Inattentive Inference*

Thomas Graeber

June 23, 2022

Abstract

This paper studies how people infer a state of the world from information structures that include additional, payoff-irrelevant states. For example, learning from a customer review about a product's quality requires accounting for the reviewer's otherwise-irrelevant taste. This creates an attribution problem common to all information structures with multiple causes. We report controlled experimental evidence for pervasive overinference about states that affect utility – a form of "omitted variable bias" in belief updating –, providing an explanation for various misattribution patterns. In studying why systematic misattribution arises, we consistently find that errors are not due to deliberate effort avoidance or a lack of cognitive capacity. Instead, people behave as if they form incomplete mental models of the information structure and fail to notice the need to account for alternative causes. These mental models are not stable but context-dependent: misattribution responds to a variety of attentional manipulations, but not to changes in the costs of inattention.

JEL Classification: C91, D01, D83, D84

Keywords: Belief Formation, Attention, Bounded Rationality

*Financial support from the Sloan Foundation and the Institute on Behavior and Inequality (briq) is gratefully acknowledged. Graeber: Harvard Business School, tgraeber@hbs.edu.

1 Introduction

The difficulty of attending to, aggregating and processing the abundance of available information in practice motivates a strand of work on errors in belief formation. For example, people may be partially inattentive to information (Bartoš et al., 2016; Caplin and Dean, 2015; Enke, 2020; Hanna et al., 2014; Malmendier and Lee, 2011; Sims, 2003) or fail to account for the relationship between different signals (Enke and Zimmermann, 2019; Eyster and Rabin, 2010; Levy and Razin, 2015). In some situations, however, the amount of information is manageable in principle and agents are capable of attending to all available pieces of information. Rather than selecting or aggregating these signals, the challenge of belief formation then often lies in selecting the right interpretation of a piece of information. In this case, the agent faces an *attribution* problem as he may struggle to figure out what a given piece of information actually means. Rather than the first type of environment with *many signals*, this paper studies attribution problems in information structures with *many causes* for a single signal. To take a stylized example, suppose that a shopper reads a positive customer review that is a function of actual product quality and the reviewer's personal taste. Learning from a positive review about underlying quality requires accounting for other, extraneous causes in the information structure, such as differing tastes. A failure to account for alternative causes creates misattribution to the causes of interest, a form of "omitted variable bias" in belief formation. For example, a decision-maker who does not factor in the role of varying tastes over-attributes a positive review to high product quality.

A collection of separately documented empirical findings is suggestive of this type of error. The defining pattern is excessive inference about a specific cause of interest, while neglecting alternative causes that are "nuisance" from the decision maker's perspective. For example, CEOs and politicians are rewarded for luck because performance evaluations and voter support partly fail to condition on external conditions such as the business climate (Bertrand and Mullainathan, 2001; Wolfers, 2002). People overstate the role of intentions relative to contextual factors and chance when explaining the behavior of others (Gurdal et al., 2013; Ross, 1977), known as the fundamental attribution error in psychology. Applied work on attention shows that people often underreact to certain elements of the price structure such as sales taxes when learning from a price about the value of a good (Abaluck and Gruber, 2011; Allcott, 2011; Chetty et al., 2009; Taubinsky and Rees-Jones, 2018). When explaining the world, we tend to narrowly focus on the determinants that matter most to us, which may result in us attributing excessive causal power to them.

This paper tackles two questions. First, how do people learn about a target state of the world from information that also depends on otherwise irrelevant states? In simple, tightly controlled updating experiments we document a systematic neglect of nuisance causes and misattribution to causes of interest. We validate the generalizability of the key finding using a naturalistic variant of the experiments that exploits an economically relevant situation and does not rely on explicit computations. Second, why does such misattribution arise? We examine this question by leveraging the distinction between "frictions" and "mental gaps" (e.g., Handel and Schwartzstein, 2018).[1] Frictions are directly linked to the costs of information processing, and may occur due to mental processing noise, capacity constraints or other forms of attentional limitations (Caplin and Dean, 2015; Gabaix, 2014; Matějka and McKay, 2015; Sims, 2003; Woodford, 2019). A mental gap describes the divergence between how people think about a problem and how they should think about it given costs. People sometimes appear to form incorrect mental models or problem representations.[2] The data from more than twenty experimental treatments that examine the cognitive mechanisms underlying the neglect of alternative causes consistently point to a mental gap: subjects are unaware of their neglect and minor attentional manipulations successfully debias respondents, whereas variations of the costs and benefits of attention have little to no effect.

We present causal evidence from laboratory and online experiments. We use a two-pronged approach with two complementary paradigms: the baseline experiment strips away the real-world context and associated ambiguities to create a maximally controlled belief updating setting, which comes at the cost of a potential lack of realism. The complementary set of vignette experiments replicates the results in settings with naturalistic task framing that is closer to real-world inference problems, but sacrifices some of the control obtained in the former. In the baseline condition of the laboratory experiment, treatment *Narrow*, subjects guess an unknown, random state of the world and are paid for accuracy. Before indicating their guess, they receive a piece of information (the *signal*) that depends both on that target state and another unobserved state. Specifically, two numbers X and Y are drawn from known distributions. In this baseline condition, subjects have to guess X, but not Y. Because Y is not a prediction target, it constitutes a *nuisance variable* from

[1]This taxonomy is representative of a collection of related classifications put forward in the literature, such as that of "bounds errors" versus "astray errors" (Rabin, 2013).

[2]See, e.g., Bordalo et al. (2020a); Enke (2020); Eyster and Rabin (2010); Gagnon-Bartsch et al. (2019); Gennaioli and Shleifer (2010); Hauser and Schwarz (2016); Jehiel (2005); Schwartzstein (2014); Spiegler (2016, 2017).

the subject's perspective: it confounds information about the target variable X, but its realization does not affect his payoff given a stated belief.[3] In a typical task, X is drawn from the simple discretized uniform distribution on $\{30, 40, 50, 60, 70\}$ and Y is drawn at random from $\{10, 20, 30, 40, 50, 60, 70, 80, 90\}$. Subjects observe a signal that depends on both states, such as the average of the drawn numbers, $S = \frac{X+Y}{2} = 70$. Crucially, inference from S about X requires accounting for the random variation in S that is due to Y. In the context of the previous example, a shopper might want to infer unobservable product quality (X) from an observable customer review (S), which is a function of both quality X and the reviewer's tastes Y. Failing to properly account for the stochasticity of Y generates misattribution of the signal to X. Subjects are informed of the simple data-generating process and the signal structure, eliminating all structural uncertainty in the information environment. We confirm that subjects are not confused about the task setup using an extensive set of control questions; we always show all relevant information on the decision screen; and we run additional control treatments to address potential misunderstandings. In this baseline condition, where subjects are incentivized to state the full distribution of their belief about X but not about Y, beliefs about X exhibit pervasive neglect of the nuisance variable Y. In the numerical example above, this is equivalent to stating that $X = 70$ with certainty, as if $S = X$. The Bayesian posterior belief about X, by contrast, assigns equal probability to 50, 60 and 70. Across all tasks, only 17% of all stated beliefs are in line with the Bayesian benchmark, whereas 62% display full neglect of Y. We refer to this as nuisance neglect and conceptualize its relationship to other forms of bias below.

In a baseline control treatment, *Broad*, a separate set of subjects is incentivized to guess the joint distribution of X and Y, rather than only X. This turns Y from a nuisance into a target variable, while keeping the overall monetary stakes as well as the objective updating problem (and thus the Bayesian posterior) exactly identical to the baseline condition. Because the information structure is unchanged, the complexity and cost of computing a posterior for X should be unchanged. Note that this treatment is a control condition that alleviates a shortcoming of recent experimental work on belief formation, because it manages to hold the objective updating problem constant across conditions.[4]

[3]We define a nuisance variable in Section 3.1.2 as one whose realization does not affect utility conditional on an action.

[4]Experimental work on belief updating routinely compares beliefs in information environments with and without a feature of interest. A manipulation of the signal structure, however, can confound the analysis if it affects other properties of the updating problem, such as the complexity of forming an update (as in, for example, Enke and Zimmermann, 2019).

In treatment *Broad*, we document a large and statistically significant treatment difference relative to treatment *Narrow*. Moreover, the median belief in *Broad* is indistinguishable from the Bayesian posterior, implying that the experimental setup is not too complex per se and subjects are in principle able to solve the task correctly. More than 70% of all stated beliefs in *Broad* correspond to the Bayesian posterior.

We examine the external validity of these findings using a set of naturalistic vignette experiments that leverage real-world scenarios, do not have the character of a math problem, include an a application with economic relevance and a variant featuring a simple choice instead of a belief incentivized with a complex scoring rule. Next to the vignette experiments, a battery of laboratory and online experiments (i) tests the robustness of nuisance neglect by varying various elements of the experimental design, such as the specific signal structure (e.g., a signal outside of the variables support), the distribution of the random states (non-uniform distributions), and the elicitation procedure,[5] (ii) documents nuisance neglect in a large and heterogeneous online population; and (iii) tests the predictions of existing theories of belief formation in this setup. Specifically, we design sharp tests of different models using systematic variations of the data and signal structures. We find that the pattern of neglect of nuisance variables in the data is not consistent with overweighting the signal (Benjamin, 2019, for a review of overinference), underweighting the base rate (Bar-Hillel, 1980; Grether, 1980) or diagnosticity-based theories of expectation formation (Bordalo et al., 2018).

The second part of the paper studies *why* nuisance neglect arises by investigating the underlying cognitive mechanisms. We adopt the distinction between frictions and mental gaps as an instructive taxonomy for the present application: the neglect of nuisance variables may be due to the (computational) difficulty of accounting for the nuisance variable Y in conjunction with X – a friction – or due to a failure to recognize the necessity to take into account Y to begin with – a form of misconstrual or mental gap.

To examine these explanations, we design additional experiments that test for a potential mental gap. We present a series of additional experiments that aim to manipulate how people think about the updating task, while keeping the cost of accounting for Y constant. If drawing people's attention to the role of Y without changing the updating problem affects the degree of nuisance neglect, a mental gap is likely to play a role.

