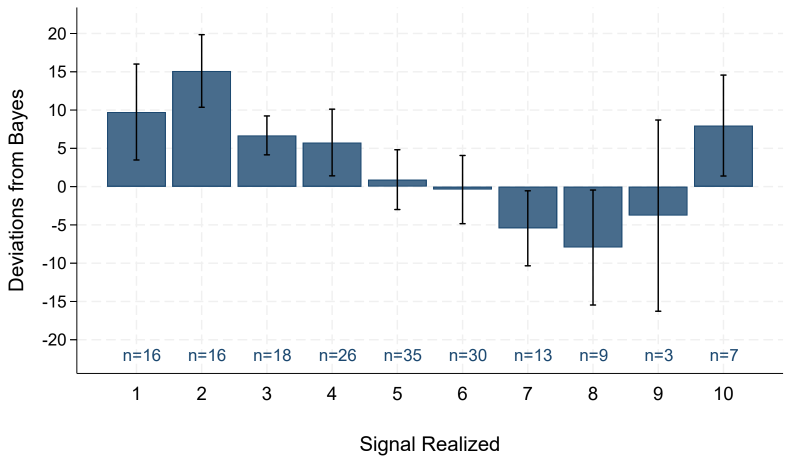
C.3 Deviations from Bayes in each group

Figure 16: Belief manipulation after a signal (all participants).



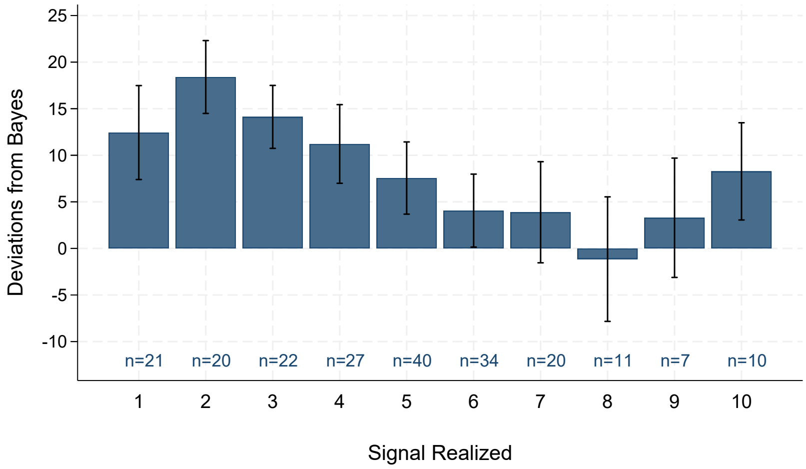
Note: The bars show the average belief manipulation after every signal realization. The averages are taken over $Y_i - \beta Y_i^{Bayes}$, where Y_i is the number of points allocated to Box 2, Y_i^{Bayes} is the number that should be allocated according to the Bayes' rule (for participants with a prior equal to zero, it is assumed to be zero as well). β is assumed to be one in order to give an equal weight to participants with a zero and non-zero prior (otherwise, we would be downweighting the deviations of people with non-zero priors). The graph is based on the data from all subjects in the treatment condition. The numbers above x-axis show the frequencies; the whiskers denote standard errors.

Figure 17: Belief manipulation after a signal in the "Exp" group.



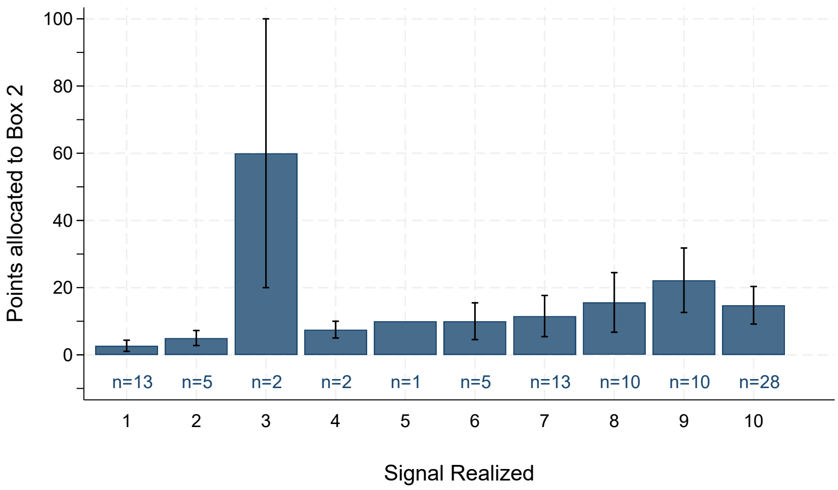
Note: The bars show the average belief manipulation after every signal realization. The averages are taken over $Y_i - \beta Y_i^{Bayes}$, where Y_i is the number of points allocated to Box 2, Y_i^{Bayes} is the number that should be allocated according to the Bayes' rule, $\beta = \alpha_1$ obtained by estimating (7) in a given group with a "good" signal defined as one of the best four signals. The graph is based on the data from subjects in the treatment condition who received a signal to which they assigned non-zero prior probability. The numbers above x-axis show their frequencies; the whiskers denote standard errors.

Figure 18: Belief manipulation after a signal in the "Exp+" group.



Note: The bars show the average belief manipulation after every signal realization. The averages are taken over $Y_i - \beta Y_i^{Bayes}$, where Y_i is the number of points allocated to Box 2, Y_i^{Bayes} is the number that should be allocated according to the Bayes' rule, $\beta = \alpha_1$ obtained by estimating (7) in a given group with a "good" signal defined as one of the best four signals. The graph is based on the data from subjects in the treatment condition who received a signal to which they assigned non-zero prior probability or was adjacent to their prior belief distribution. In the later case, I replaced the zero prior with 1% and calculated the Bayesian benchmark as in the main text. The numbers above x-axis show their frequencies; the whiskers denote standard errors.

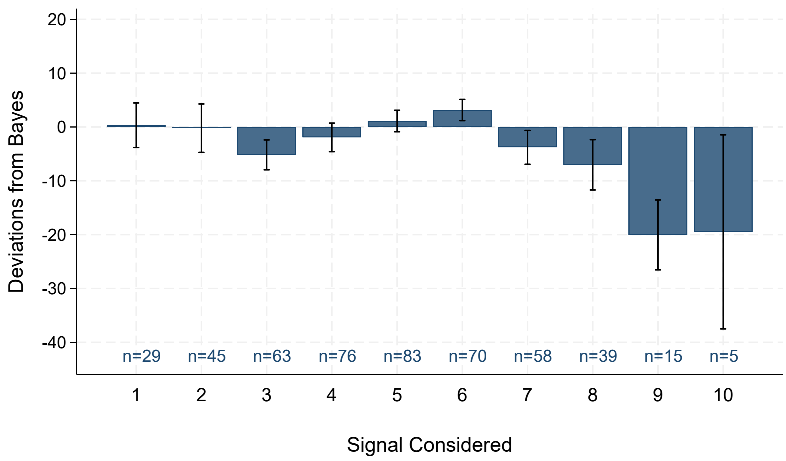
Figure 19: Belief manipulation after a signal in the "Unexp" group.



Note: The bars show the average belief manipulation after every signal realization. They are based on the data from subjects in the treatment condition who received a signal to which they assigned zero prior probability and was further than 1 rank from their prior belief distribution. The averages are taken over Y_i (note that, in this case, $Y_i^{Bayes} = 0$). The numbers above x-axis show their frequencies; the whiskers denote standard errors.

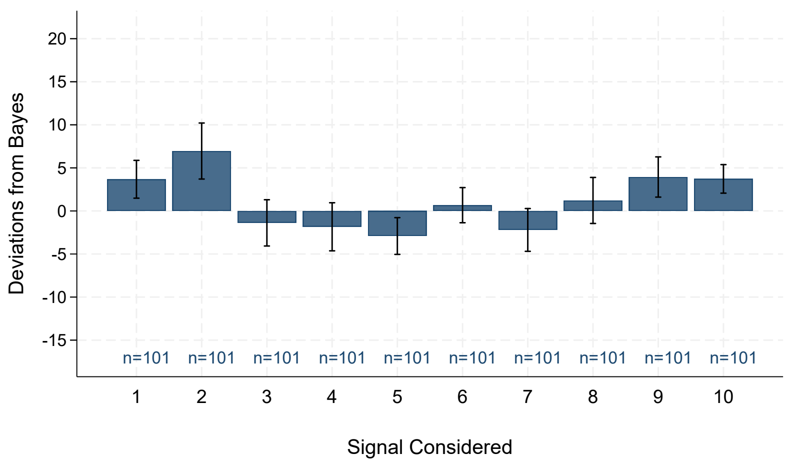
C.4 Belief manipulation in the control condition

Figure 20: Belief manipulation in the control condition ("Exp").



Note: The bars show the average belief manipulation for every signal realization. The averages are taken over $Y_i - \beta Y_i^{Bayes}$, where Y_i is the number of points allocated to Box 2, Y_i^{Bayes} is the number that should be allocated according to the Bayes' rule, $\beta = \alpha_1$ obtained by estimating (7) in a given group with a "good" signal defined as one of the best four signals. The graph is based on the data from subjects in the control condition who considered a signal to which they assigned non-zero prior probability. The numbers above x-axis show their frequencies; the whiskers denote standard errors.

Figure 21: Belief manipulation in the control condition (all signals).



Note: The bars show the average belief manipulation after every signal realization. The averages are taken over $Y_i - \beta Y_i^{Bayes}$, where Y_i is the number of points allocated to Box 2, Y_i^{Bayes} is the number that should be allocated according to the Bayes' rule (for participants with a prior equal to zero, it is assumed to be zero as well). β is assumed to be one in order to give an equal weight to participants with a zero and non-zero prior (otherwise, we would be downweighting the deviations of people with non-zero priors). The graph is based on the data from all subjects in the control condition. The numbers above x-axis show the frequencies; the whiskers denote standard errors.

C.5 Testing regression coefficients

Table 9: Testing equality of coefficients at the "Bayes" variable across the conditions.

H_0: α_1 at Y^{Bayes} in Table 3 is equal to α_1 at Y^{Bayes} in Table 4			
	All	*Exp*	*Exp+*
Prob > chi2	0.640	0.936	0.394
	(0.03)	(0.07)	(0.04)
	1311	656	864

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

Standard errors clustered at the individual level. Their values in parentheses.

C.6 Heterogeneity in confidence

I analyzed the data from the following recent experiments on asymmetric updating: Drobner (2022), Drobner and Goerg (2024), Möbius et al. (2022), and Zimmermann (2020), as well as this study. The criteria used to classify a subject as over- or underconfident were rather conservative, as I allowed for some inaccuracy in beliefs, thus, characterizing a considerable fraction of subjects as unbiased. For experiments with 10 ranks, I define a person as overconfident (underconfident) if the mean belief is higher than the actual rank, with subjects reporting beliefs within the ± 0.5 rank interval from their actual rank classified as unbiased. This classification reveals 48.5% of participants to be overconfident, 41.5% underconfident, and 10% unbiased in Zimmermann (2020), whereas in the current paper, the corresponding fractions are 56%, 31%, and 13%. Using the data from Möbius et al. (2022), I defined a subjects as overconfident (underconfident) if his belief of being in the top 50% of test-takers was higher or equal (lower or equal) than their actual percentile placement in the distribution plus (minus) 5 percentage points. I obtained the following ratios: 56% overconfident, 32% underconfident, and 12% unbiased subjects. A similar definition was adopted to analyze the data from Drobner (2022), however, in this dataset, I had to calculate the placement in the distribution of all test-takers myself. The resulting fractions are 44.6% overconfident, 44.6% underconfident, and 11% unbiased participants. Lastly, Drobner and Goerg (2024) used a different structure, with 4 ranks, and probabilistic belief on being assigned to each rank. I used the same definition as for the 10 ranks, with a proportional margin to account for inaccurate reporting (± 0.2 rank instead of ± 0.5 rank). The results suggest 45.5% overconfident, 44% underconfident, and 10.5% unbiased subjects in the sample.

C.7 Comparing standard deviations

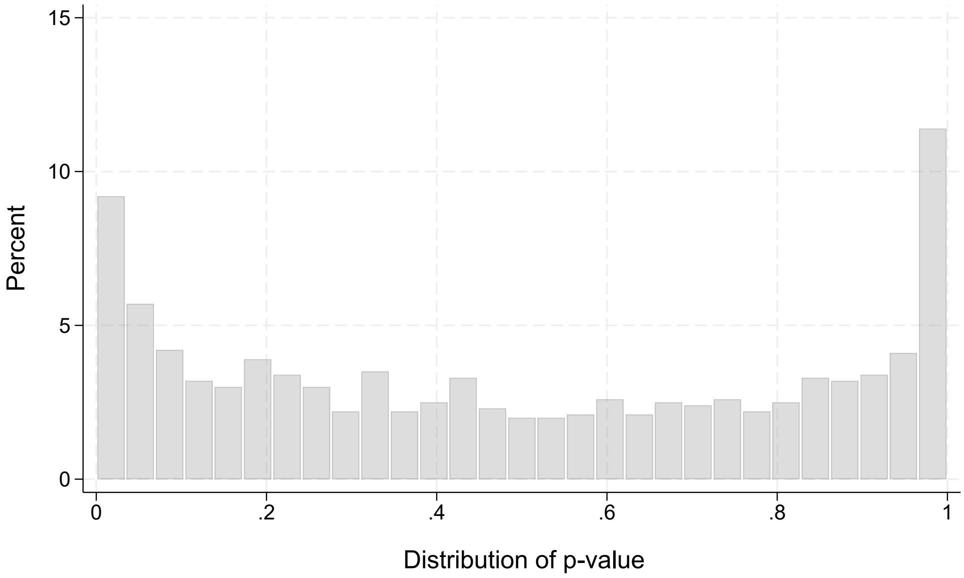
In this section, I argue that participants in the control condition do not make more mistakes than subjects in the treatment. To this end, I compare subjects' decisions to the Bayesian benchmark and show that they do not deviate more from it in the control condition. In Table 10, I compare the standard deviations in the two conditions and show that there is no significant difference between the two. However, the observations in the control condition are not independent, and a proper test should take this into account. I solve the problem as follows. I randomly draw 1 decision for each participant in the control condition (100 decisions), and a decision from one out of every three participants in the treatment (100 decisions). I compare the standard deviations between these samples and save the p-value of the F-test (with the null hypothesis that the ratio of the standard deviations from the two samples is equal to one, and the alternative hypothesis that the ratio is less than one). I repeated the procedure 1000 times, thereby obtaining the distribution of p-values presented in Figure 22. The results show that we can reject the null hypothesis only in a very few cases.

Table 10: Standard deviations in the two conditions.

	Treatment	**Control**		Ratio < 1	Ratio $\neq 1$	Ratio > 1
SD	24.944	24.830	*p-value:*	0.546	0.908	0.454
*Standard deviation of points allocated to Box 2 minus the Bayesian benchmark.						
N	301	1010				

Note: "Ratio" denotes the ratio of standard deviations in the two conditions.
The conclusions do not change if I use subjects' decisions instead of their deviations from benchmark.

Figure 22: Standard deviations of decisions in the two conditions.



Note: The graph shows p-values of F-test that the standard deviations in the two conditions are equal under the alternative hypothesis that the standard deviation in the treatment condition is higher. In order to calculate SD in comparable samples, I randomly drew 1 decision for each participant in the control condition (100 decisions), and a decision from one out of every three participants in the treatment (100 decisions). I compare the standard deviations between the samples and save the p-value of the F-test. I repeated the procedure 1000 times, thereby obtaining the distribution of p-values. The conclusions do not change if I use subjects' decisions instead of their deviations from the Bayesian benchmark.

D Robustness checks

In this section, I present the same tables as in the main body (Tables 1-4) with changes that allow for checking robustness, which I describe in the table notes. Some tables are omitted, as they do not add much to the argument, or are presented elsewhere. For example, the equivalent of Table 3 is omitted in Appendix D.7 because I only exclude participants from the control condition—it has no impact on the results based on the data from treatment; Tables 3 and 4 are omitted from Appendix D.2, as I find it unnecessary to present specifications that control for both the Bayesian benchmark and the prior.

D.1 Controlling for the entire belief distribution

Table 11: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	3.892*	4.237**	0.295	1.077	0.015	1.122	8.975***	9.868***
	(2.11)	(2.14)	(2.73)	(2.98)	(2.46)	(2.63)	(3.33)	(3.33)
Good Signal	3.065	3.055	-1.348	-0.545	-0.782	-0.354	5.158	3.612
	(2.41)	(2.41)	(2.88)	(2.93)	(2.51)	(2.53)	(3.93)	(4.00)
Treat × Good	4.552	4.152	12.422**	10.824**	11.967***	10.193**	-13.220**	-14.688**
	(3.80)	(3.83)	(4.96)	(5.05)	(4.30)	(4.33)	(6.24)	(6.69)
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. In both specifications, the controls include: the prior assigned to the rank = signal, and the subject's rank. Additionally, in Specification (1), I control for the following measures of individual belief distribution: the 1^{st} quartile, the median, and the 3^{rd} quartile. In Specification (2), I control for the number of points allocated to each of the 10 ranks.

Table 12: Matching.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	7.074*	14.167**	12.201**	-9.263
	(4.22)	(5.58)	(5.03)	(6.57)
Observations	301	173	212	89

The reported errors are bootstrap standard errors based on 500 replications.

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

Note: The dependent variable is the difference between the actual decision and a counterfactual constructed using the nearest neighbor matching (the number of neighbors is set to 3). To predict the counterfactual outcome, I use all observations from the control condition and the following characteristics: the signal received (exact match), the prior assigned to it, the subject's rank, and the following measures of individual belief distribution: the 1^{st} quartile, the median, and the 3^{rd} quartile.

Table 13: The effect of signal valence in the treatment condition.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Good Signal	5.142*	3.611	9.743**	8.342	8.832**	9.879**	-5.015	-5.788
	(2.84)	(3.24)	(3.95)	(5.07)	(3.49)	(4.31)	(8.56)	(8.91)
Bayes	0.710***	0.705***	0.937***	0.987***	0.798***	0.809***		
	(0.04)	(0.04)	(0.10)	(0.10)	(0.06)	(0.06)		
Observations	301	301	173	173	212	212	89	89

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. In both specifications, the controls include: the prior assigned to the rank = signal, and the subject's rank. Additionally, in Specification (1), I control for the following measures of individual belief distribution: the 1^{st} quartile, the median, and the 3^{rd} quartile. In Specification (2), I control for the number of points allocated to each of the 10 ranks.

Table 14: The effect of signal valence in the control condition.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Good Signal	2.636	2.638	-0.911	-0.996	-0.368	-0.384	4.734	4.786
	(2.41)	(2.42)	(3.27)	(3.29)	(2.72)	(2.73)	(4.01)	(4.18)
Bayes	0.678***	0.678***	0.878***	0.875***	0.711***	0.710***		
	(0.03)	(0.03)	(0.08)	(0.09)	(0.04)	(0.04)		
Observations	1010	1010	483	483	652	652	358	358

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. In both specifications, the controls include: the prior assigned to the rank = signal, and the subject's rank. Additionally, in Specification (1), I control for the following measures of individual belief distribution: the 1^{st} quartile, the median, and the 3^{rd} quartile. In Specification (2), I control for the number of points allocated to each of the 10 ranks.

D.2 Alternative specifications

Table 15: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	3.916*	3.767*	0.698	0.692	0.680	0.258	8.682**	8.790**
	(2.14)	(2.15)	(2.76)	(2.78)	(2.51)	(2.50)	(3.46)	(3.39)
Good Signal	3.068	3.064	-0.807	-0.843	-0.766	-0.684	7.979**	7.644**
	(2.40)	(2.40)	(2.89)	(2.87)	(2.55)	(2.52)	(3.60)	(3.52)
Treat × Good	3.098	4.298	10.367**	11.822**	9.651**	11.587***	-14.056**	-13.688**
	(3.67)	(3.80)	(4.40)	(4.87)	(4.00)	(4.27)	(6.58)	(6.40)
Controls		✓		✓		✓		✓
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). In Specification (1), the controls include only the prior assigned to the rank corresponding to the signal received. In Specification (2), the aforementioned prior and the subject's rank.

Table 16: Matching.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	4.333	10.703**	9.018**	-10.173
	(3.58)	(4.95)	(4.29)	(5.53)
Observations	301	173	212	89

The reported errors are bootstrap standard errors based on 500 replications.

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

Note: The dependent variable is the difference between the actual decision and a counterfactual constructed using the nearest neighbor matching (the number of neighbors is set to 3). To predict the counterfactual outcome, I use all observations from the control condition and the following characteristics: the signal received (exact match) and the prior assigned to it.

"Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further from their priors ("Unexp").

Table 17: Matching.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.226	10.190*	8.943*	-8.708
	(4.43)	(5.50)	(4.93)	(7.23)
Observations	301	173	212	89

The reported errors are bootstrap standard errors based on 500 replications.

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

Note: The dependent variable is the difference between the actual decision and a counterfactual constructed using the nearest neighbor matching (the number of neighbors is set to 3). To predict the counterfactual outcome, I use all observations from the control condition and the following characteristics: the signal received (exact match), the prior assigned to it, and the participant's rank.

"Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further from their priors ("Unexp").

D.3 Alternative definition of a "good" signal

Table 18: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	6.845**	3.697*	3.237	1.782	4.301	0.895	8.211**	8.652***
	(2.65)	(1.93)	(2.89)	(2.53)	(2.81)	(2.28)	(3.38)	(3.27)
Good Signal 1-3	-0.694	0.295	-2.468	-3.997	-5.034	-4.576*	7.798**	4.909
	(3.62)	(2.37)	(3.59)	(3.35)	(3.44)	(2.67)	(3.45)	(3.82)
Treat × Good 1-3	5.788	6.568*	9.189	13.663***	10.305*	13.943***	-13.523**	-13.410**
	(5.62)	(3.94)	(5.71)	(5.08)	(5.64)	(4.32)	(6.78)	(6.76)
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable in Specifications (1) and (2) is the number of points allocated to Box 2. "Good Signal 1-3" is an indicator variable that takes the value 1 if a subject received one of the best three signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank.

Table 19: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal 1-3	4.922*	9.928***	8.916***	-5.725
	(2.96)	(3.60)	(3.25)	(6.44)
Bayes	0.710***	0.904***	0.808***	
	(0.04)	(0.09)	(0.06)	
Observations	301	173	212	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal 1-3" is an indicator variable taking value 1 if a subject received one of the best three signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 20: The effect of a "good" signal on beliefs in the control condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal 1-3	1.612	-0.291	-1.925	7.798**
	(2.36)	(3.27)	(2.73)	(3.46)
Bayes	0.677***	0.869***	0.718***	
	(0.03)	(0.07)	(0.04)	
Observations	1010	483	652	358

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

Standard errors clustered at the individual level. Their values in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal 1-3" is an indicator variable taking value 1 if a subject received one of the best three signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 21: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	5.049	4.847**	-0.373	1.880	1.745	1.212	8.905**	9.338***
	(3.08)	(2.34)	(3.99)	(3.44)	(3.71)	(2.92)	(3.52)	(3.43)
Good Signal 1-5	13.517***	5.931**	4.339	3.027	7.588***	3.552	7.592**	4.734
	(3.10)	(2.36)	(3.00)	(2.66)	(2.77)	(2.47)	(3.40)	(3.34)
Treat × Good 1-5	6.486	1.462	9.550*	6.448	8.675*	6.673	-13.694**	-13.745**
	(4.77)	(3.70)	(5.11)	(4.88)	(4.94)	(4.28)	(6.37)	(6.12)
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable in Specifications (1) and (2) is the number of points allocated to Box 2. "Good Signal 1-5" is an indicator variable that takes the value 1 if a subject received one of the best five signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank.