We present three main findings from the analysis of mental gaps. First, nuisance ne-

[5]For example, we disentangle the elicitation procedure from prediction incentives by having subjects state the joint distribution when only X is incentivized – unlike in *Narrow* –, or by having them state a marginal belief about X first when both variables are incentivized – unlike in *Broad*.

glect is reduced substantially and beliefs are pre-dominantly Bayesian once attention is drawn to Y. A contextual cue to attend to Y is sufficient to reduce nuisance neglect, while maintaining Y's role as a nuisance variable and holding constant the difficulty of accounting for it. In treatment *Hint*, subjects only guess X but see an additional verbal statement on each elicitation screen: "Also think about the role of Y." The hint produces a large and statistically significant treatment difference relative to the baseline condition *Narrow*.

Second, while the exogenous manipulations of attention have the potential to debias, we find that subjects are able to overcome nuisance neglect on their own when nudged to reconsider their solution strategy. In treatment *Enforced Deliberation*, we implement a thirty-second deliberation time on the elicitation screen before the input fields are activated. The objective is to encourage subjects to deliberate their problem interpretation *before* they form their posterior. Enforced deliberation time substantially reduces nuisance neglect and is roughly half as effective as an explicit hint.

Note that the effect of minor attentional manipulations is striking in the sense that even in condition *Narrow*, all relevant pieces of information are displayed on the screen and we ensure that subjects are not confused by the setup. However, they may still fail to realize the necessity to account for the variation of Y to begin with and would consequently be *unaware* of committing an error. In a third step, we test this lack-of-awareness hypothesis directly by measuring confidence in beliefs using incentivized willingness-to-pay (WTP) to have a guess replaced by an optimal guess. Exploiting causal variation, we find that nuisance neglect is associated with similar confidence levels as Bayesian updating, indicating that subjects are unaware of the neglect.

These three findings consistently suggest an underlying mental gap: Attentional manipulations that plausibly hold the cost of information processing fixed close the mental gap of failing to attend to Y, which subjects seem to be unaware of to begin with.

In a companion exercise, we empirically investigate the friction mechanism for nuisance neglect. Why do people systematically neglect elements of an information structure, even in simple contexts? One candidate explanation is that such model simplifications reflect a strategy to economize on cognitive costs.[6] We report two main findings on the applicability of this cognitive cost-benefit perspective to our setting. First, we find that increasing the stake size tenfold (in the laboratory) or fivefold (in online experiments)

[6]A prominent view in cognition research holds that humans are "cognitive misers" who continuously seek strategies to avoid thinking (Fiske and Taylor, 2013). Similarly, a large class of models in economics relies on weighing the expected benefits against the cognitive costs of attention (Caplin and Dean, 2015; Gabaix, 2014), prominently including theories of rational inattention (Sims, 2003, 2006).

substantially increases effort as measured by response times, but does *not* reduce the prevalence of nuisance neglect, at odds with an underlying lack of effort. Second, we directly test for the presence of cost-benefit considerations by manipulating the specific monetary loss incurred from committing nuisance neglect, based on how much "noise" and "bias" the presence of Y introduces into the posterior of an agent who mistakenly updates as if $S = X$. Strikingly, the presence of nuisance neglect does *not* respond to the monetary loss associated with its expected (in)accuracy.

Taken together, the analysis of mechanisms suggests that nuisance neglect occurs when subjects do not mentally account for Y to begin with. They are unaware of this omission, and it does not reflect a lack of effort. People seem to initially "fail to notice" the necessity of accounting for the variation in Y, which may lead them to form a misspecified problem representation. Attentional cues that nudge subjects into re-considering the problem (conditions *Enforced Deliberation* and *Hint*) improve updating substantially. The combined evidence is more consistent with a mental gap interpretation of misattribution.

The paper proceeds as follows. Section 2 embeds the paper in the existing literature. In Section 3, we present the baseline design and results from the laboratory and online experiments, as well as extensions that include a replication in a naturalistic context and robustness exercises. In Section 4, we examine why nuisance neglect occurs based on the distinction between mental gaps and cost-benefit considerations. Section 5 concludes.

2 Related Literature

The paper contributes to several literatures. In the experimental literature, this study of misattribution in the basic case of interpreting a *single* piece of information complements recent work on updating errors in situations that require the joint processing and aggregation of *many* pieces of information (Enke, 2020; Enke and Zimmermann, 2019). Other related work highlights failures of hypothetical thinking (Esponda and Vespa, 2014, 2019; Martínez-Marquina et al., 2019) and the failure to notice important features of the available data (Hanna et al., 2014). Benjamin (2019) reviews a voluminous body of empirical research on probabilistic reasoning. His meta-study concludes that beliefs often tend to be less sensitive to variation in problem parameters – such as the base rate, diagnosticity and sample size – than postulated by Bayes' rule. That people do respond to parameters albeit too little differs from the type of discrete neglect of a part of the signal structure documented here. Moreover, we show that subjects do not follow a compelling intuition when committing nuisance neglect, which underlies many judgment errors studied in the heuristics and biases literature (Kahneman and Tversky, 1982; Morewedge and Kahneman, 2010; Tversky and Kahneman, 1983). Finally, this paper contributes a new perspective to the long-standing debate on the conditions for overreaction versus underreaction to information (Adam et al., 2017; Bordalo et al., 2020b; Coibion and Gorodnichenko, 2012, 2015; Frydman and Nave, 2016; Greenwood and Shleifer, 2014; Landier et al., 2017).[7] Nuisance neglect simultaneously generates overreaction to payoff-relevant causes and underreaction to nuisance causes, providing testable predictions on their relative likelihood of occurrence.

In studying *why* updating errors occur, our findings on the source of nuisance neglect in attribution problems square with those of Enke (2020), who finds that people sometimes narrowly focus on visible parts of the information structure in signal aggregation tasks. Enke (2020) argues that people form simplified mental models of a problem that respond to the computational complexity of a task. Comparable findings in the updating environments studied here hint at a common cognitive mechanism underlying belief errors in both signal aggregation and attribution problems: an unwitting neglect of parts of the structure of updating problems. While Enke (2020) shows that this neglect can be context-driven by varying which signals are visible, the present paper highlights a different channel: heuristic model simplifications may often be determined by people's incentive structure, irrespective of which parts of the information structure are visible.

On the applied side, this paper speaks to a collection of separately documented misattribution patterns. One line of work starting with Chetty et al. (2009) shows inattention to specific features of the decision context (Abaluck and Adams, 2017; Abaluck and Gruber, 2011; Allcott, 2011; Taubinsky and Rees-Jones, 2018). For example, Taubinsky and Rees-Jones (2018) find that people underreact to sales taxes. While their experiment does explicitly pose an inference problem, the results are consistent with consumers systematically overinferring from price (S) about the value of the product (X) while neglecting the sales tax (Y). Other phenomena that can be interpreted through the lens of nuisance neglect are outcome bias in punishing decisions that are based on luck (Y) rather than effort (X) alone (Brownback and Kuhn, 2019; Gurdal et al., 2013); in consumer choice, there is misattribution of positive experience to the intrinsic value of an outcome (X) while ne-

[7] Championed by Kahneman and Tversky and prominent in finance is the view that beliefs move too much (De Bondt and Thaler, 1985; Bordalo et al., 2019; Shiller, 1981; Tversky and Kahneman, 1971), while an older psychology literature and the dominant view in macroeconomics maintains that beliefs tend to move too little (Benjamin, 2019; Edwards, 1968; Mankiw and Reis, 2002; Rabin and Schrag, 1999).

glecting reference-dependent surprise (Y) (Bushong and Gagnon-Bartsch, 2018) and to the quality of a consumption good (X) rather than a contextual state such as the weather (Y) (Haggag et al., 2018); and in social learning contexts, people overinfer about a person's private information (X) from their action, neglecting that the action also embeds private information from earlier movers (Y) (Eyster et al., 2018). In much of this work, inattention specifically occurs to problem features that are plausibly nuisance variables. At the same time, findings from previous experimental work on environments with many signals do not apply to these settings (e.g., Bartoš et al., 2016; Enke and Zimmermann, 2019).

Research in cognitive science has studied related phenomena that speak to the external validity of our results. The "causal frame problem" shows that people often form incomplete causal models of a problem. Work on biases in causal reasoning finds that people employ cognitive shortcuts that can result in the neglect of alternative causes (Fernbach et al., 2010; Fernbach and Rehder, 2013; Sloman and Lagnado, 2015). This body of work indicates that the findings from the highly controlled but stylized experimental environments studied here carry over to environments with a more naturalistic task framing.

Finally, the paper speaks to a large theoretical literature. Work on the rational inattention paradigm focuses on rational information acquisition given cognitive capacity constraints or processing costs. Rational inattention models do not generate systematic misinference conditional on processing a piece of information, as they posit Bayesian inference from the information that an agent actually attends to (Caplin and Martin, 2015; Caplin et al., 2020; Matejka and McKay, 2014; Sims, 2003; Wiederholt, 2010). The discrete neglect of certain dimensions in the data is reminiscent of the sparsity-based model of Gabaix (2014), applied to belief updating. The lack of responsiveness to variation in costs and benefits, however, is at odds with sparse maximization. A key characteristic of the combined evidence is that people appear to form inaccurate mental representations of problems because they are looking at the problem the wrong way, rather than trading off the benefits and costs of more accurate representations. This appears more compatible with frameworks of mental gaps than models of rational inattention. Inaccurate priors may lead to self-serving misattribution (Hestermann and Yaouanq, 2020) or discrimination (Chauvin, 2020) but are unlikely to be at play here because priors are controlled experimentally. The most closely related theoretical frameworks view incomplete representations as reflecting incorrect beliefs about which variables matter (Gagnon-Bartsch et al., 2019; Schwartzstein, 2014). All of these models share the prediction that representations should look fairly consistent across problems. The evidence in this paper highlights that they miss how heuristic model simplifications may often not be stable but constructed *on-the-fly* in response to task demands, environmental cues and even suggestions to reconsider a representation.