Table 22: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal 1-5	4.637	8.199**	8.269***	-6.102
	(2.81)	(3.47)	(3.11)	(6.13)
Bayes	0.687***	0.841***	0.765***	
	(0.04)	(0.09)	(0.06)	
Observations	301	173	212	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal 1-3" is an indicator variable taking value 1 if a subject received one of the best three signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 23: The effect of a "good" signal on beliefs in the control condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal 1-5	4.189*	2.306	2.027	7.592**
	(2.40)	(2.96)	(2.62)	(3.40)
Bayes	0.662***	0.866***	0.715***	
	(0.03)	(0.07)	(0.04)	
Observations	1010	483	652	358

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

Standard errors clustered at the individual level. Their values in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal 1-3" is an indicator variable taking value 1 if a subject received one of the best three signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

D.4 Guessing a random number

Figure 23: Decisions after observing a number from Box 1 or Box 2.

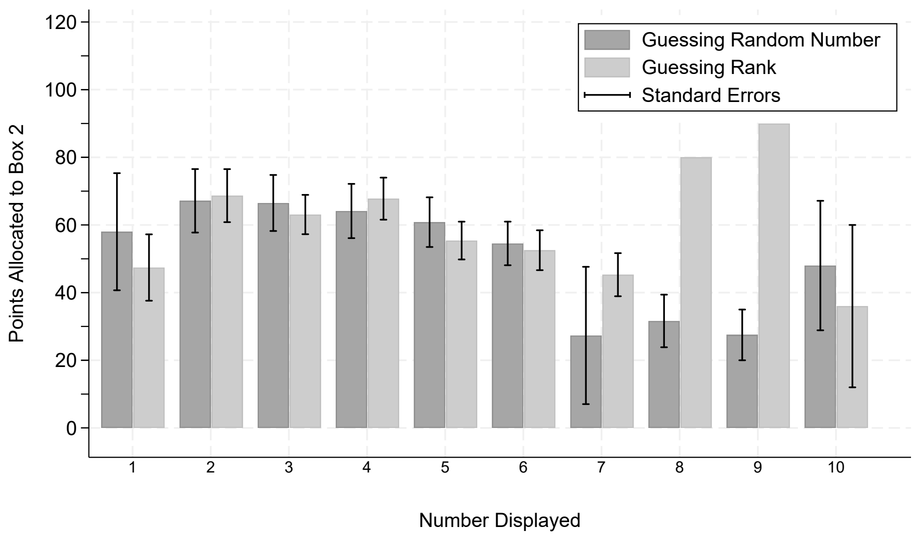


Table 24: The effect of a random signal's valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	4.096	5.347*	-2.310	-0.314	-0.266	1.214	10.336**	11.322**
	(3.71)	(2.91)	(4.65)	(3.82)	(4.30)	(3.48)	(4.86)	(4.83)
Good Signal	5.851*	3.073	-0.417	-1.187	-0.238	-0.849	7.979**	4.833
	(3.54)	(2.41)	(3.17)	(2.97)	(3.09)	(2.54)	(3.60)	(3.90)
Treat × Good	2.992	1.599	16.860**	14.360**	13.937**	11.099*	-13.949	-13.912*
	(6.84)	(5.02)	(7.06)	(6.50)	(6.99)	(5.66)	(8.46)	(8.39)
Observations	1154	1154	551	551	743	743	411	411

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. I only include participants who assessed a number from Box 1 in the treatment condition.

Table 25: The effect of a random "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	3.940	9.794*	8.880*	-5.970
	(4.50)	(5.86)	(5.20)	(8.27)
Bayes	0.640***	0.771***	0.660***	
	(0.07)	(0.15)	(0.09)	
Observations	144	68	91	53

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. I only include participants who assessed a number from Box 1 in the treatment condition.

D.5 Time effects: the results based on the first vs the last decisions.

Table 26: Decisions in the control condition: first five vs last five reports.

	Average Report			p-value		
	Decision 1-5	Decision 6-10		Diff < 0	Diff $\neq 0$	Diff > 0
All	31.54	28.52	*p-value:*	0.064	0.128	0.936
	(1.413)	(1.388)				
Observations:	505	505				
Exp	49.344	51.230	*p-value:*	0.232	0.464	0.768
	(1.786)	(1.848)				
Observations:	253	230				
Exp+	40.641	42.814	*p-value:*	0.183	0.365	0.817
	(1.634)	(1.754)				
Observations:	351	301				
Unexp	10.812	7.436	*p-value:*	0.940	0.119	0.060
	(1.904)	(1.208)				
Observations:	154	204				

Standard errors in parentheses.

Table 27: Time effects: the results based on the first vs the last decisions in control.

	All		*Exp*		*Exp+*		*Unexp*	
Decisions:	*1-5*	*6-10*	*1-5*	*6-10*	*1-5*	*6-10*	*1-5*	*6-10*
Treatment	4.690	7.028**	0.593	0.182	2.657	0.855	9.271**	8.167**
	(3.231)	(2.985)	(3.832)	(3.718)	(3.555)	(3.709)	(3.592)	(3.764)
Good Signal	7.711**	4.709	-2.184	1.219	-0.883	0.045	18.854***	1.779
	(3.849)	(3.940)	(4.040)	(4.195)	(3.643)	(4.207)	(6.154)	(3.632)
Treat × Good	5.783	8.784	14.217**	10.815*	13.513**	12.586**	-24.931***	-7.856
	(5.492)	(5.556)	(5.692)	(5.803)	(5.471)	(5.862)	(8.275)	(6.611)
Observations	806	806	426	403	563	513	243	293

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. No controls included. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). In the first specification (Decisions: 1-5), the results are based on the data from the first 5 decisions made the subjects in the control condition. In the second specification (Decisions: 6-10), the results are based on the data from the last 5 decisions made the subjects in the control condition. The number of observations differ between specifications because the first (last) five decisions might contain more (less) signals to which a subject assigned non-zero prior.

Table 28: Time effects: the results based on the first vs the last decisions in control.

	All		*Exp*		*Exp+*		*Unexp*	
Decisions:	*1-5*	*6-10*	*1-5*	*6-10*	*1-5*	*6-10*	*1-5*	*6-10*
Treatment	2.775	4.900**	0.172	0.469	-0.409	0.618	8.632**	9.696***
	(2.326)	(2.218)	(3.163)	(3.107)	(2.788)	(2.777)	(3.616)	(3.491)
Good Signal	5.966**	1.202	-1.050	-0.700	0.167	-1.322	16.408**	-0.962
	(2.892)	(2.671)	(3.408)	(3.798)	(2.822)	(3.379)	(6.567)	(3.842)
Treat × Good	2.102	5.167	11.475**	11.257**	11.478**	10.843**	-23.159***	-9.525
	(3.982)	(4.070)	(5.128)	(5.447)	(4.449)	(4.906)	(7.826)	(6.234)
Observations	806	806	426	403	563	513	243	293

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). In the first specification (Decisions: 1-5), the results are based on the data from the first 5 decisions made the subjects in the control condition. In the second specification (Decisions: 6-10), the results are based on the data from the last 5 decisions made the subjects in the control condition. The number of observations differ between specifications because the first (last) five decisions might contain more (less) signals to which a subject assigned non-zero prior.

Table 29: The effect of a "good" signal in the control condition (the first five decisions).

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	4.321	0.630	-0.966	18.854***
	(2.923)	(3.337)	(2.958)	(6.155)
Bayes	0.642***	0.939***	0.712***	
	(0.043)	(0.103)	(0.050)	
Observations	505	253	351	154

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

Standard errors clustered at the individual level. Their values in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results are based on the first five decisions in the control conditions. The estimates in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 30: The effect of a "good" signal in the control condition (the last five decisions).

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	1.429	-0.600	0.458	1.779
	(2.704)	(3.758)	(3.404)	(3.631)
Bayes	0.701***	0.804***	0.730***	
	(0.038)	(0.102)	(0.057)	
Observations	505	230	301	204

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

Standard errors clustered at the individual level. Their values in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results are based on the last five decisions in the control conditions. The estimates in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

D.6 Excluding "default" responses

Table 31: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	6.570**	4.601**	0.391	0.883	2.323	0.789	9.236***	9.546***
	(2.90)	(2.11)	(3.31)	(2.79)	(3.18)	(2.49)	(3.41)	(3.34)
Good Signal	5.692	2.885	-0.442	-1.674	-0.504	-1.144	8.227**	5.162
	(3.66)	(2.45)	(3.45)	(3.09)	(3.31)	(2.61)	(3.57)	(3.87)
Treat × Good	7.806	4.303	12.996**	12.136**	13.403**	11.767***	-14.304**	-14.173**
	(5.42)	(3.87)	(5.36)	(5.05)	(5.33)	(4.41)	(6.57)	(6.37)
Observations	1260	1260	616	616	817	817	443	443

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. The sample excludes people who took the "default" action: choosing the default probability of 0.5.

Table 32: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.307*	9.760***	8.955***	-6.077
	(2.82)	(3.37)	(3.07)	(6.23)
Bayes	0.698***	0.859***	0.780***	
	(0.04)	(0.09)	(0.06)	
Observations	297	169	208	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp"). The sample excludes people who took the "default" action: choosing the default probability of 0.5.

Table 33: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	2.829	3.108	0.342	0.726	2.354	0.749	10.016*	10.907**
	(3.03)	(2.53)	(3.43)	(2.89)	(3.22)	(2.63)	(5.78)	(5.45)
Good Signal	1.026	1.331	-1.642	-1.946	-1.084	-1.439	10.314	8.656
	(3.59)	(3.08)	(3.62)	(3.20)	(3.31)	(2.81)	(6.22)	(6.53)
Treat × Good	10.354*	5.457	14.967***	13.219**	11.821**	10.989**	-18.620	-18.809*
	(5.44)	(4.58)	(5.58)	(5.30)	(5.37)	(4.67)	(11.91)	(10.96)
Observations	909	909	587	587	729	729	180	180

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. The sample excludes people who took the "default" action (choosing the default probability of 0.5) or chose the exact probability prescribed by the Bayes' rule.

Table 34: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.799*	10.335***	8.769***	-8.306
	(3.30)	(3.50)	(3.26)	(10.90)
Bayes	0.572***	0.850***	0.726***	
	(0.05)	(0.09)	(0.06)	
Observations	237	162	191	46

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp"). The sample excludes people who took the "default" action (choosing the default probability of 0.5) or chose the exact probability prescribed by the Bayes' rule.

Table 35: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	2.930	3.028	0.375	0.612	2.476	0.708	9.792*	10.415*
	(3.07)	(2.61)	(3.51)	(2.98)	(3.26)	(2.72)	(5.81)	(5.50)
Good Signal	0.876	1.299	-2.143	-2.166	-1.594	-1.622	11.058*	9.783
	(3.64)	(3.22)	(3.72)	(3.34)	(3.35)	(2.94)	(6.51)	(6.86)
Treat × Good	8.249	5.493	14.241**	14.484***	10.187*	11.460**	-19.364	-19.208*
	(5.66)	(4.80)	(5.94)	(5.57)	(5.62)	(4.87)	(12.07)	(11.16)
Observations	856	856	543	543	679	679	177	177

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. The sample excludes people who took the "default" action (choosing the default probability of 0.5) or chose the probability prescribed by the Bayes' rule allowing for mistakes (± 1 point).

Table 36: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.272	10.610***	8.570**	-8.306
	(3.52)	(3.83)	(3.54)	(10.90)
Bayes	0.553***	0.835***	0.705***	
	(0.05)	(0.10)	(0.07)	
Observations	222	148	176	46

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp"). The sample excludes people who took the "default" action (choosing the default probability of 0.5) or chose the probability prescribed by the Bayes' rule allowing for mistakes (± 1 point).

D.7 Excluding people who confused rank order

Table 37: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	5.353*	3.259	-0.410	0.044	1.224	-0.435	8.676**	8.843***
	(2.87)	(2.06)	(3.17)	(2.73)	(3.09)	(2.45)	(3.47)	(3.35)
Good Signal	5.195	1.675	-1.236	-0.740	-0.913	-0.811	5.736*	3.589
	(3.57)	(2.22)	(3.10)	(2.87)	(3.06)	(2.51)	(3.31)	(3.97)
Treat × Good	8.298	5.332	13.269***	12.754***	13.543***	12.424***	-11.813*	-11.937*
	(5.30)	(3.75)	(5.06)	(4.84)	(5.10)	(4.26)	(6.43)	(6.20)
Observations	1291	1291	651	651	857	857	434	434

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank. Excluding people who confused rank order in the control condition (2 subjects).

Table 38: Matching.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	6.644	13.382**	12.263**	-9.099
	(4.48)	(5.64)	(5.20)	(7.83)
Observations	301	173	212	89

The reported errors are bootstrap standard errors based on 500 replications.

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

Note: The dependent variable is the difference between the actual decision and a counterfactual constructed using the nearest neighbor matching (the number of neighbors is set to 3). To predict the counterfactual outcome, I use all observations from the control condition and the following characteristics: the signal received (exact match), the prior assigned to it, and the participant's rank.

"Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further from their priors ("Unexp"). Excluding people who confused rank order in the control condition (2 subjects).

Table 39: The effect of a "good" signal on beliefs in the control condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	1.447	-1.041	-1.042	5.736*
	(2.26)	(2.87)	(2.53)	(3.31)
Bayes	0.692***	0.894***	0.734***	
	(0.03)	(0.07)	(0.04)	
Observations	990	478	645	345

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

Standard errors clustered at the individual level. Their values in parentheses.

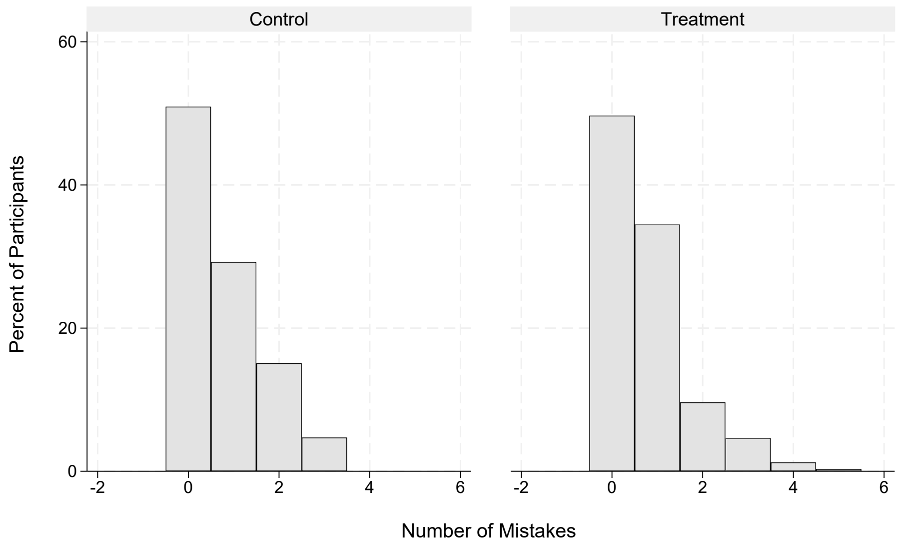
Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp"). Excluding people who confused rank order in the control condition (2 subjects).

D.8 Including people with many mistakes in control questions

D.8.1 Mistakes in control questions

The experimental tasks required a good understanding of the instructions. For this reason, before the main task, participants had to solve a set of control questions. While I allowed confused participants to finish the experiment, I collected the data on the number of mistakes and removed the most mistaken subjects from the main analysis. In the end, I excluded 25 participants who made three or more mistakes in five control questions.[29] Additionally, one participant reported to the assistant after the session that he mixed up the two boxes. His cubicle number was noted and this observation was removed from the analysis (the participant with a number 340).

Figure 24: Mistakes in the five control questions in the two conditions.



[29]One of the control questions (Question 2) was slightly different in the two conditions, and participants in the control condition made more mistakes in their version. I did not include this question in the measure of subjects' mistakes.

D.8.2 Results based on an entire sample

Table 40: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	5.517**	3.554*	-0.042	-0.010	1.811	-0.181	8.400***	8.803***
	(2.78)	(2.06)	(3.18)	(2.71)	(3.07)	(2.46)	(3.22)	(3.12)
Good Signal	6.046*	2.883	-0.654	-2.230	-0.429	-1.455	7.885**	4.894
	(3.40)	(2.32)	(3.08)	(2.83)	(2.98)	(2.45)	(3.44)	(3.61)
Treat × Good	7.025	4.131	12.149**	11.687**	11.761**	10.948**	-14.231**	-14.219**
	(5.08)	(3.72)	(5.00)	(4.84)	(4.98)	(4.23)	(6.26)	(6.03)
Observations	1381	1381	685	685	904	904	477	477

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 41: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	4.283	8.201**	7.532**	-6.346
	(2.69)	(3.27)	(2.95)	(5.90)
Bayes	0.697***	0.835***	0.776***	
	(0.04)	(0.09)	(0.06)	
Observations	321	183	224	97

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 42: The effect of a "good" signal on beliefs in the control condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	2.533	-0.599	-0.567	7.885**
	(2.31)	(2.86)	(2.49)	(3.44)
Bayes	0.669***	0.850***	0.709***	
	(0.03)	(0.07)	(0.04)	
Observations	1060	502	680	380

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

Standard errors clustered at the individual level. Their values in parentheses.

D.9 Gender Differences

Table 43: Differences between men and women.

	Men	**Women**		Diff < 0	Diff $\neq 0$	Diff > 0
IQ test score	5.323	5.250	*p-value:*	0.572	0.856	0.428
Rank	5.517	5.602	*p-value:*	0.392	0.784	0.607
Measures of Belief Distribution:						
Mean Belief	4.508	4.716	*p-value:*	0.137	0.274	0.863
1^{st} Quartile	3.752	4.009	*p-value:*	0.090	0.179	0.910
Median Belief	4.497	4.704	*p-value:*	0.144	0.289	0.856
3^{st} Quartile	5.233	5.398	*p-value:*	0.205	0.410	0.795
Range	4.986	5.102	*p-value:*	0.258	0.516	0.742
N	294	108				

Table 44: Decisions about signals with non-zero prior probability.

	Men	**Women**		Diff < 0	Diff $\neq 0$	Diff > 0
Decision Treatment	56.146	56.060	*p-value:*	0.508	0.984	0.492
	(2.491)	(3.563)				
N	123	50				
Decision Control	49.905	51.198	*p-value:*	0.329	0.659	0.671
	(1.501)	(2.486)				
N	357	126				

Note: Decision Treatment (Control) denotes the number of points allocated to Box 2 after observing a signal (considering a signal) in the treatment (control) condition. Standard errors in parentheses.

Table 45: Decisions about signals with zero prior probability.

	Men	**Women**		Diff < 0	Diff $\neq 0$	Diff > 0
Decision Treatment	12.740	20.438	*p-value:*	0.061	0.123	0.939
	(2.190)	(5.538)				
N	96	32				
Decision Control	11.214	12.388	*p-value:*	0.292	0.583	0.709
	(1.057)	(1.949)				
N	393	134				

Note: Decision Treatment (Control) denotes the number of points allocated to Box 2 after observing a signal (considering a signal) in the treatment (control) condition. Standard errors in parentheses.

D.9.1 Gender differences: data analysis

Table 46: The treatment effect controlling for subject's gender.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	3.630	1.410	-0.899	-2.174	0.630	-1.966	5.533	5.884*
	(3.29)	(2.33)	(3.89)	(3.28)	(3.67)	(2.82)	(3.35)	(3.35)
Good Signal	6.293	3.123	-1.257	-2.288	0.172	-0.531	6.858*	4.072
	(4.19)	(2.88)	(3.89)	(3.65)	(3.75)	(3.19)	(3.88)	(4.27)
Treat × Good	11.382*	6.708	16.247***	15.364***	15.424**	12.826**	-6.680	-7.127
	(6.24)	(4.40)	(6.18)	(5.85)	(6.15)	(5.08)	(7.78)	(7.70)
Female	2.252	0.908	-0.145	-0.584	0.733	0.297	1.683	0.474
	(3.85)	(2.71)	(4.16)	(3.90)	(3.72)	(3.35)	(2.97)	(2.84)
Female × Good	-1.716	-0.224	3.308	3.887	-1.583	-1.243	3.952	4.417
	(7.76)	(5.07)	(6.40)	(5.89)	(6.43)	(4.94)	(9.11)	(9.39)
Female × Treat	8.270	9.644*	4.544	9.161	4.526	7.820	14.199	14.910
	(6.58)	(5.18)	(6.99)	(6.28)	(6.98)	(5.85)	(10.20)	(10.30)
Fem × Treat × Good	-13.739	-8.928	-13.535	-11.956	-8.677	-4.480	-29.729**	-28.273*
	(11.63)	(8.36)	(10.76)	(9.62)	(10.96)	(8.66)	(14.68)	(14.91)
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. "Female" is an indicator variable taking value 1 if a subject is female. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank.

Table 47: The effect of a "good" signal in the treatment controlling for gender.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.052*	9.547***	8.696***	-6.509
	(2.77)	(3.28)	(3.00)	(6.20)
Female	4.761	4.080	3.753	8.854
	(2.97)	(3.60)	(3.27)	(6.30)
Bayes	0.698***	0.873***	0.784***	
	(0.04)	(0.09)	(0.06)	
Observations	301	173	212	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Female" is an indicator variable taking value 1 if a subject is female. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 48: The effect of a"good" signal in the treatment controlling for gender.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	7.735**	11.560***	9.991***	0.178
	(3.27)	(3.89)	(3.58)	(7.10)
Female	8.434**	7.102	5.665	15.882**
	(3.79)	(4.79)	(4.35)	(7.28)
Female $\times$ Good	-9.404	-6.920	-4.413	-25.778*
	(6.08)	(7.22)	(6.60)	(13.95)
Bayes	0.694***	0.868***	0.781***	
	(0.04)	(0.09)	(0.06)	
Observations	301	173	212	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Female" is an indicator variable taking value 1 if a subject is female. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 49: The effect of a“good” signal in the control condition, controlling for gender.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	2.714	-1.040	0.150	6.858*
	(2.83)	(3.60)	(3.18)	(3.88)
Female	1.448	0.062	1.012	1.683
	(2.68)	(4.01)	(3.45)	(2.97)
Bayes	0.673***	0.870***	0.721***	
	(0.03)	(0.07)	(0.04)	
Observations	1010	483	652	358

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. “Good Signal” is an indicator variable taking value 1 if a subject received one of the best four signals. “Female” is an indicator variable taking value 1 if a subject is female. “Bayes” is the number of points that should be allocated given one’s prior beliefs. The results in the first column are based on the entire sample (“All”). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal (“Exp”), received a signal close to their prior belief distribution (“Exp+”), or received an “unexpected” signal that was further away from their priors (“Unexp”).

Table 50: The effect of a“good” signal in the control condition, controlling for gender.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	2.714	-1.040	0.150	6.858*
	(2.83)	(3.60)	(3.18)	(3.88)
Female	1.448	0.062	1.012	1.683
	(2.68)	(4.01)	(3.45)	(2.97)
Female × Good	-0.202	3.782	-1.697	3.952
	(5.19)	(6.00)	(5.11)	(9.10)
Bayes	0.673***	0.870***	0.721***	
	(0.03)	(0.07)	(0.04)	
Observations	1010	483	652	358

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the number of points allocated to Box 2. “Good Signal” is an indicator variable taking value 1 if a subject received one of the best four signals. “Female” is an indicator variable taking value 1 if a subject is female. “Bayes” is the number of points that should be allocated given one’s prior beliefs. The results in the first column are based on the entire sample (“All”). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal (“Exp”), received a signal close to their prior belief distribution (“Exp+”), or received an “unexpected” signal that was further away from their priors (“Unexp”).