3 Evidence for Nuisance Neglect

3.1 Baseline Experiments

To causally examine the role of nuisance variables in information structures for belief updating, the experimental design aims to satisfy the following requirements: (i) a fully controlled and transparent data-generating process and information structure that is known to subjects, (ii) an experimental manipulation of the presence of nuisance causes, (iii) limited complexity to minimize confusion, and (iv) an incentive-compatible procedure to extract beliefs.

3.1.1 Design

Experimental variation in the presence of nuisance causes can be achieved by changing the information structure, but this also affects the complexity of updating beliefs across conditions. We instead design a simple setting that implements this variation without changing the information structure or data-generating process. The basic updating task features two unobserved random numbers X and Y, generated by stochastic processes known to subjects. To simplify, these numbers are independently drawn from two discrete uniform distributions with small sample spaces. Subjects receive a signal $S = s$ on the two unknown draws: depending on the task, they see either the sum or the average of the two numbers.[8] The signal structure maps two inputs, i.e., the realizations of random variables X and Y, to a one-dimensional output, i.e., the observed signal s. The experiment creates exogenous between-subject variation in whether the agent's payoff depends on the realizations of only one or both of the inputs in the signal structure. There are two experimental conditions: In *Narrow*, subjects are paid to guess only X, while in *Broad*,

[8]In the baseline tasks, the signal is an unbiased estimator of the mean of X. Either subjects receive the average of the drawn numbers and the prior distributions of X and Y have identical means, or they see the sum of the drawn numbers and Y has a mean of zero. We study more general signal structures from Section 3.2.

subjects are paid to guess both X and Y. The experimentally controlled prior, the signal structure, and the Bayesian posterior are identical in *Narrow* and *Broad*. A Bayesian agent thus forms identical beliefs in both conditions. By randomly choosing only one of the guesses in *Broad* for payment, the size of the monetary incentive is kept constant.

Table 1: Overview of baseline task specifications

Sample space of X	Sample space of Y	Signal structure	Signal realization
30, 40, 50, 60, 70	10, 20, 30, 40, 50, 60, 70, 80, 90	$(X+Y)\div 2$	60
230, 240, 250, 260, 270	210, 220, 230, 240, 250, 260, 270, 280, 290	$(X+Y)\div 2$	230
180, 190, 200, 210, 220	180, 190, 200, 210, 220	$(X+Y)\div 2$	200
80, 90, 100, 110, 120	-30, -20, -10, 0, 10, 20, 30	$X+Y$	80
130, 140, 150, 160, 170	-25, -15,-5, 0, 5, 15, 25	$X+Y$	165

Notes: Overview of the five baseline belief tasks in the laboratory study. The distributions of X and Y as well as the signal structure are identical in both treatment conditions. X and Y are independently drawn from two discrete uniform distributions, i.e., every indicated outcome is equally likely. In the baseline study, all subjects received the same (random) signal realization. In the complementary online experiments, signal realizations were drawn at the subject level.

Subjects complete the five updating tasks of Table 1 in random order without receiving feedback in between. For example, in the first task of Table 1, X is one of five numbers, 30, 40, 50, 60 or 70 with equal probability, while Y is independently drawn with equal probability from 10, 20, 30, 40, 50, 60, 70, 80 and 90. Subjects learn that the average of X and Y is 60 and then state their belief as described in detail in Section 3.1.3.[9] To solve this task, subjects need to identify all (X, Y) combinations with an average of 60, that is (30, 90), (40, 80), (50, 70), (60, 60), (70, 50). Both numbers being drawn uniformly and independently, it follows that each of these outcomes is equally probable. Additional task specifications and treatment variations address the robustness of the baseline results and examine the nature of updating rules (see Section 3.2). For example, we replicate our findings with more general data and signal structures, for example when the signal falls outside of the support of a variable.

A key feature of this design is that unlike related empirical studies of updating errors (Caplin et al., 2011; Dean and Neligh, 2019; Enke, 2020; Enke and Zimmermann, 2019), this experimental setup holds the information structure fixed across conditions.

[9]Note that in the baseline study, all subjects received the same (random) signal realization. In complementary online experiments, signal realizations were drawn at the subject level.

3.1.2 Definition of Nuisance Variables and Predictions

To fix ideas, we delineate basic concepts underlying the baseline treatment comparison in a setting that loosely follows Gabaix (2019). Assume an agent (he) who states a belief $b \in \mathbb{R}$ about the two-dimensional vector of random states in the updating task, $(x, y) \in \mathbb{R}^2$. Without loss of generality, we normalize $\mu_X = \mu_Y = 0$, such that (x, y) denote deviations from their respective means. The agent chooses b to maximize linear-quadratic utility after observing a signal s about the random draw (x, y). Crucially, the signal is generated by a deterministic function of *both* random variables, i.e., $s = f(x, y)$, with $\frac{\partial s}{\partial x} \neq 0$ and $\frac{\partial s}{\partial y} \neq 0$. The utility function

$$u(b, x, y) = -\frac{1}{2}\left(b - \eta_x x - \eta_y y\right)^2 \tag{1}$$

yields the following optimal belief:

$$b^r(s) = \max_b \ \mathrm{EU}|s = \max_b \ \mathbb{E}\left[-\frac{1}{2}\left(b - \eta_x x - \eta_y y\right)^2 |s\right] \tag{2}$$

$$= \mathbb{E}[\eta_x x + \eta_y y|s] = \eta_x \mathbb{E}[x|s] + \eta_y \mathbb{E}[y|s]. \tag{3}$$

This optimal belief is a function of the Bayesian conditional posterior expectation of x, $\mathbb{E}[x|s]$, and y, $\mathbb{E}[y|s]$, as well as weight parameters (η_x, η_y) that reflect how strongly the agent's utility depends on the realization of each variable. The definition of a nuisance variable directly follows from the weight parameters.

Definition. *$Z \in \{X, Y\}$ is a nuisance variable in an updating problem if its realization z does not affect the agent's expected utility conditional on a stated belief. Formally, $\forall b \in \mathbb{R}$: $\frac{\partial [EU|b]}{\partial z} = 0$. This is the case iff $\eta_z = 0$.*

Intuitively, the agent's expected payoff in a belief formation task does not respond to the realization of a nuisance variable for any given stated belief. A nuisance variable is payoff-irrelevant after stating a belief. Importantly, this does not mean that a nuisance variable is irrelevant for the agent's *optimal* belief b^r, which is clear from equation (3): even if $\mu_y = 0$, b^r mechanically depends on y through the conditional expectation $\mathbb{E}[x|s]$. $\mathbb{E}[x|s]$ is determined by the signal structure S that we assumed is a function of Y.[10] The definition of a nuisance variable highlights that optimal beliefs can depend on variables whose realizations are payoff-irrelevant conditional on a stated belief. This points to a

[10] Note that $\mathbb{E}[x|s] = \mathbb{E}[x \mid f(X, Y) = s]$.

crucial distinction between incentives provided through payoffs on the one hand, and the necessity of taking into account all elements of an information structure to form a Bayesian posterior on the other hand.

We now apply this idea to the treatment variation. In *Broad*, the agent is paid for accuracy of his joint posterior about (x, y), so that his utility depends on both realizations given a stated belief, i.e., $\eta_x \neq 0$ and $\eta_y \neq 0$. Thus neither X nor Y are nuisance variables. In *Narrow*, however, the agent's expected utility given a belief only depends on the realized state of X, but not on that of Y, i.e., $\eta_x \neq 0$ but $\eta_y = 0$.

Note, however, that since the Bayesian belief about (x, y) is independent of the prediction incentives, the treatment manipulation is designed in such a way that it is inconsequential under Bayesian updating. While we focus on a tightly controlled, stylized setting for the reasons outlined above, note that nuisance variables are readily identifiable in applied contexts: they are sources of stochasticity that are materially irrelevant to an agent beyond the necessity to account for them in an inference problem.

The thrust of the baseline prediction is that people neglect nuisance variables in the updating problem. A priori, this neglect of Y could take on a number of different forms. The decision maker may implicitly neglect the variance of Y, replace Y with a "default" value (as in Gabaix, 2014), or apply a particular non-Bayesian updating rule. We investigate the precise form of nuisance neglect using additional experimental variations, see Section 3.2.1. A candidate form of neglect is that the agent interprets the signal as if it only depends on X, but not Y. Nuisance neglect in condition *Narrow* may then be characterized by the agent taking the signal as fully revealing about X, *as if* generated by an alternative deterministic signal structure $\tilde{S} = g(X)$. The neglectful agent forms his belief based on a flawed posterior $P(X|\tilde{S} = s)$ instead of $P(X|S = s)$.

Prediction 1. *Beliefs exhibit nuisance neglect.*

(a) Beliefs in condition Narrow imply a neglect of Y. Specifically, subjects take the signal as fully revealing about X.

(b) Beliefs in condition Broad are Bayesian.

Prediction 1 directly implies a treatment difference between stated beliefs in conditions *Narrow* and *Broad*. The above simplistic notion of neglecting states that are payoff-irrelevant given a stated belief abstracts from the specific features of attention. It merely serves to set the stage for our in-depth analysis of the nature of attention and the more explicit framework of the origins of neglect. In particular, we will later argue and show that being a nuisance variable is not a sufficient condition for neglect in the inference problem and disentangle between endogenously chosen attention and exogenous attentional cues.

3.1.3 Procedures

Subjects in condition *Broad* guess the joint distribution of X and Y and are randomly paid for their accuracy in guessing either of these (decision screen in Appendix Figure 33). Subjects in condition *Narrow* only guess the marginal distribution of X (Appendix Figure 30).[11] The design unobtrusively obfuscates the study's objective: subjects receive their signal in encrypted form and have to decipher it using a simple two-step decoding protocol.[12] Note that no subject failed to implement the protocol. In a control treatment (*Simplification*, see also Appendix C.4) and all online experiments (Section 3.2.1), this feature was removed. The findings absent the obfuscation indicate that the obfuscation would not have been necessary in the baseline experiment. Each belief elicitation (excluding the deciphering stage) is subject to a five-minute time limit. The findings are robust to removing both the deciphering and the time limit (Section 3.2).

The elicitation procedure aims at providing a full characterization of subjective beliefs by having subjects indicate the entire posterior distribution instead of a point prediction. At the end, one of the tasks is randomly selected to be paid out based on the Binarized Scoring Rule with a prize of 10 euros (Hossain and Okui, 2013).[13] Subjects receive extensive instructions and have to complete eight control questions that test their understanding of the instructions, the data-generating process and signal structure, as well as the elicitation protocol (see Appendix G). In two unpaid practice tasks, subjects are trained to indicate a verbally described belief in a way that maximizes their payoff. This training stage is identical across treatments.

The belief updating problems are followed by a questionnaire. To shed light on cor-

[11]There is a treatment difference in the elicitation protocol, i.e., whether X and Y or only X is elicited. Additional treatment variations harmonize the elicitation protocol, i.e., subjects with *Narrow* incentives predict both X and Y, and subjects with *Broad* incentives predict first the marginal of X, and then the marginal of Y on a separate subsequent page. All main findings persist. See Section 3.2.

[12]Subjects see a sequence of letters. First, each letter has to be translated into a digit based on a decoding key displayed on the screen. Then the number 20 has to be added to the result. Subjects are familiarized with the deciphering process in the practice stage. See also the instructions in Appendix G.

[13]The scoring rule proposed by Hossain and Okui (2013) remains incentive compatible if subjects are risk averse. We adopt the approach suggested by Hossain and Okui (2013) to incentivize the entire stated distribution based on the sum of squared deviations between the probability mass allocated to each value of the distribution and the corresponding mass that should be allocated after learning the realized outcome. See Appendix G for further details.

relates of subject-level heterogeneity in belief formation, we measure performance on an incentivized test of cognitive capacity (10 Raven matrices, 0.2 euros per correct answer) and elicit a measure of risk preferences (Falk et al., 2016).

144 student subjects (72 in each treatment) participated in six sessions of the baseline experiment run at the University of Bonn's BonnEconLab in July 2017. Treatment status was randomized within session. We implemented the study in oTree (Chen et al., 2016). Mean earnings amounted to 11.40 euros – including a 5-euro show-up fee – for an average session duration of 57 minutes.

3.1.4 Results

We begin with an analysis of stated beliefs at the aggregate level before exploring their heterogeneity in Section 3.1.5. Figure 1 illustrates raw beliefs in each baseline task. It shows the sample average of stated belief distributions in both treatment conditions, alongside the Bayesian belief and the signal realization. The average subject in *Broad* forms beliefs that are closely aligned with the Bayesian posterior. In *Narrow*, by contrast, subjects on average assign too much probability mass to outcomes close to the signal value, as implied by inattention to Y.

Table 2 provides summary statistics and non-parametric tests by task. Median beliefs in *Narrow* (column 3) and *Broad* (column 4) closely correspond to the observed signal realization (column 1) and the Bayesian benchmark (column 2), respectively. Column 7 shows that belief distribution means and belief distribution variances are significantly different between treatments at the 0.1% level (M-W U tests).[14] Note that the median variance of stated distributions in *Narrow* is far too low, indicating that subjects hold too precise beliefs.

[14]This holds for all tasks except the distribution means in task (3), in which the signal realization coincides with the Bayesian posterior mean.

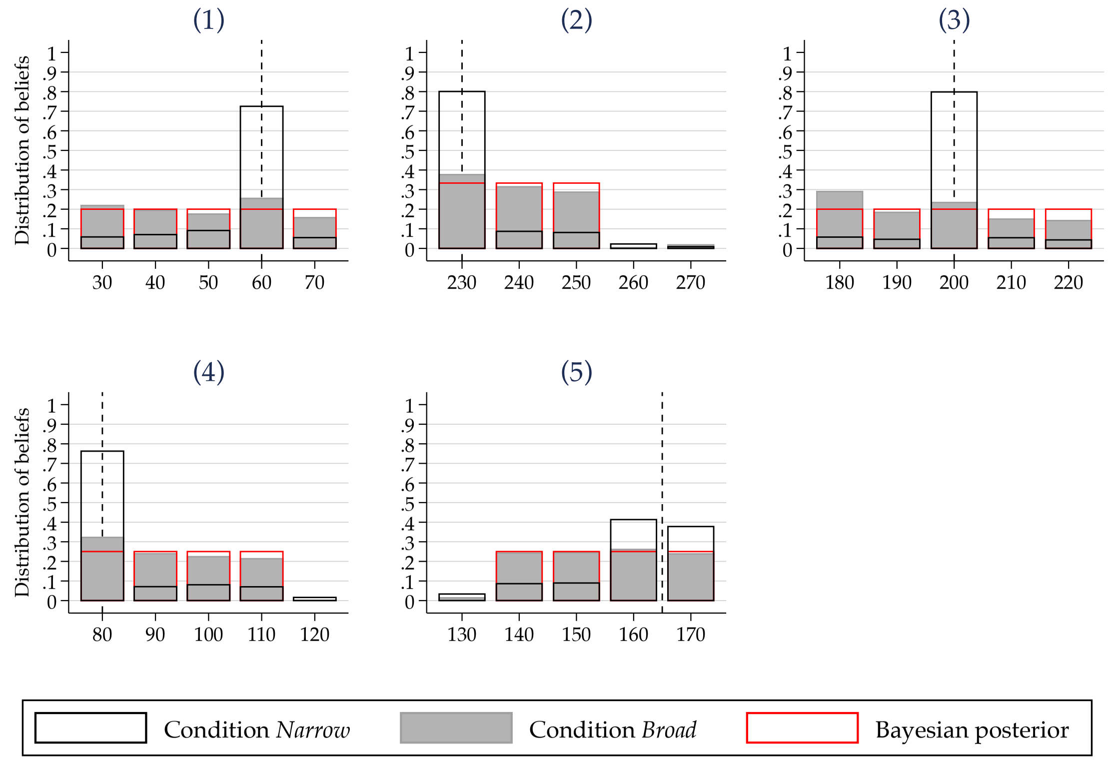


Figure 1: Treatment averages of stated belief distributions about X for each of five baseline tasks. N=72 for each treatment in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the vertical dashed line. In all five tasks, X and Y follow independent discrete uniform distributions. The task order was randomized at the subject level. The distributions and signal structure for each task is shown in Table 1. Subjects observed the mean of the drawn numbers in tasks (1), (2) and (3), and they saw the sum in (4) and (5).

Result 1. *Beliefs display nuisance neglect.*

(a) *The median belief in Narrow exhibits exact nuisance neglect, i.e.,* $P(X|X = s)$.

(b) *The median belief in Broad equals the Bayesian posterior.*

(c) *There are significant treatment differences in stated posterior distributions between Narrow and Broad.*

Three implications of these results are that (i) there is no systematic confusion about the experimental setup, since the average belief in *Broad* is nearly Bayesian, (ii), in *Narrow*, the average belief *overshoots* in the direction of the signal and (iii) is *overprecise* relative to both the Bayesian benchmark and beliefs stated in *Broad*. Overprecision is the common finding in belief research that the implied variance of stated beliefs is too low, indicating people's excessive confidence in their own judgments (Moore et al., 2015). In the present context, overprecision in *Narrow* is solely generated by the presence of a nuisance variable, as the information structure does not change relative to *Broad*. Task (3) in Figure 1 exemplifies the role of overprecision. Since the signal realization coincides with the mean of the Bayesian posterior distribution, subjects in *Narrow* form unbiased beliefs on average about X, i.e., they correctly guess the expected value of X given the signal. However, they express too much certainty that this expected value of X equals the actual draw. This finding could not be identified from point predictions about X alone.

Nuisance neglect implies a sizeable monetary cost for subjects. The average expected payoff for the beliefs stated in the baseline tasks is 53% higher in *Broad* than in *Narrow* (5.86 versus 3.82 euros, $p < 0.001$, M-W U test).[15]

3.1.5 Heterogeneity

Next, we examine what are *typical* beliefs in each condition. We characterize each stated belief by its relative proximity to the Bayesian posterior as opposed to the nuisance neglect posterior. Using that each observation is a distribution, we calculate the Hellinger metric distance (Hellinger, 1909) between the stated posterior b_X and the Bayesian posterior

[15]Actual earnings for the baseline tasks also significantly differ across groups (means of 4.56 in *Narrow* and 2.22 euros in *Broad*, $p = 0.005$, M-W U test), but these further depend on randomness induced by the binarized scoring rule as well as an additional choice by subjects that affects their payoff (see Section 4.2.3).

Table 2: Beliefs about X in baseline tasks

Signal realization	Bayesian posterior distribution	Stated posterior distribution		Sign test of median		M-W U test
		Narrow N=72	*Broad* N=72	*Narrow* vs. Bayesian	*Broad* vs. Bayesian	*Narrow* vs. *Broad*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
	distribution mean (distribution variance)	*median of distribution means (median of distribution variances)*		*p-value: distribution of means (p-value: distribution of variances)*		
60	50 (200)	60 (0)	50 (200)	< 0.001 (< 0.001)	0.664 (0.011)	< 0.001 (< 0.001)
230	237.6 (71.7)	230 (0)	240 (67)	< 0.001 (< 0.001)	< 0.001 (< 0.001)	< 0.001 (< 0.001)
200	200 (200)	200 (0)	200 (200)	1.000 (< 0.001)	0.012 (0.004)	0.024 (< 0.001)
80	95 (125)	80 (0)	95 (125)	< 0.001 (< 0.001)	0.508 (0.180)	< 0.001 (< 0.001)
165	155 (125)	165 (25)	155 (125)	< 0.001 (< 0.001)	1.000 (0.180)	< 0.001 (< 0.001)

Notes: Beliefs in *Narrow* and *Broad* by task. Each stated belief is a distribution, summarized here by its mean and variance. The table shows medians of stated distribution means and stated distribution variances for each condition (Columns (3) and (4)) and compares these to the mean and variance of the Bayesian posterior distribution (Columns (5) and (6)). Column (7) shows treatment comparisons. The task order was randomized at the subject level.