E Belief Elicitation II

In this section, I explore the data from the second belief elicitation. First, I compare it to the prior beliefs and show how learning depends on a signal realization. I show that the final belief is consistent with the decision about the box, and there is no additional asymmetry in how participants translate the belief about a signal into the belief about the rank. Second, I demonstrate that the results continue to hold if I use the final belief about the rank instead of the belief about the box elicited in the main task (I replicate the results from Tables 1-4 using this final beliefs in Section E.2). Lastly, I look at the changes in distributions of beliefs across the two conditions (Section E.4). I confirm that beliefs move in the predicted direction and the changes are correlated with subjects' emotional states in the treatment but not in the control condition. The results provide strong evidence that the treatment manipulation worked as expected.

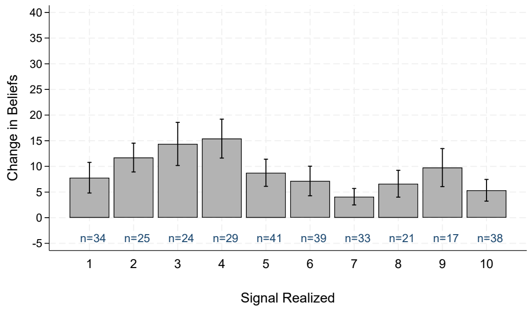
Several points should be kept in mind when interpreting the data. In experiments that measure beliefs multiple times, one common problem is people's desire to be seen as consistent decision-makers (Falk and Zimmermann, 2017). Despite the best efforts to ensure anonymity and instruct subjects to treat each part of the experiment independently, these motives are likely to influence beliefs reported in Belief Elicitation II.[30] Thus, "stickiness" of beliefs is to be expected, with participants holding on to their prior beliefs also in the treatment condition. Second, while I explained in intuitive terms how one can translate prior beliefs into a posterior about the box, I provided no such guidance on how to translate the beliefs about the signal into the posterior distribution about the rank. The uncertainty about how to approach the problem might exacerbate the tendency to hold on to the prior belief distribution.

[30] This concern is alleviated in the main analysis, for two reasons. First, I elicited posterior beliefs in a different way: by asking about the box. Second, the analysis is based on a comparison between the treatment and the control condition, and there is no reason to suspect that consistency motives operate differently in the two conditions.

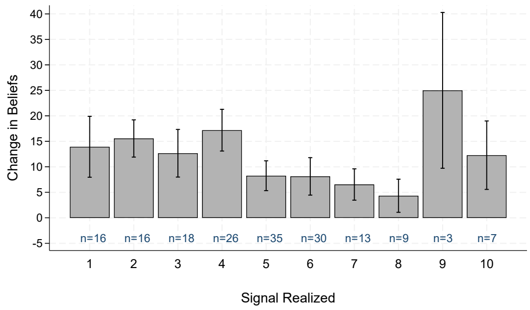
E.1 Change in beliefs about the rank

On average, participants in the treatment condition allocated 32.86 points to the rank corresponding to the received signal in the second belief elicitation. This value is 11.36 points higher than the number of points allocated to the same rank in Belief Elicitation I. Figure 25 shows how the difference in the number of allocated points depends on the signal realization. In Panel a), the results are based on the data from all participants, and in Panel b), participants in the "Exp" group (defined as previously). For example, in the "Exp" group, 16 subjects who received a signal "2" allocated, on average, 15 points more to Rank 2 in the second than in the first belief elicitation. Therefore, they revealed a 15 pp higher probability that "2" is their rank. One can notice that the change in beliefs depends on the signal value. The average change in beliefs is 80% higher after signals 1 to 4, compared to signals 5 to 10 (p-value of one-tailed t-test = 0.0097). The difference remains positive and significant if I control for prior beliefs or the Bayesian benchmark. The results based on the data from participants in the "Exp+" and the "Unexp" group are presented in Figure 26. The results in the Panel a) are very similar to the "Exp" group in Figure 25, whereas participants in the "Unexp" group (Panel b) seem to allocate more points to higher (worse) ranks.

Figure 25: Changes in the number of points allocated to the rank = signal.



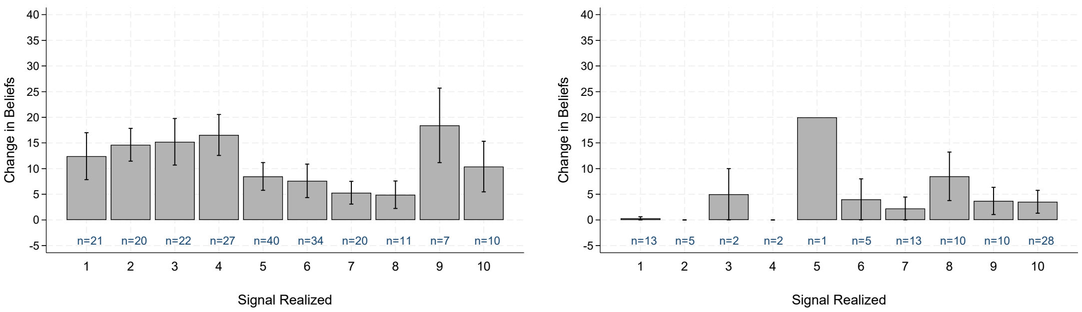
(a) All subjects.



(b) Expected signals.

Note: The bars show changes in the number of points allocated to the rank indicated by a signal in Belief Elicitation I and II. The whiskers denote standard errors for a difference between two means. The graph in Panel a) is based on the data from all subjects, and the graph in Panel b) is based on the data from subjects who assigned non-zero prior probability to the rank corresponding to the signal received. The graphs use only the data from the treatment conditions.

Figure 26: Changes in the number of points allocated to the rank = signal.



(a) Expected or adjacent signals.

(b) Unexpected signals.

Note: The bars show changes in the number of points allocated to the rank indicated by a signal in Belief Elicitation I and II. The whiskers denote standard errors for a difference between two means. The graph in Panel a) is based on the data from the "Exp+" group, and the graph in Panel b) is based on the data from subjects who received unexpected signals. The graphs use only the data from the treatment.

E.2 Analysis based on change in beliefs about the rank

Table 51: The effect of signal valence across the two conditions.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Treatment	7.180***	7.013***	10.107***	9.205***	8.866***	8.234***	4.212***	4.534***
	(1.10)	(1.11)	(1.87)	(1.83)	(1.54)	(1.51)	(1.35)	(1.34)
Good Signal	1.007	1.089	1.167	-0.000	0.726	0.445	1.994*	0.354
	(0.75)	(0.80)	(0.87)	(0.99)	(0.72)	(0.74)	(1.19)	(0.80)
Treat × Good	4.269**	4.607**	5.492*	6.147**	5.996**	6.806***	-5.685***	-6.168***
	(2.16)	(2.13)	(2.99)	(2.99)	(2.61)	(2.61)	(1.86)	(1.96)
Controls		✓		✓		✓		✓
Observations	1311	1311	656	656	864	864	447	447

Standard errors clustered at the individual level. Their values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the change in final beliefs about the rank corresponding to the number displayed on-screen. "Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp"). No control variables are included in the first specification. In Specification (2), the controls include the prior probability assigned to the rank equivalent to the signal received, the mean of the prior belief distribution, and the subject's rank.

Table 52: Matching.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	5.369**	6.375**	6.722**	-3.839**
	(2.19)	(3.18)	(2.71)	(1.61)
Observations	301	173	212	89

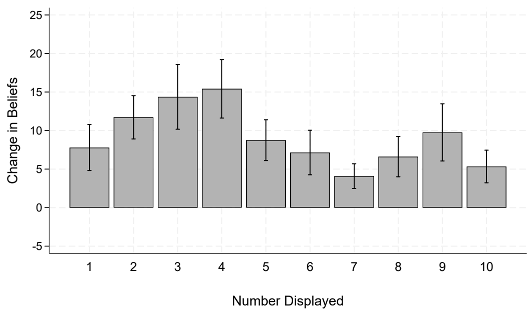
The reported errors are bootstrap standard errors based on 500 replications.

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

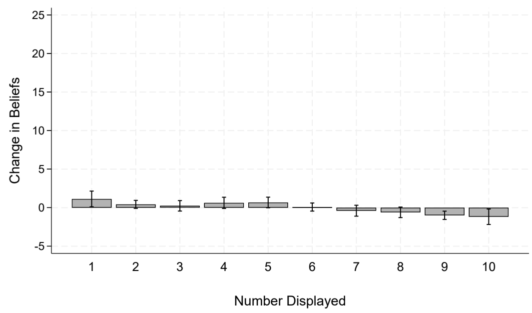
Note: The dependent variable is the difference between the final belief and a counterfactual constructed using the nearest neighbor matching (the number of neighbors is set to 3). To predict the counterfactual outcome, I use all observations from the control condition and the following characteristics: the subject's rank, the signal received (exact match), the prior assigned to it, and the mean of the prior belief distribution.

"Good Signal" is an indicator variable that takes the value 1 if a subject received one of the best four signals. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further from their priors ("Unexp").

Figure 27: Comparison between the treatment and the control condition.



(a) Treatment condition.



(b) Control condition.

Note: The bars show how seeing a number affected the change in the average number of points allocated to the corresponding rank. The change is defined as the difference in points allocated in the final and the initial belief elicitation. The whiskers denote standard errors. The graph in Panel a) is based on the data from the treatment condition, and the graph in Panel b) is based on the data from subjects in the control group.

Table 53: The effect of a "good" signal on beliefs in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	4.054**	6.194**	6.319**	-3.692
	(1.91)	(2.80)	(2.44)	(2.37)
Bayes	0.094***	0.149**	0.074*	
	(0.03)	(0.07)	(0.04)	
Observations	301	173	212	89

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the change in final beliefs about the rank corresponding to the number displayed on-screen. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

Table 54: The effect of a "good" signal on beliefs in the control condition.

	All	*Exp*	*Exp+*	*Unexp*
Good Signal	1.007	1.167	0.726	1.994*
	(0.75)	(0.87)	(0.72)	(1.19)
Observations	1010	483	652	358

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

Standard errors clustered at the individual level. Their values in parentheses.

Note: The dependent variable is the change in final beliefs about the rank corresponding to the number displayed on-screen. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. "Bayes" is the number of points that should be allocated given one's prior beliefs. The results in the first column are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the signal ("Exp"), received a signal close to their prior belief distribution ("Exp+"), or received an "unexpected" signal that was further away from their priors ("Unexp").

E.3 Consistency

I examine the relation between the decisions about the box and the posterior about the rank using a regression analysis. The dependent variable is the number of points allocated to the relevant rank in Belief Elicitation II. I regress this value on two independent variables. "Prior Belief" denotes the number of points allocated to the relevant rank in Belief Elicitation I. The independent variable "Belief Box" denotes the number of points allocated to Box 2. The results are gathered in Table 55. The estimates in the first column show that both variables have a positive and significant effect on the final belief. In the second specification, I add an independent variable "Good Signal" defined as in the previous section. The coefficient is not significant, meaning that there is no additional effect of a "good" signal beyond the effect it had on the decision about the box. The results validate the assumption that beliefs about the box reveal one's beliefs about the rank. One can also view them as an additional check that subjects understood the main task.

Table 55: The effect of the decision about the box on the posterior belief.