$P(X|S)$ distributions:[16]

$$H_B = \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{k}\left(\sqrt{b_{X_i}} - \sqrt{P(X_i|S=s)}\right)^2} \tag{4}$$

Given an analogous distance to the inattentive posterior distribution, H_N,[17] we define an inattention score θ that captures the distance of the subjective belief distribution to the Bayesian distribution, relative to the sum of the distances of the subjective distribution to the inattentive and the Bayesian posterior:

$$\theta = \frac{H_B}{H_B + H_N} \tag{6}$$

[16]The Hellinger distance is a bounded metric used to characterize the similarity between two probability distributions (Bandyopadhyay et al., 2016). It is suited for the present purpose as it is a proper metric, unlike, e.g., the Kullback-Leibler divergence, which does not satisfy symmetry.

[17]H_N is calculated as:

$$H_N = \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{k}\left(\sqrt{b_{X_i}} - \sqrt{P(X_i|\tilde{S}=s)}\right)^2} \tag{5}$$

A Bayesian belief corresponds to $\theta = 0$ and nuisance neglect to $\theta = 1$. The parameter θ is computed individually for each stated belief. First, Figure 2 provides a histogram of empirical inattention parameters by treatment condition. This analysis pools all stated beliefs in a treatment condition across tasks and subjects. More than 70% of beliefs in *Broad*, but less 20% in *Narrow* can be characterized as close to Bayesian ($\theta < 0.1$). By contrast, about 60% of beliefs in *Narrow* are close to nuisance neglect ($\theta > 0.9$), with the remaining 20% located in between the two extremes. The vast majority of stated beliefs are either fully sophisticated or fully inattentive to Y. This measure of inattention suggests a markedly bi-modal distribution of beliefs. Second, we analyze the within-subject heterogeneity of beliefs by counting how often each subject states a belief that is close to Bayesian ($\theta < 0.1$) or nuisance neglect ($\theta > 0.9$), as opposed to a belief that corresponds to neither of the two ($\theta \in [0.1, 0.9]$). 58% of subjects in *Broad* but only 6% in *Narrow* state *all* of their beliefs in line with Bayes' rule. 44% of subjects in *Narrow* (but none in *Broad*) exhibit nuisance neglect in all of their stated beliefs. Consequently, a share of 62% in *Broad* and 50% in *Narrow* switch at least once between the three updating modes specified here. Kernel density estimates of the subject-level average of θ display a pronounced peak around zero mean inattention in *Broad* and a less pronounced clustering of subjects with mean inattention above 0.8 in *Narrow* (see Appendix Figure 6). This suggests that while a considerable fraction of subjects is consistently inattentive, most subjects in *Narrow* exhibit some heterogeneity, with 15.5% reporting both a fully Bayesian belief and a belief implying exact nuisance neglect at least once. Strikingly, we find that a staggering 93% of beliefs stated in rounds that followed a close-to-Bayesian belief ($\theta < 0.1$) are also close to Bayesian. This fraction was only slightly lower in condition *Narrow* (82%) than in *Broad* (95%). This finding highlights the role of "insight": once people figure out the right strategy, they consistently apply it throughout subsequent problems. This provides a first indication for the relevance of the mental model of cognitive solution approach.

3.2 Robustness and External Validity

The baseline study documents nuisance neglect in a specific configuration of the information environment and experimental setup. Using further experiments we test the robustness of the findings, address potential confounds and examine the generalizability of the baseline result.

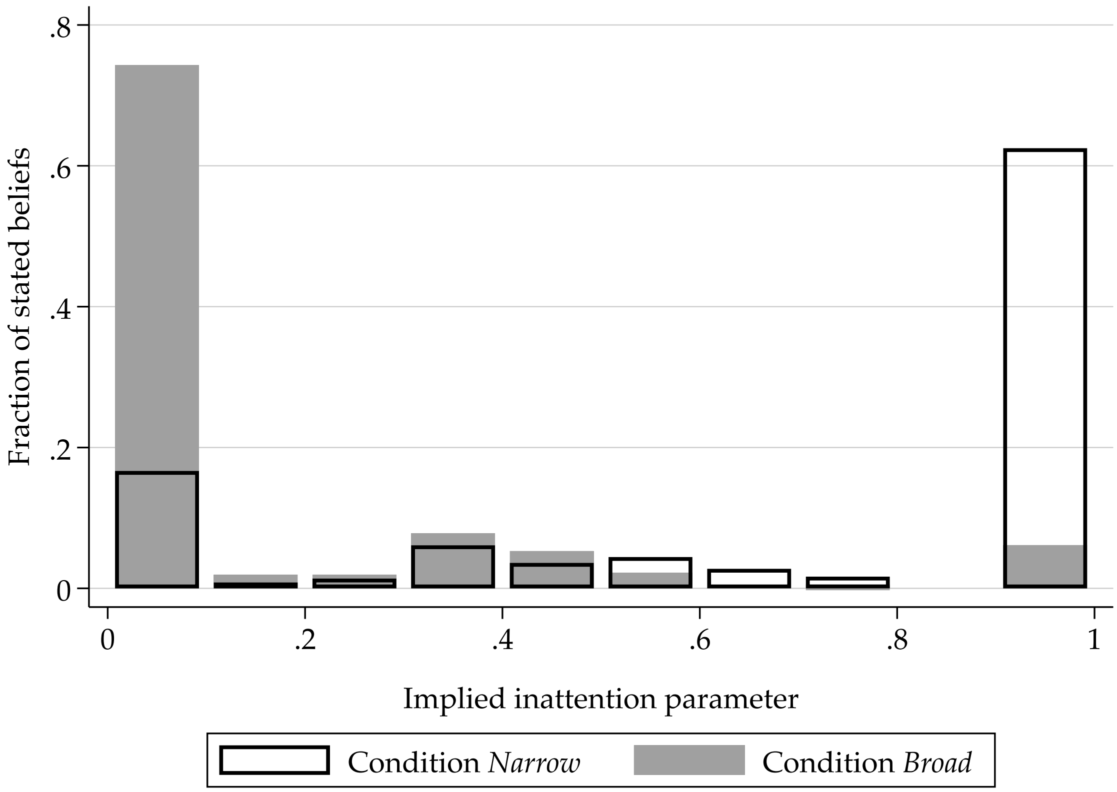


Figure 2: Distribution of implied inattention scores by treatment condition. N=1135. Displayed are binned histograms for the implied inattention parameter based on all beliefs stated in the five baseline tasks. Inattention scores are calculated as $\theta = \frac{H_B}{H_B+H_N}$, where H_B and H_N denote the Hellinger distances of the stated belief distribution to the Bayesian posterior and the nuisance neglect posterior, respectively. A parameter of $\theta = 0$ corresponds to Bayesian updating. $\theta = 1$ implies nuisance neglect.

3.2.1 Robustness and Extensions

This section summarizes a collection of robustness exercises and extensions that include (i) additional tasks introducing various departures from the simple discrete uniform case, (ii) a direct test of a signal anchoring heuristic, (iii) two treatments that exactly align the elicitation procedure across conditions, (iv) a simplified version that removes the deciphering stage and time limits, (v) a test of a face value heuristic, and (vi) an examination of the form of nuisance neglect across information structures. The following provides a brief discussion of these analyses, with all details relegated to the Appendix.

Adding to the baseline tasks in Table 1, four additional tasks were presented in random order after the baseline tasks. Highly significant treatment effects persist ($p < 0.001$, M-W U tests) under continuous uniform, normally distributed, or correlated data structures, or a case in which the signal realization is outside of the range of X. See Appendix C.1 for the robustness task specifications and detailed results.

A potential concern is that the treatment manipulation in the baseline study not only varies the incentive structure as postulated by the definition of a nuisance variable, but also the elicitation procedure: subjects in *Narrow* only state a belief about X, whereas subjects in $Broad$ guess both X and Y.[18] To better understand the extent to which the treatment effect is due to the difference in elicitation procedures, two additional treatments are designed to obtain a 2 (incentives *Narrow* vs. *Broad*) × 2 (elicitation of: only X vs. X and Y) between-subjects design. We find that given an incentive structure, i.e., *Narrow* or *Broad*, harmonizing the elicitation protocol reduces the treatment effect by roughly one third, while all differences in estimated inattention scores remain highly significant (see Appendix C.3). Put differently, most of the treatment effect is driven by prediction incentives as opposed to the elicitation procedure.

Finally, drastically simplifying condition *Narrow* by removing the deciphering stage as well as all time limits induces a reduction in the implied inattention parameter ($p < 0.001$), but the treatment effect persists in a conservative comparison against the baseline condition *Broad* which included both deciphering and time limits ($p < 0.001$; Appendix C.4).

A possible explanation of nuisance neglect is that subjects use the heuristic of reporting back the signal value, akin to exact anchoring or taking the signal at face value. Treatment *Computation* tests the face value explanation by adding a simple algebraic computation into the information structure, in such a way that it remains equally plausible to anchor on the observed signal value. For example, instead of $S = \frac{X+Y}{2}$, subjects receive the modified signal $S = \frac{X+Y}{2} - (2 \cdot 10) + 30$. We find minimal evidence for anchoring on the observed signal. Instead, subjects are able and willing to invert the computations, but still do not account for Y.[19]

Next, we complement the baseline evidence from the laboratory with online experiments in a large, more heterogeneous population. We implement several design modifications for these experiments that are discussed in Appendix D. While the design of the lab study was suitable for a sample of highly attentive student subjects, adjustments were necessary to adapt this experiment to the plausibly less attentive online worker population.

[18]This is a deliberate design choice: making a prediction in itself provides a (non-monetary) incentive to pay attention, or constitutes a form of "cue" as studied in Section 4. Note that the information set at the time of stating a belief is held exactly constant across treatments, so that subjects in *Narrow* do not have to memorize the distribution of Y.

[19]Further treatment details, figures and results are relegated to Appendix C.2.

Correspondingly, the online study did not serve the purpose of an exact replication, but instead aimed at documenting the prevalence of nuisance neglect under less controlled conditions and a more diverse sample, and served as the basis for investigating different features of the phenomenon.

The main finding of substantial nuisance neglect replicates in the online study. Specifically, 53% of stated beliefs imply an attention parameter θ above 0.9. In addition, we document evidence for an additional updating mode, "signal neglect" or non-updating, a frequent finding in belief formation studies (Coutts, 2019; Henckel et al., 2018; Möbius et al., 2014). Using additional variation in the online experiment, we make some progress towards a characterization of the form of nuisance neglect across information structures. Our results indicate that rather than the (possibly implicit) use of a distorted distribution of Y or a non-Bayesian updating rule, nuisance neglect is best characterized as a strong form of ignorance about the existence of Y: people seem to apply a modified signal structure S_i that excludes Y.