	All		*Exp*		*Exp+*		*Unexp*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Prior Belief	0.898***	0.894***	0.836***	0.843***	0.771***	0.779***		
	(0.08)	(0.08)	(0.11)	(0.11)	(0.10)	(0.10)		
Belief Box	0.231***	0.224***	0.341***	0.326***	0.321***	0.306***	0.042	0.036
	(0.04)	(0.04)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
Good Signal		2.779		3.107		3.550		-3.473
		(1.81)		(2.66)		(2.29)		(2.39)
Observations	301	301	173	173	212	212	89	89

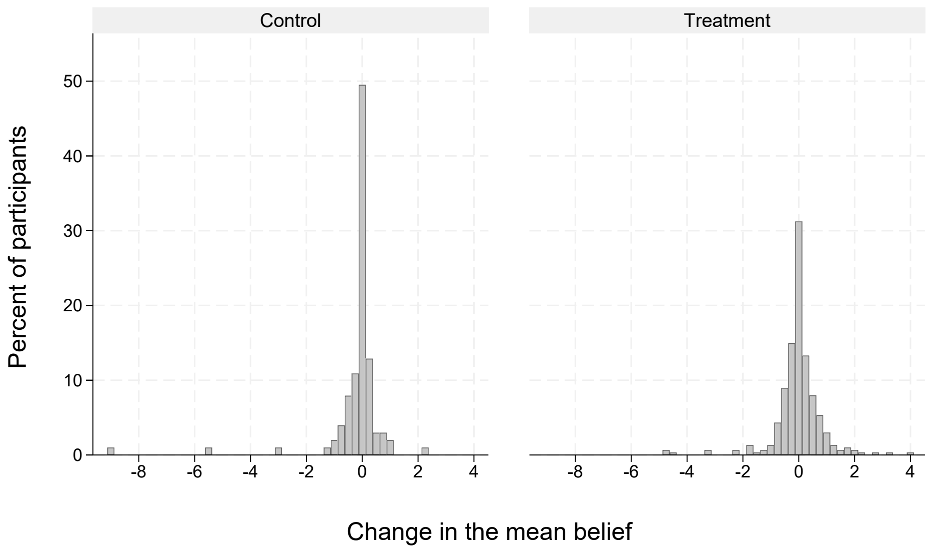
Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of points allocated to the rank corresponding to the received signal. "Prior Belief" is the number of points allocated to the rank corresponding to the received signal in the first belief elicitation. "Belief Box" is the number of points allocated allocated to Box 2 (indicative of one's rank) in the main task. "Good Signal" is an indicator variable taking value 1 if a subject received one of the best four signals. The results in the first two columns are based on the entire sample ("All"). In the remaining columns, the results are based on a sample of participants who assigned a non-zero prior probability to the number displayed on screen ("Exp"), a sample of participants who received a signal close to their prior belief distribution ("Exp+"), and participants who received an "unexpected" signal that was further away from their priors ("Unexp").

E.4 Manipulation check

Figure 28 shows changes in beliefs in the treatment and in the control condition. There are several things to be noted. First, the fraction of participants who did not change their beliefs is much higher in the control condition (50% of subjects compared 30% in the treatment condition). Second, even though participants in Control did not receive any new information, half of them reported a different belief distribution after the task. This might be due to reconsidering one's ability over the course of experiment, or difficulty in entering the same distribution (95% of changes in the control condition are no higher than ± 1 rank). Moreover, two outliers in the control condition visible in Figure 28 are most likely people who confused the ordering of ranks (mistakenly considered Rank 10 to be the best rank). Participant no. 230 allocated 100 points to Rank 10 in the first belief elicitation and 100 points to Rank 1 in the second. Similarly, participant no. 458 allocated all points to the last three ranks in the first elicitation, and to the first three ranks in the second; his second report is the mirror image of the first. I exclude those participants from the analysis presented in this section.[31,32]

Figure 28: Change in the mean belief in the two conditions.



[31] If not excluded, the two observations increase the average change in beliefs in the control condition to -0.19, a value four times higher than the average of the remaining observations.

[32] In Section D.7, I take a more systematic approach to identify subjects who confused the rank order in both conditions, and I re-do the main analysis excluding those participants. All other results presented in the paper are unaffected by those subjects.

In Table 56, I compare the average prior and posterior in the two conditions. Since high and low signals pool the mean in opposite directions, I distinguish subjects who received a signal that was higher than the mean of their prior belief distribution, and those who received a signal lower or equal. No similar distinction is made for participants in control, as they considered both types of signals (distinguishing subjects who saw more "high" than "low" signals does not change the results). As one can see in the upper part of Table 56, participants updated their beliefs in the treatment but not in the control condition. Second, the change in beliefs is significantly higher in the treatment compared to the control (see the lower part of Table 56). The results confirm that treatment manipulation worked as expected: only realized signals shifted subjects' beliefs.

Moreover, changes in beliefs are correlated with participants' emotional states in the treatment but not in the control condition. The results of a regression analysis are presented in Table 57. The dependent variable is the difference between the mean of individual belief distribution reported in the first and the second belief elicitation. I

Table 56: The average prior and posterior belief in the two conditions.

	Posterior	**Prior**		Diff < 0	Diff $\neq 0$	Diff > 0
Treatment, $s \leq \bar{p}_0$	4.620	4.942	*p-value:*	0.000	0.000	1.000
	(0.171)	(0.166)				
Treatment, $s > \bar{p}_0$	4.384	4.188	*p-value:*	0.999	0.002	0.001
	(0.120)	(0.109)				
Control	4.673	4.722	*p-value:*	0.178	0.356	0.822
	(0.166)	(0.170)				
	Treatment, $s \leq \bar{p}_0$	**Control**		Diff < 0	Diff $\neq 0$	Diff > 0
Change in beliefs:	-0.322	-0.049	*p-value:*	0.003	0.005	0.997
	(0.077)	(0.053)				
N	117	99				
	Treatment, $s > \bar{p}_0$	**Control**		Diff < 0	Diff $\neq 0$	Diff > 0
Change in beliefs:	0.194	-0.049	*p-value:*	0.996	0.008	0.004
	(0.053)	(0.060)				
N	184	99				

Note: Standard errors in parenthesis. $s > \bar{p}_0$ denotes a signal higher than the mean of individual belief distribution. Since Rank 10 denote the lowest performance, higher signals indicate worse states.

regress the difference in beliefs on a treatment dummy, an achievement emotion, and its interaction with the treatment. The coefficients at the interactions are positive and significant for negative emotions: anger, anxiety, shame, and hopelessness. Those experiencing more negative emotions reveal the final beliefs that is higher (less optimistic) than the prior. The result provides indirect evidence on relationship between instantaneous utility and belief formation. I examine this issue further in Appendix G.

Table 57: Predicting change in beliefs.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.231 (0.23)	0.480* (0.29)	0.304 (0.21)	0.023 (0.22)	-0.267 (0.19)	-0.441** (0.17)	-0.298* (0.18)	-0.179 (0.16)
Emotion 1: Enjoyment	-0.020 (0.05)							
Treat × Emotion 1	-0.043 (0.06)							
Emotion 2: Hope		0.000 (0.05)						
Treat × Emotion 2		-0.095 (0.06)						
Emotion 3: Pride			-0.022 (0.05)					
Treat × Emotion 3			-0.082 (0.06)					
Emotion 4: Relief				-0.035 (0.06)				
Treat × Emotion 4				0.011 (0.07)				
Emotion 5: Anger					-0.049 (0.05)			
Treat × Emotion 5					0.114** (0.06)			
Emotion 6: Anxiety						-0.079 (0.05)		
Treat × Emotion 6						0.226*** (0.06)		
Emotion 7: Shame							-0.030 (0.05)	
Treat × Emotion 7							0.128** (0.05)	
Emotion 8: Hopelessness								-0.030 (0.06)
Treat × Emotion 8								0.131* (0.07)
Observations	389	389	389	389	389	389	389	389

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The dependent variable is the difference in mean beliefs reported in the first and the second belief elicitation. All specifications include a control for the agent's rank. The sample consists of participants in the treatment and the control condition. The number of observations (389) is different from the baseline sample (402), because in one session, the data from the final questionnaire on emotions and emotion regulation was not saved due to an error. Moreover, I exclude 2 participants in the control condition who confused the rank order (the outliers in Figure 28). The effect of negative emotions in treatment is no less strong if these two observations are included in the sample.

F Literature: Design Comparison

The experimental design developed in this paper differs from the designs used in the literature. My main goal was to develop an updating task that induces a strong emotional reaction to a signal. I compare it to the experiments conducted in the past in Table 58. In this review, I focus on papers that study updating about ego-relevant characteristics and do so by asking subjects to update their beliefs about their *relative* performance.[33] The papers gathered in the first column in Table 58 are categorized based on one of the relevant design features. In the second column, I describe the design used in my experiment. The last column presents the rationale behind choosing this particular feature for my work.

One design feature that requires an additional comment is the information structure. In almost all of the work reviewed in this section, the information structure follows the scheme presented in Figure 29.[34] There are two states of the world H and L indicating whether one's score was in the upper or the lower half of the test score distribution. Subjects receive a signal that is informative about the state with known precision, e.g., 75% in Möbius et al. (2022). However, this signal structure becomes more complicated if extended to a larger signal and state space (see Figure 30) and I am not aware of any experimental work that implements it. Papers that used 10 states of the world, Eil and Rao (2011) and Zimmermann (2020), use binary signals shown in Figure 31. A signal informs a subject whether or not he ranked higher than another participant who was randomly drawn from a group of 10 (I denote these binary signals with H and L). The precision depends on the state and, for the first signal, takes one of the following values: 55.6%, 66.7%, 77.8%, 88.9% or 100% (for the second signal it is 50%, 62.5%, 75%, 87.5% or 100%, as comparisons are made without replacement).

[33] For a review of the literature on learning about absolute performance as well as updating about non-ego-relevant parameters, I refer the reader to Barron (2021) and Coutts (2019). An even broader review of the literature on errors in probabilistic reasoning can be found in Benjamin (2019).

[34] See Table 58 for the references.

Figure 29: Design used in the literature (2 states).

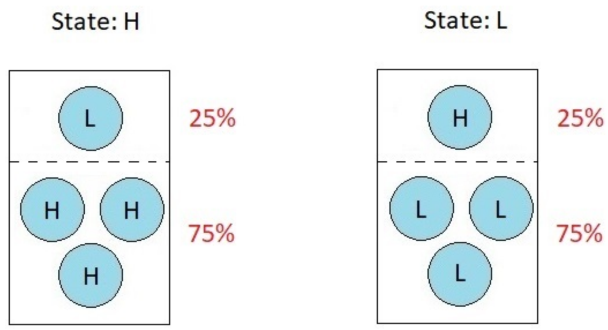


Figure 30: Design used in the literature extended to 10 states.

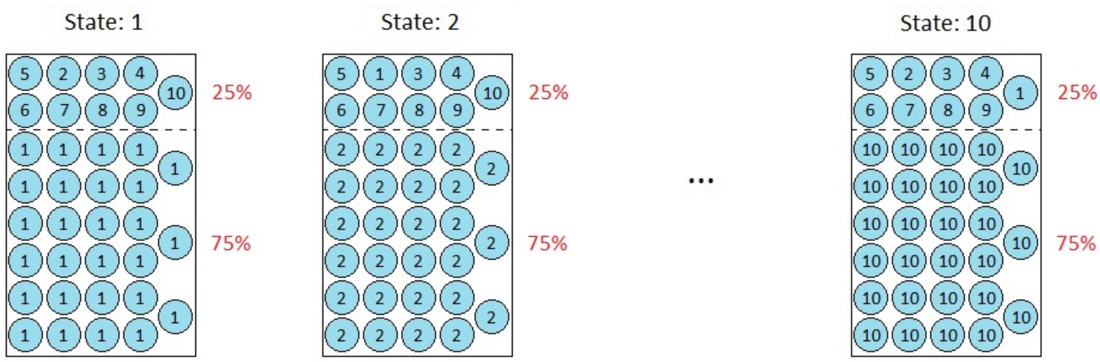
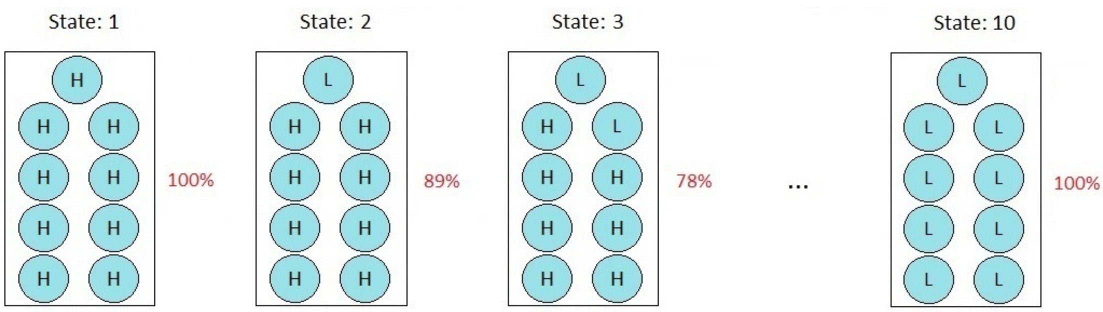


Figure 31: Design used in Eil and Rao (2011) and Zimmermann (2020).



The design commonly used in the literature (Figure 29) extended to 10 states can be simplified by distinguishing two urns: one with balls indicating the state (the "IQ" urn), and one with every possible number (the "Random" urn).[35] Figures 32 and 33 present the simplified design for 2 and 10 states of the world. The information structure in Figure 32 is equivalent to the one depicted on Figure 29, assuming either urn can be selected with equal probability $\frac{1}{2}$. If the state is H, a ball indicating H is drawn with probability $0.5 \cdot 0.5 + 0.5 \cdot 1 = 0.75$, the same as in Figure 29. Similarly, Figure 33 is equivalent to the information structure in Figure 30 with the signal precision of 55%.

Figure 32: Design developed in this paper (2 states).

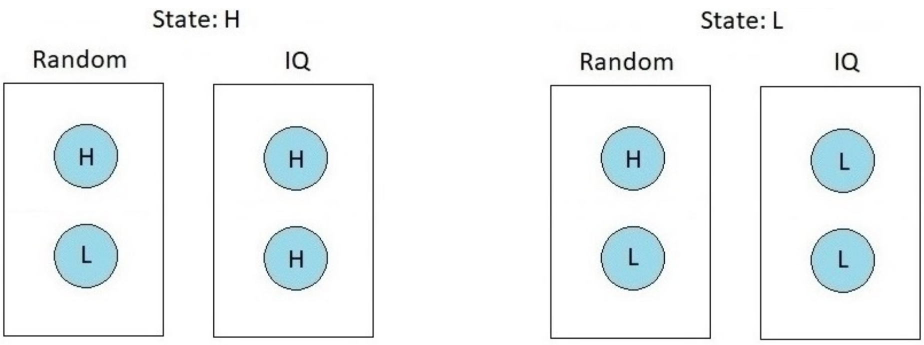
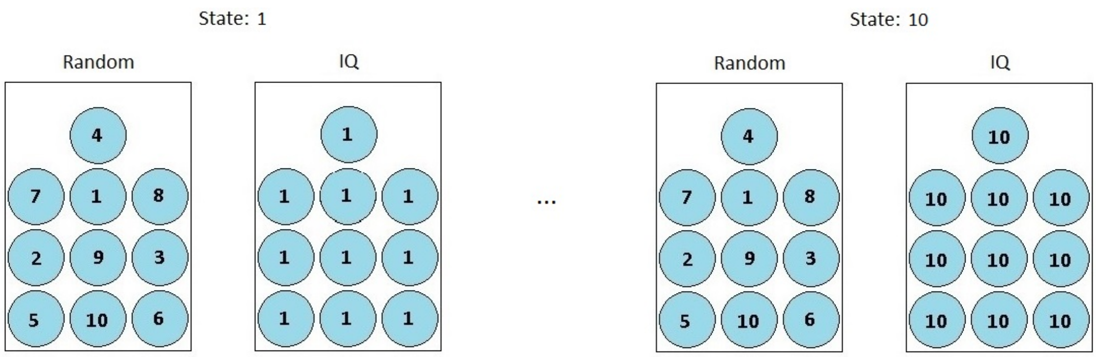


Figure 33: Design developed in this paper (10 states).



[35]One could also distinguish the two urns along the dashed line in Figure 30, with the Random urn containing all numbers except the one that indicates the state. This design, however, lacks the intuitive interpretation of "a random urn" from which *any number* can be drawn with *the same* probability, hence it might be more difficult to explain to the participants.

Table 58: Literarure review: design comparison.

Previous Work	This Paper	Goal
1. Number of signals:		
– more than 1 signal Buser et al. (2018), Coutts (2019), Drobner and Goerg (2024), Eil and Rao (2011), Möbius et al. (2022), and Zimmermann (2020).	– 1 signal	– separating reaction to signals from information aggregation.
– 1 signal Drobner (2022), Ertac (2011), and Schwardmann and Van der Weele (2019).		
2. State space, signal space, signal precision:		
– 2 states (above or below 50%; above or below 85% in Coutts, 2019), – 2 signal values, – signal precision: 67% Coutts (2019), Drobner (2022), and Drobner and Goerg (2024).	– 10 states (deciles of the distribution) – 10 signal values	– richer state space and signal space to induce a stronger emotional reaction to a signal (based on the observation that it is more painful for subjects to be in the bottom 10% than in the bottom 50%).
– 2 states (above or below 50%) – 2 signal values – signal precision: 70% Buser et al. (2018).		
– 2 states (above or below 50%) – 2 signal values – signal precision: 75% Möbius et al. (2022) and Schwardmann and Van der Weele (2019).		
– 3 states (lower 20%, middle 60%, or upper 20%) – 2 signal values – perfectly informative but coarse signals Ertac (2011).		
– 10 states (deciles of the distribution) – 2 signal values – signal precision depends on the state: 56%, 67%, 78%, 89% or 100%. Eil and Rao (2011) and Zimmermann (2020).		

Previous Work	**This Paper**	**Goal**
3. Information structure and implementation:		
– information structure as in Figure 29 – a signal is true or false with precision known to the subjects Buser et al. (2018), Coutts (2019), Drobner and Goerg (2024), Möbius et al. (2022), and Schwardmann and Van der Weele (2019). Drobner (2022), uses the same information structure (Figure 29), but the signal is a comparison with another subject.	– information structure shown in Figure 33 – it is equivalent to Figure 30 with a signal precision of 55%	– it would not be possible to introduce richer state and signal space using any other information structure from the literature.
– information structure as in Figure 31 – a signal is a pairwise comparison with another subject Eil and Rao (2011) and Zimmermann (2020).		
– a signal is always true, but only reveals whether the subject is in the top or the bottom half of the distribution, and not precisely the state Ertac (2011).		
4. Comparison group:		
– a group of 4 Drobner (2022) and Schwardmann and Van der Weele (2019).	– 300 subjects	– a larger comparison group makes for a stronger signal (e.g., in the case of a group of four, one can easily attribute a negative signal to being assigned to a particularly strong pair of subjects). When there is another way of “explaining” a bad signal, there may be no need for (costly) belief distortion.
– a group of 8 Buser et al. (2018).		
– a group of 10 Eil and Rao (2011), Ertac (2011), and Zimmermann (2020).		
– a group larger than 10 Coutts (2019), Drobner and Goerg (2024), and Möbius et al. (2022).		
5. Timing of revealing information:		
– In most of the papers mentioned above it is unclear whether and when subjects expected the resolution of uncertainty (see Drobner, 2022, for a comprehensive literature review). This problem was noticed and tested in the recent work by Drobner (2022).	– online access one week after the session	– to describe the behavior with a one-period model without the dynamic concerns – to bring the design closer to the real-world situations: grades are rarely immediate, need to be checked etc.

G Questionnaires

In this section, I describe the analysis of the data from questionnaires. The evidence is correlational: I use a regression analysis to investigate whether subjects' responses correlate with their decisions in the main task. I examine the data along two dimensions: 1) expected versus unexpected signals, and 2). "good" versus "bad" signals. Section G.1 aims to uncover which variables predict decisions in the treatment condition. In Section G.1.1, I directly show that these variables affect decisions differently after expected and unexpected signals.

G.1 Emotions and decisions in the main task

First, I examine whether the decisions in the main task are correlated with the psychological measures. I include eight measures of achievement emotions, four of which are positive (enjoyment, hope, pride, and relief) and four negative (anger, anxiety, shame, and hopelessness), two measures of habitual use of emotion regulation strategies (reappraisal and suppression), five measures of personality (BIG-5: extraversion, conscientiousness, openness, neuroticism, and agreeableness), and two measures of anxiety (anxiety as a state and as a trait; higher values indicating less anxious individual).[36]

In Table 59, I regress the dependent variable, the number of points allocated to Box 2, on the Bayesian benchmark and the above-mentioned measures. The columns correspond to the samples defined as in the main text. As we see in the first column, none of the measures is correlated with the decisions for all individuals. The remaining columns reveal why this is the case: participants' decisions are driven by different factors depending on whether the signal was expected or not. For expected signals, there is a strong and significant correlation with reappraisal: using more reappraisal in one's daily life is correlated with placing higher probability on the displayed number. This correlation has the opposite sign in the sample of people with unexpected signals. At the same time, there is a positive correlation between feeling hopeless (a negative anticipatory emotion) and the decisions after unexpected signals.

Table 60 reveals that the negative feelings are correlated with decisions after "bad" signals. In this table, I divide the sample depending on the signal value. The middle

[36]The later differs from the achievement emotion "anxiety" with respect to the time it was elicited (immediately after the IQ test, prior to learning about the task), and the nature of the measure (a general measure, no reference to the test).

columns show results based on signals 1 to 5 (the best signals, median split), and 6 to 10 (the worst signals, median split). The outer columns gather the results based on the best four and the worst four signals (1 to 4, and 7 to 10, respectively). The idea is to examine whether the results in the middle columns are stronger for more extreme signals. Unfortunately, the small sample size prevents us from simultaneously distinguishing expected and unexpected signals. I postpone examining the intersection of the two dimensions until Section G.1.1.

One disadvantage of the approach is the number of observations per variable. For this reason, I turn to a Lasso procedure—a method of choosing a subset of independent variables that are the best predictors of the variable of interest. Due to a small sample size, which precludes splitting the sample for cross-validation, I choose different λ for each sample such that the model includes only the first variable selected by the procedure. Tables 61 and 62 gather the variable selected for each sample (in the case of unexpected signals, I included two variables as they are both selected at the first place). As before, I look separately at the expected vs unexpected, and "good" vs "bad" signals.

The analysis reveal that the decisions after unexpected signals are associated with anxiety and hopelessness, whereas the decisions after expected signals correlate with participants' ability to regulate these emotions via the use of cognitive reappraisal.[37] A positive anticipatory emotion "hope" was selected for explaining decisions after "good" signals. The anticipatory nature of selected emotions should not come as a surprise—the essence of belief-based utility is enjoying something that is unknown at the moment, but likely to be beneficial in the future. In the following section, I look at the variables selected by the procedure and study their effect across different groups of signals.

[37]It is worth noting that the next variable chosen by the LASSO procedure in two groups was "Anxiety: State", which denotes the STAI measure of anxiety experienced at the moment. The results obtained using this variable are very similar to the results based on the self-reported measure presented in the next section. This is not surprising given a strong correlation between the two variables (Pearson's correlation coefficient $\rho = 0.62$ is significant at the 1%-level), and lends credibility to the self-reported measure.

Table 59: Emotions and decisions in the treatment condition.