3.2.2 External Validity: Nuisance Neglect in a Naturalistic Setting

While the baseline experimental paradigm provides evidence in a tightly controlled setup, it lacks the ecological validity of real-life contexts in which people typically encounter inference problems. To address this issue and examine the generalizability of nuisance neglect in more naturalistic settings, we designed additional, pre-registered experiments. This series of experiments (i) relies on more real-world contexts that subjects may have some familiarity with, (ii) is not limited to an abstract situation that "feels like a math problem," (iii) leverages one application with more immediate economic relevance, and (iv) includes a version where subjects take a simple choice rather than state a belief given a complex scoring rule.

Design. We preserve the basic structure of the baseline experiments, with two random variables (X and Y) that causally affect a third variable (S), but simplify by binarizing all three variables. We specify the base rates of as well as the causal relationships between the two generative causes (X, Y) and the effect variable (S). Each subject took decisions in two naturalistic contexts (see complete instructions and decision screens in Appendix G.6). In context *Earnings*, a hypothetical company makes a quarterly earnings announcement, which either surpasses or falls short of an analyst prediction. In this scenario, there are two generative causes, the company's operational performance and the general business climate. When the company exceeds the operational performance goal, this *causes* realized earnings to surpass analyst expectations with probability 70%, irrespective of the business climate (see below for a discussion of how to model posterior beliefs given such a causal structure). Conversely, when the business climate is good, this *causes* realized earnings to surpass analyst expectations with probability 90%, irrespective of operational performance. Exceeding operational performance and good business climate are independent of one another and each occur with 50% probability. Subjects then find out that the company's earnings actually surpassed the analyst prediction. Given all this information, a Bayesian would infer that there is a 65% chance that the company exceeded the operational performance goal, and a 73% chance that the business climate was good. Similar to the baseline experiment, our main interest was in the treatment comparison between *Broad*, in which subjects stated their beliefs about both operational performance and business climate, and *Narrow*, in which subjects were only asked about operational performance.

The second vignette, *Restaurant*, leverages a context that plausibly taps into subjects' real-life experience. Subjects were asked to imagine having dinner at a new restaurant. The dining experience either exceeds or fails their expectation based on similar restaurants. The actual restaurant quality (which is outstanding or not with equal probability) causes the dining experience to exceed expectations with probability 95%, and good luck that is unrelated to restaurant quality – such as a good mood, sunny weather or enjoyable company (which happens with 50% probability) – causes an exceeding dining experience with probability 80%. The corresponding Bayesian posterior was 71% for outstanding restaurant quality and 65% for good luck.

Treatment conditions and outcomes. We ran a total of six between-subjects treatment conditions that cross the main treatment manipulation (*Broad* vs. *Narrow*) with different outcome measures: in *Belief Probabilistic*, subjects received all probabilistic information that was necessary to form a Bayesian posterior similar to the baseline experiment, and were incentivized using a binarized scoring rule. In *Action*, subjects took an action instead of stating a belief. For one (in *Narrow*) or both (in *Broad*) causes, they were endowed with \$1 each and could either keep this money or bet it on the occurrence of the respective cause. If the cause occurred, their \$1 bet would be tripled, and if the cause did not occur, the money would be lost. In *Belief Simple*, we replaced all numerical probability information with verbal descriptions (e.g., 95% was described as an "extremely high" probability and 80% as a "high" probability) and only asked subjects to indicate which state of the target cause (in *Narrow*) or of both causes (in *Broad*) they thought was more likely, without any monetary incentive to avoid the added complexity of a scoring rule. See all decision screens in Appendix G.6.

The rationale behind this series of treatments is that while the first outcome variant (*Belief Probabilistic*) is closest in spirit to a simplified version of the baseline experiments and mainly adds a naturalistic cover story, the second variant (*Action*) removes the artificiality of stating a belief (and the corresponding complex scoring rule), and the third variant (*Belief Simple*) simplifies even further by removing all explicit probabilistic information and the incentive scheme.

Procedures and Pre-registered Predictions. The vignettes specified the likelihood with which each of the causes changes the state of the effect; the so-called "causal power" (see, e.g., Cheng, 1997). The standard Boolean Noisy-Or parameterization for non-deterministic disjunctive interactions between causes of an effect allows us to specify the normative equations for causal inference (for details, see, e.g., Pearl, 2014). We pre-registered two types of predictions.[20] First, we predicted a treatment effect of the main manipulation, i.e., that subjects are more likely to believe that the target cause occurred (or bet on its occurrence) in *Narrow* than in *Broad*. Second, we predicted that the point belief in condition *Narrow* for subjects facing the *Belief Probabilistic* condition would be significantly higher than the Bayesian posterior (as implied by a Noisy-Or model).[21] We pre-registered a total sample of 600 completed responses across the six treatment conditions. A (pseudo-) representative online sample of the US population was collected on Prolific in September 2021.

Results. Figure 3 illustrates the results from all six experimental conditions, separately for each of the vignettes. We confirm both pre-registered predictions. First, we find a treatment effect between *Narrow* and *Broad* across all six vignette-outcome pairs.[22] Second, we predicted that beliefs in *Narrow* significantly exceed the Bayesian benchmark for *Belief Probabilistic*, which is also confirmed. Finally, we observe that *Broad* beliefs in *Belief Probabilistic* are indistinguishable from the Bayesian posterior in *Earnings*, but not in

[20] See `https://aspredicted.org/w2qi8.pdf`.

[21] Note that there are no Bayesian point predictions for the two other types of outcomes.

[22] *Belief Probabilistic*: one-sided t-tests yield $p < 0.001$ in *Earnings* and $p = 0.007$ in *Restaurant*. *Action*: two-sample tests of proportion yield $p = 0.007$ in *Earnings* and $p = 0.019$ in *Restaurant*. *Action*: two-sample tests of proportion yield $p = 0.022$ in *Earnings* and $p = 0.085$ in *Restaurant*.

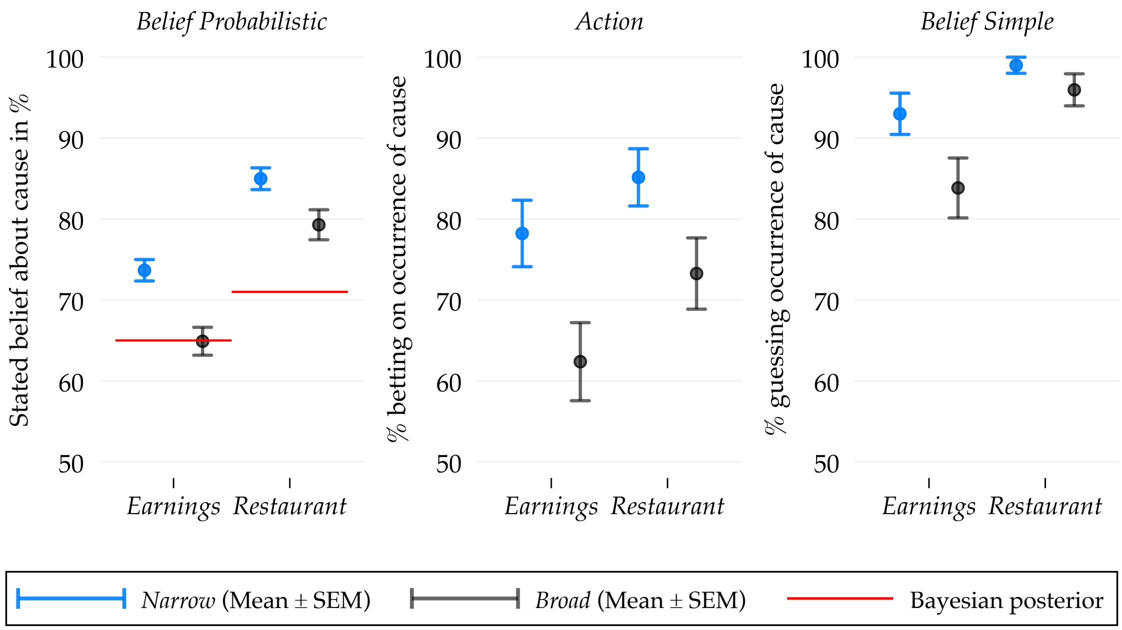


Figure 3: Results from experiments with naturalistic task framing. Participants were randomly assigned to either state a probabilistic belief (*Belief Probabilistic*, $N = 199$), take an action by placing a bet (*Action*, $N = 202$) or guess the realization of the focal variable (*Belief Simple*, $N = 199$). Within each outcome group, subjects were randomly assigned to either condition *Narrow* or *Broad*. Each participant completed both the *Earnings* and the *Restaurant* vignettes in random order. Displayed are the mean decisions and standard errors of the mean. The sample size, the treatment effect (*Narrow* vs. *Broad*) and the deviation of the point belief in *Narrow* from the Bayesian posterior in *Belief Probabilistic* were pre-registered.

Restaurant. We did not pre-register a prediction about point beliefs in *Broad*. The reason is that there is a multitude of potential explanations for non-Bayesian point beliefs that vary across the vignettes. For example, point beliefs are likely affected by the idiosyncratic features of the real-life applications. In sum, this series of treatments strongly support the external validity of the baseline experiments in more naturalistic task settings.

4 Cognitive Mechanisms: Mental Gap or Friction?

4.1 Conceptual Considerations

Research in behavioral economics has produced a collection of deviations from rationality in information processing, many of which are studied and modeled in isolation. Understanding the mechanisms behind updating errors can help identify any common primitives of different anomalies, potentially advancing the convergence of models (Fudenberg, 2006) and informing the design of interventions that target specific mechanisms with what Handel and Schwartzstein (2018) label "mechanism policies."