	All	*Exp*	*Exp+*	*Unexp*
	(1)	(2)	(3)	(4)
Emotion: Enjoyment	-0.684	-1.657	-2.240*	0.391
	(0.98)	(1.36)	(1.22)	(1.76)
Emotion: Hope	-0.089	0.552	0.111	-2.557
	(1.22)	(1.48)	(1.33)	(2.65)
Emotion: Pride	1.249	1.461	1.209	2.094
	(1.40)	(1.71)	(1.55)	(3.00)
Emotion: Relief	0.556	0.716	0.484	0.774
	(1.41)	(1.79)	(1.60)	(2.91)
Emotion: Anger	-1.156	-1.049	-1.017	-2.232
	(1.02)	(1.38)	(1.20)	(1.95)
Emotion: Anxiety	2.856	3.008	2.957	4.395
	(1.81)	(2.17)	(2.01)	(3.79)
Emotion: Shame	0.059	0.831	0.629	-1.498
	(1.13)	(1.37)	(1.24)	(2.48)
Emotion: Hopelessness	1.331	-0.939	-1.399	6.440**
	(1.65)	(2.07)	(1.95)	(3.13)
Extraversion	-0.150	-0.554	-0.351	0.219
	(0.37)	(0.50)	(0.42)	(0.77)
Conscientiousness	0.254	-0.087	0.149	0.565
	(0.46)	(0.64)	(0.55)	(0.86)
Openness	0.152	-0.547	-0.380	1.280
	(0.41)	(0.51)	(0.46)	(0.85)
Neuroticism	-0.314	-0.596	-0.246	-0.739
	(0.49)	(0.66)	(0.56)	(0.98)
Agreeableness	-0.051	0.230	0.206	-0.314
	(0.50)	(0.69)	(0.59)	(0.92)
Anxiety State	0.167	0.434	0.462*	-0.392
	(0.23)	(0.30)	(0.27)	(0.42)
Anxiety Trait	-0.214	-0.509*	-0.345	-0.060
	(0.22)	(0.30)	(0.25)	(0.45)
Reappraisal	0.082	3.843**	3.912**	-4.838*
	(1.45)	(1.88)	(1.73)	(2.74)
Suppression	-0.499	-3.878*	-2.382	0.770
	(1.62)	(2.11)	(1.86)	(3.46)
Observations	290	166	205	85

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the decision in the main task (the number of points allocated to Box 2). In every specification, I control for the Bayesian benchmark.

Table 60: Emotions and decisions in the treatment condition.

	Signals: 1-4	*Signals: 1-5*	*Signals: 6-10*	*Signals: 7-10*
	(1)	(2)	(3)	(4)
Emotion: Enjoyment	-1.638	-0.704	0.263	0.022
	(1.53)	(1.33)	(1.40)	(1.60)
Emotion: Hope	2.293	2.276	-3.378*	-0.596
	(2.00)	(1.70)	(1.77)	(2.28)
Emotion: Pride	2.864	0.588	2.161	5.214*
	(2.28)	(1.87)	(2.03)	(2.67)
Emotion: Relief	-1.024	-0.736	2.233	1.141
	(2.14)	(1.90)	(1.97)	(2.56)
Emotion: Anger	1.421	0.901	-3.340**	-4.255**
	(1.93)	(1.44)	(1.36)	(1.64)
Emotion: Anxiety	-1.135	-0.502	3.489	7.711*
	(2.86)	(2.27)	(3.05)	(3.95)
Emotion: Shame	2.256	-0.187	1.156	0.484
	(1.88)	(1.63)	(1.56)	(1.90)
Emotion: Hopelessness	2.878	-0.136	2.965	3.563
	(2.69)	(2.10)	(2.54)	(3.32)
Extraversion	0.350	0.119	-0.134	0.411
	(0.55)	(0.48)	(0.56)	(0.64)
Conscientiousness	0.009	0.453	0.054	0.337
	(0.75)	(0.66)	(0.61)	(0.74)
Openness	-0.401	-0.140	0.483	0.292
	(0.65)	(0.55)	(0.58)	(0.70)
Neuroticism	-0.843	-0.625	-0.106	0.893
	(0.96)	(0.70)	(0.65)	(0.85)
Agreeableness	0.989	0.898	-0.658	-1.398*
	(0.94)	(0.74)	(0.65)	(0.83)
Anxiety State	0.365	0.403	-0.222	-0.195
	(0.38)	(0.33)	(0.32)	(0.37)
Anxiety Trait	-0.537	-0.448	0.025	0.314
	(0.36)	(0.30)	(0.32)	(0.39)
Reappraisal	-0.096	1.231	-0.676	-1.291
	(2.53)	(2.06)	(2.04)	(2.54)
Suppression	-1.619	1.371	-2.084	1.496
	(2.70)	(2.09)	(2.43)	(2.99)
Observations	106	144	146	107

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the decision in the main task (the number of points allocated to Box 2). In every specification, I control for the Bayesian benchmark.

Table 61: Emotions and decisions in the treatment condition (lasso procedure).

	All	*Exp*	*Exp+*	*Unexp*
	(1)	(2)	(3)	(4)
Emotion: Anxiety	2.427**			5.294*
	(1.02)			(2.77)
Emotion: Hopelessness				4.353*
				(2.61)
Reappraisal		2.295	2.650*	
		(1.59)	(1.46)	
Observations	290	166	205	85

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the decision in the main task—the number of points allocated to Box 2. The independent variables in every sample (shown in columns) have been selected by the lasso procedure. In every specification, I control for the Bayesian benchmark.

Table 62: Emotions and decisions in the treatment condition (lasso procedure).

	Signals: 1-4	*Signals: 1-5*	*Signals: 6-10*	*Signals: 7-10*
	(1)	(2)	(3)	(4)
Emotion: Hope	2.284	2.814**		
	(1.48)	(1.24)		
Emotion: Anxiety			6.068***	7.070***
			(1.33)	(1.66)
Observations	106	144	146	107

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the decision in the main task—the number of points allocated to Box 2. The independent variables in every sample (shown in columns) have been selected by the lasso procedure (the first variable selected). In every specification, I control for the Bayesian benchmark.

G.1.1 Emotions and decisions: unexpected signals

In Tables 63 and 64, I gather the results of a regression analysis that correlates decisions in the main task with the relevant variables from the previous section. The sample includes all participants in the treatment condition for whom I have the data on emotions. In all regressions, the dependent variable is the decision in the treatment condition: the number of points allocated to Box 2. It is regressed on a variable returned in the Lasso procedure (hopelessness, anxiety, hope, or reappraisal), a dummy variable indicating an unexpected signal, a dummy variable indicating a "good" signal (defined as in the main text), and their interactions. For every emotion, I present two specifications: one that includes a triple interaction and one without. In every specification, I control for the Bayesian benchmark. In Table 63, I show the results for two negative anticipatory emotions (anxiety and hopelessness), whereas Table 64 gathers the results for a positive anticipatory emotion (hope) and reappraisal.

The results show a strong effect of negative anticipatory emotions on decisions after unexpected signals. An increase in anxiety by 1 point (on a scale from 1 to 7) is correlated with allocating 6.8 points more to Box 2. This effect is mostly driven by unexpected "bad" signals—the coefficient at the triple interaction goes in the opposite direction, but is not significant. A similar relationship emerges for hopelessness. An increase by 1

Table 63: Emotions and decisions: expected vs unexpected signals.

	(1)	(2)	(1)	(2)
Good Signal	13.301**	10.752*	10.542**	8.498
	(5.55)	(5.75)	(4.82)	(5.38)
Unexpected Signal	-10.590*	-13.351**	-8.799	-10.515*
	(6.39)	(6.59)	(5.53)	(5.88)
Good × Unexp	-11.781*	9.807	-12.556*	-5.911
	(6.53)	(14.79)	(6.50)	(10.11)
Emotion: Anxiety	1.667	1.247		
	(1.38)	(1.40)		
Good × Anxiety	-2.315	-0.998		
	(2.35)	(2.48)		
Unexp × Anxiety	6.824***	8.112***		
	(2.33)	(2.45)		
Good × Unexp × Anxiety		-12.332		
		(7.59)		
Emotion: Hopelessness			0.641	0.287
			(1.39)	(1.45)
Good × Hopeless			-1.319	0.005
			(2.35)	(2.81)
Unexp × Hopeless			7.303***	8.385***
			(2.20)	(2.54)
Good × Unexp × Hopeless				-4.398
				(5.12)
Observations	290	290	290	290

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the decision in the main task (the number of points allocated to Box 2). "Unexpected Signal" is dummy variable indicating a signal far from individual belief distribution (more than 1 rank away from the beginning or the end of distribution). In every specification, I control for the Bayesian benchmark.

point (on a scale from 1 to 7) is correlated with allocating 7.3 more points to Box 2. The results are robust to controlling for the signal value. At the same time, hope is positively correlated with decisions about "good" signals that were expected by subjects, as shown in Table 64. The effect of reappraisal is negative and significant after unexpected signals. Participants who tend to use more cognitive reappraisal in their daily life also report lower probability after unexpected "bad" signals. I note that subjects who received an unexpected signal and those who received a close signal do not differ in the average level of anxiety, hopelessness, or reappraisal (see Table 65). While there is a difference in anxiety and hopelessness between people who received a "good" vs a "bad" signal (with

Table 64: Emotions and decisions: expected vs unexpected signals.

	(1)	(2)	(1)	(2)
Good Signal	-14.457	-15.713	9.961	4.598
	(9.69)	(10.96)	(12.87)	(14.85)
Unexpected Signal	19.847**	18.582	35.400***	30.401**
	(10.03)	(11.29)	(12.73)	(14.48)
Good × Unexp	-14.732**	-9.336	-13.961**	6.329
	(6.55)	(22.90)	(6.60)	(28.72)
Emotion: Hope	-1.785	-1.880		
	(1.26)	(1.32)		
Good × Hope	4.935**	5.206**		
	(1.97)	(2.26)		
Unexp × Hope	-3.443*	-3.179		
	(1.94)	(2.22)		
Good × Unexp × Hope		-1.141		
		(4.64)		
Reappraisal			2.804	2.390
			(1.81)	(1.90)
Good × Reapp			-0.339	0.868
			(2.80)	(3.26)
Unexp × Reapp			-7.289***	-6.179**
			(2.72)	(3.12)
Good × Unexp × Reapp				-4.645
				(6.40)
Observations	290	290	290	290

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: The dependent variable is the decision in the main task (the number of points allocated to Box 2). "Unexpected Signal" is dummy variable indicating a signal far from individual belief distribution (more than 1 rank away from the beginning or the end of distribution). In every specification, I control for the Bayesian benchmark.

Table 65: Difference between participants who received different signals.

	Expected	**Unexpected**		Diff < 0	Diff $\neq 0$	Diff > 0
Emotion: Anxiety	2.051	1.976	*p-value:*	0.669	0.661	0.331
	(0.094)	(0.136)				
Emotion: Hopelessness	1.639	1.600	*p-value:*	0.593	0.814	0.407
	(0.088)	(0.144)				
Reappraisal	4.465	4.339	*p-value:*	0.822	0.356	0.177
	(0.072)	(0.121)				
N	205	85				

Note: Standard errors in parenthesis. "Expected" denotes a group of participants who received a signal within or close to their prior belief distribution. "Unexpected" denotes participants who received a signal far from their priors (defined as before).

the later group experiencing more negative emotions), within a "bad" ("good") signal group, the emotion levels are no different for expected and unexpected signals.

The results presented in this section bring us two important insights. First, the anticipatory emotions are correlated with belief formation. Second, the relations between the anticipatory emotions and decisions are very different in the case of expected and unexpected signals. The decisions after unexpected signals are strongly correlated with the negative anticipatory emotions: anxiety and hopelessness, and these relations that are mostly driven by unexpected "bad" signals. One interpretation is that, when an agent receives a "bad" signal that he did not expect, he might be using his negative emotions as additional information—an "internal" signal that he might have performed worse than he expected. If an agent is able to regulate his emotions by using reappraisal, he will form a less pessimistic belief after an unexpected signal. The results emphasize the role of expectations in belief formation.