Previous research classifies the sources of deviations from optimality into different categories. We adopt the distinction between "mental gaps" and "frictions" (Handel and Schwartzstein, 2018) here as a productive organizing structure that is representative of other, similar classifications. First, a friction occurs if people understand a problem correctly, but do not accurately execute all necessary steps to arrive at the normatively optimal solution due to, e.g., computational errors, noisy processing or limited attentional capacity. The corresponding class of models which includes rational inattention frameworks assume that beliefs formation reflects cost-benefit considerations in the presence of some psychological cost of processing information.

Second, a mental gap occurs if people develop an incorrect understanding of the situation to begin with, so that non-Bayesian reasoning is due to *how* they approach and think about a problem given costs. A recent strand of literature in economics examines the implications of misspecified mental models.[23] This work builds on a prominent theme in cognitive science that studies people's mental representations, i.e., their subjective models of a problem (Clark, 2013; Fodor and Pylyshyn, 1988; Newell and Simon, 1972; Pitt, 2018). Misspecified subjective representations have been characterized through their automaticity, i.e., they emerge quickly and effortlessly, they tend to be simple, low-complexity models, and it requires some form of cue to trigger a different representation. This notion of default mental models is related to the intuition-based "System 1" that provides automatic, effortless responses to problems according to dual-process theories (Evans and Stanovich, 2013; Kahneman, 2003). Dual-process theories also feature the idea that System 1 override by the deliberate, effortful System 2 does not occur automatically but requires situational cues (Kahneman, 2003; Stanovich and West, 2008).

This Section aims to shed light on whether nuisance neglect is better characterized as rooted in a mental gap or a friction. If a mental gap is at the source of nuisance neglect, the bias should respond to attentional manipulations that alter how a subject thinks about the task. If nuisance neglect reflects a friction, its prevalence should depend on the relative size of the benefits and cognitive costs associated with an updating problem.

In the following we present a sequence of analyses that aim to disentangle a mental gap from a friction explanation. Friction explanations have been widely studied in the

[23]See, e.g., Barron et al. (2019); Bohren (2016); Bohren and Hauser (2019); Bushong et al. (2019); Enke (2020); Enke and Zimmermann (2019); Esponda and Pouzo (2016); Eyster and Rabin (2014); Gabaix (2014); Gennaioli and Shleifer (2010); Hanna et al. (2014); Heidhues et al. (2018); Schwartzstein (2014); Spiegler (2016).

literature, which includes psychometric designs used the in recent economics literature (e.g., Caplin et al., 2020). Our primary focus is on testing for the presence of a mental gap, as work on this topic has received comparably less attention in previous work (but there are exceptions, e.g., Enke, 2020). Note that the objective is to test *whether* a mental gap is at the source of the bias in the experimental setup under consideration. It is beyond the scope of this paper to characterize what mental models look like in general.

We proceed in two steps. First, in Section 4.2, we test for a mental gap explanation while trying to hold constant the costs and benefits associated with nuisance neglect. Second, in Section 4.3, we briefly outline our findings from tests of whether nuisance neglect responds to its associated costs while plausibly holding the mental representation that people form constant.

4.2 On Mental Gaps

In the following, we present three tests of a mental gap: an explicit hint at the nuisance variable, an implicit nudge for subjects to reconsider their problem representation, and an analysis of subjects' awareness about their nuisance neglect. Appendix Table A gives an overview of all experimental treatments.

4.2.1 The Effect of a Hint

Treatment *Hint* only provides incentives for estimating X (similar to condition *Narrow*), but adds a contextual cue that shifts attention to the nuisance variable. On every elicitation screen, subjects see a statement that reads "Also think about the role of Y." Note that the hint does not provide direct instructions on how to solve the updating problem. If subjects are aware of the relevance of Y in the updating problem, this hint should have no effect. Moreover, the hint itself neither changes the cognitive costs associated with accounting for Y in processing the signal, nor the incentive for accuracy.

We conduct online experiments on MTurk following the procedures described in Section D and using the task specifications listed in Appendix Table 10. The results of treatment *Hint* and other mechanism treatments are summarized in Figure 4. For each treatment, we pool data from all five updating tasks and display the fraction of beliefs that can be characterized as Bayesian, nuisance neglect, and signal neglect – a posterior belief that equals the prior –, as well the remaining fraction of beliefs that does not correspond to any of these updating rules.[24] We compare the change in the prevalence of different updating rules relative to the baseline treatment.

Figure 4 documents a substantial and highly significant decrease in the fraction of nuisance neglect by almost two thirds upon adding the hint (χ^2 test, $p < 0.001$). At the same time, Bayesian updating significantly increases from a fraction of below 30% to roughly 50% ($p < 0.001$). Without changing monetary incentives or cognitive costs, the hint has a substantial effect on updating, in support of the idea that subjects are in principle willing and able to update in Bayesian fashion, but fail to notice the need to account for Y to begin with.

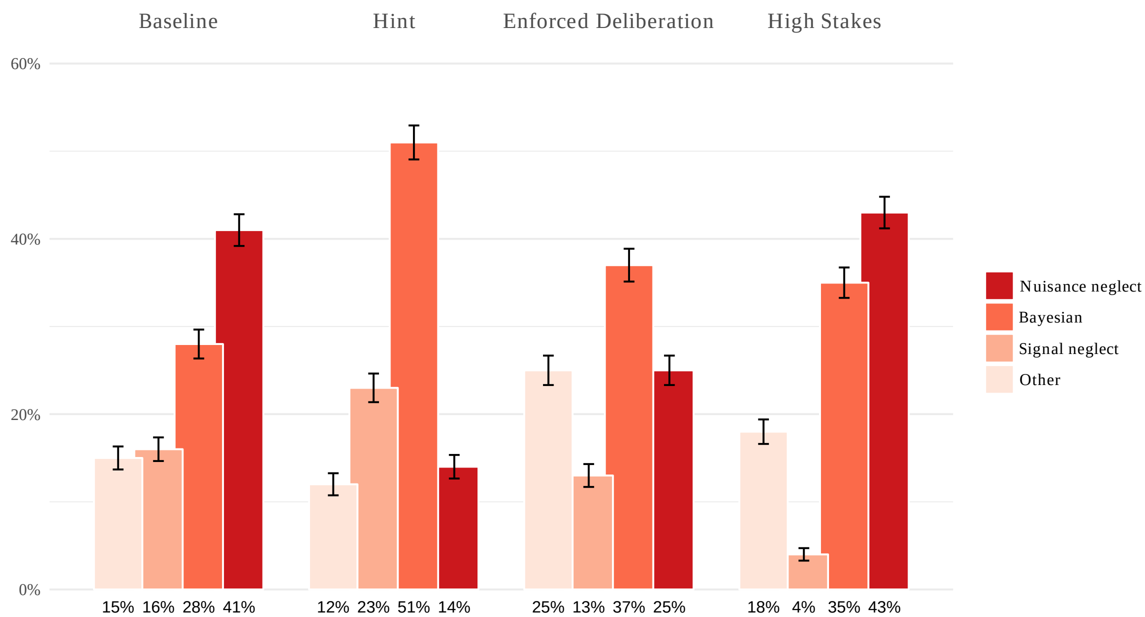


Figure 4: Fraction of stated beliefs in line with nuisance neglect, Bayesian updating and signal neglect, as well as the remaining share of beliefs, separately for the baseline online experiment (see tasks specifications in Appendix Table 10) and three mechanisms treatments discussed in Sections 4.2.1 (*Hint*), 4.2.2 (*Enforced Deliberation*) and 4.3 (*High Stakes*). All stated beliefs, pooled across updating tasks. Error bars indicate standard errors of the proportion. Stated beliefs are classified as Bayesian if they are within +/-1 percentage points of the Bayesian posterior, and as nuisance neglect or signal neglect if they exactly corresponded to stating $X = s$ or $X = \mu_X$, respectively.

[24]Specifically, stated beliefs are classified as Bayesian if they fall into a window of +/-1 percentage points of the Bayesian posterior, and as nuisance neglect or signal neglect if they exactly correspond to stating $X = s$ or $X = \mu_X$, respectively.

4.2.2 The Effect of Enforced Deliberation Time

The external hint directs attention to the neglected part of the task. Do subjects notice the nuisance variable on their own if they are nudged to re-consider their solution strategy?

In condition *Enforced Deliberation*, subjects face a 30-second waiting time on each elicitation screen, during which they cannot enter a guess or submit the page. The input fields are only activated after that time is up. This variant of enforced waiting time aims at having people deliberate their approach towards solving the problem – rather than the execution of the subsequent computations –, potentially leading them to recognize the need to account for Y. Figure 4 shows that this is the case: The share of nuisance neglect in *Enforced Deliberation* falls substantially from 41% to 25% ($p < 0.001$), roughly by half as much as the effect size of a hint.

This result is consistent with an interpretation according to which subjects' solution strategy may be divided into two successive steps: first, parsing the problem description into a mental problem representation, and second, implementing a solution based on that representation. Noticing the neglect may require that subjects specifically reconsider their problem interpretation, rather than their downstream implementation.

4.2.3 Awareness of Nuisance Neglect

The effects of a hint and enforced deliberation demonstrate that even minimal interventions that bring attention to Y suffice to substantially reduce nuisance neglect. This indicates that subjects initially fail to think about about Y and are thus *unaware* about committing an error. We provide a correlational analysis of this lack-of-awareness hypothesis by measuring confidence in stated beliefs. If subjects who commit nuisance neglect are aware of their distorted beliefs, they will be less confident than subjects who form optimal beliefs. If instead the simplification of ignoring Y occurs outside of the agent's control, subjects may deliberately execute the subsequent computations and still exhibit high confidence in their beliefs. In stage *Confidence* that directly follows the belief tasks in the baseline laboratory experiment (Section 3), subjects indicate their willingness-to-accept (WTA) to give up the uncertain payoff associated with each previously stated belief. They are again presented with each individual updating task together with their own stated belief. Participants are asked to indicate whether they prefer to be paid out for the accuracy of their belief or receive a certain monetary amount. They make this binary decision for different fixed amounts ranging from 0 euros to 6 euros in increments of 0.25 euros. These choices are presented using the multiple-price list method, which Andreoni and

Kuhn (2019) argue is particularly easy to understand for subjects. If the task is randomly selected for payout in the end, a subject's decision in one of the rows of the list is implemented. Note that the *Confidence* tasks (i) have no time limit such that subjects could freely rethink their stated belief, and (ii) the subjective valuation in each task provides an incentivized measure of confidence in the belief distribution itself, beyond the variance of the stated belief distribution.[25]

Table 3 shows results from regressions in which the dependent variable is the subjective valuation of a stated belief, i.e., the minimal certain amount preferred over a having the stated belief paid out. A higher value corresponds to higher confidence in a stated belief. Columns (1) to (3) show that more inattentive beliefs are not significantly associated with lower reservation prices. Even after reconsidering the updating problem and their own belief, subjects fail to recognize the necessity to account for Y and are equally confident in their own guess. Reassuringly, the variance of the indicated belief distribution is negatively correlated with confidence. While these analyses are correlational in nature, we exploit the causal variation in θ generated by the treatment manipulation (between conditions *Narrow* and *Broad*) in a regression reported in column (4) of Table 3, which uses treatment status as an instrument for inattention θ. The two-stage least squares procedure yields a similar coefficient estimate, again indicating no significant relationship between inattention and confidence. Restricting the sample to beliefs stated in *Narrow* (column (5)), we again find no relationship between the valuation of a stated belief and implied inattention.

Taking stock, the effectiveness of simple attentional manipulations and subjects' unawareness of nuisance neglect implied by the confidence measure points to a behavioral mechanism related to how subjects mentally construe the updating problem. Moreover, the findings suggest that selective processing of problem features at least partly depends on factors that are unrelated to the cost of information processing.

Result 2. *Nuisance neglect is reduced by simple attentional cues to Y.*

[25]The WTA, however, depends on the curvature of the utility function, which motivates robustness analyses below that take into account subjects' risk attitudes.

Table 3: Determinants of confidence in stated beliefs

Dependent variable:	**Confidence**: Valuation for stated belief				
Condition:	*Narrow* and *Broad*				*Narrow*
Estimation method:	OLS			IV	OLS
	(1)	(2)	(3)	(4)	(5)
0 if *Broad*, 1 if *Narrow*	-0.497	-0.499	-0.104		
	(0.316)	(0.317)	(0.300)		
Inattention θ	-0.808	-0.801	-0.369	-0.566	-0.487
	(0.509)	(0.508)	(0.512)	(0.429)	(0.436)
Treatment dummy * Inattention θ	0.714	0.705	0.006		
	(0.620)	(0.619)	(0.612)		
Variance of belief distribution		-0.000***	-0.000*	-0.000*	-0.000
		(0.000)	(0.000)	(0.000)	(0.002)
Willingness to take risks			0.555***	0.557***	0.645***
			(0.134)	(0.134)	(0.175)
Constant	4.550***	4.555***	3.694***	3.689***	3.011***
	(0.180)	(0.181)	(0.637)	(0.619)	(0.613)
Task fixed effects			Yes	Yes	Yes
Additional controls			Yes	Yes	Yes
R^2	0.02	0.02	0.13	0.13	0.15
# Observations	1135	1135	1135	1135	607

Notes: Least squares and IV regressions. Inattention scores are calculated as $\theta = \frac{H_B}{H_B+H_N}$, where H_B and H_N denote the Hellinger distances of the stated distribution to the Bayesian posterior and the nuisance neglect posterior (as defined in Section 3), respectively. In column (4), implied inattention scores θ are instrumented with an indicator for treatment status (1 if *Narrow*, 0 if *Broad*). The additional controls include gender, age, income and task-fixed effect. Robust standard errors clustered at participant level in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.3 On Cost-Benefit Considerations

Why does the attribution error in inference arise even though it creates a substantial monetary loss? One explanation is that it reflects a simplification strategy that economizes on cognitive resources. A more parsimonious problem representation arguably draws less cognitive capacity, and optimization within a simpler model may allow a quicker and less effortful solution. The avoidance of cognitive effort has a been a long-standing theme in cognitive science that has led to the notion of humans as "cognitive misers" or "motivated tacticians" (Fiske and Taylor, 2013; Stanovich, 2009), with some arguing that most biases in judgment and decision-making reflect effort-reduction strategies (Shah and Oppenheimer, 2008). In economics, a growing literature shows that simplifications and inattention can reflect rational, constrained optimization in the presence of cognitive costs or capacity limitations (Caplin and Dean, 2015; Gabaix, 2014; Sims, 2003; Wiederholt, 2010). A common prediction of this class of models is that deviations from rationality respond to their associated cost. If nuisance neglect is driven by underlying cost-benefit considerations – explicit or implicit, i.e., without the agent's awareness –, then its prevalence should respond to the cognitive costs and the expected benefits of optimal belief updating.

Next, we outline the main findings from examining the effect of variation in the costs of nuisance neglect on its occurrence within the paradigm of this experiment. Put differently, we focus on the sensitivity of updating patterns to changing costs.[26] All details are relegated to Appendix E.

In treatment *High Stakes*, the available prize is raised five-fold relative to the baseline online experiment. Under higher incentives, effort as measured by response times increases significantly, both overall and within each subgroup (pairwise t tests, all $p < 0.001$). We find that the prevalence of Bayesian updating increases statistically significantly, but the share of nuisance neglect remains roughly constant. In fact, the increase in Bayesian updating occurs fully at the expense of signal neglect. This means, given higher incentives, subjects try harder, but that only affects non-updating, reducing the fraction of subjects that ignore the signal altogether. On average, higher stakes do not reduce nuisance neglect, however. Compellingly, a tenfold increase of the stake size in the laboratory experiment leads to a similar pattern, see Appendix Section E.4. This indicates that psychic costs, cognitive miserliness, laziness or effort reduction may explain non-updating, but have limited explanatory power for nuisance neglect.

In Appendix E.5, we investigate whether the *specific* monetary cost associated with nuisance neglect affects its prevalence. In economic models of rational belief formation, the likelihood of committing a specific error depends on its expected cost in utility terms (Caplin and Dean, 2015; Gabaix, 2014; Wiederholt, 2010). On that account, the preva-

[26]While the responsiveness of errors to their cost is a central, testable prediction of a friction explanation, a lack of responsiveness does not imply that a mental gap may not in itself be rooted in some form of psychological cost.

lence of nuisance neglect should vary systematically with its expected loss of accuracy in a given information environment. We vary the expected cost of nuisance neglect using variations of the signal-to-noise ratio and the directional bias implied by the bias. The results suggest that the prevalence nuisance neglect does not systematically respond to these variation in its expected costliness.

5 Conclusion

A collection of previous empirical findings implies that misattribution is a pervasive feature of human decision making. The extant literature by and large treats these patterns as unconnected phenomena. This paper contributes in two ways. First, by cleanly documenting the neglect of nuisance variables in both tightly controlled and naturalistic environments, it provides a potential conceptual link between various attribution errors. Because people tend to narrowly focus on explanations that appear subjectively most relevant, they disproportionately assign casual power to these explanations and neglect alternative causes. Second, by studying the precise behavioral mechanisms underlying misattribution, this paper extends the recent literature on updating problems given many pieces of information to attribution problems where agents face a single piece of information. Our conclusion that a similar mechanism of misspecified mental representations may be at play in both problem classes sheds light on the primitives of a theoretical framework that may successfully capture different types of anomalies.

Limitations and directions for future work. While the paper provides controlled evidence on the simplest type of attribution problem, it has a number of limitations. First, the combined evidence from over twenty treatments in the abstract and naturalistic paradigms is confined to *static* updating tasks with two random variables. A natural question is how the neglect of alternative causes extends to sequential updating tasks, more complicated signal structures and environments with more than two variables. Second, the presented evidence from the naturalistic paradigm remains suggestive. This type of extension that transfers structured updating problems to real-world applications merits more work in the future and can help shed light on the relevance and generalizability of belief updating patterns such as nuisance neglect. Third, while this paper concludes that updating errors here may plausibly be due to a mental gap, it does not claim to characterize in any generality when a mental gaps occurs, why mental gaps arise and what they look like in different contexts.

A broader challenge for research on bounded rationality is whether two seemingly conflicting directions in the literature can be reconciled. Evidence on incomplete mental models tends to favor overreaction and “jumping to conclusions,” which is broadly in line with the heuristics and biases program and classical work by Kahneman and Tversky. This paper falls into this category. A separate strand of the literature highlights the role of noise and imprecision in human cognition (see Woodford, 2019, for an overview). Noisy processing motivates a class of models that predominantly predict insensitivities and underreaction, which is supported by mounting evidence on both lower-level perceptual processes and higher-level reasoning (Enke and Graeber, 2022a,b; Frydman and Jin, 2019; Gabaix and Laibson, 2017; Khaw et al., 2019; Steiner and Stewart, 2016). Future work may help shed light on whether these forces operate simultaneously or apply in distinct environments, and if so what characterizes their respective scope of application.

References

Abaluck, Jason and Abi Adams, "What do consumers consider before they choose? Identification from asymmetric demand responses," Technical Report, National Bureau of Economic Research 2017.

_ **and Jonathan Gruber**, "Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program," *American Economic Review*, 2011, *101* (4), 1180–1210.

Adam, Klaus, Albert Marcet, and Johannes Beutel, "Stock price booms and expected capital gains," *American Economic Review*, 2017, *107* (8), 2352–2408.

Allcott, Hunt, "Consumers' perceptions and misperceptions of energy costs," *American Economic Review*, 2011, *101* (3), 98–104.

Andreoni, James and Michael A Kuhn, "Is It Safe to Measure Risk Preferences? A Comparison of Four Techniques," *Working Paper*, 2019.

Bandyopadhyay, Prasanta, Gordon Brittan, and Mark Taper, *Belief, Evidence, and Uncertainty: Problems of Epistemic Inference*, Springer International, 2016.

Bar-Hillel, Maya, "The base-rate fallacy in probability judgments," *Acta Psychologica*, 1980, *44* (3), 211–233.

Barron, Kai, Steffen Huck, and Philippe Jehiel, "Everyday econometricians: Selection neglect and overoptimism when learning from others," *WZB Discussion Paper SP II 2019*, 2019.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka, "Attention discrimination: Theory and field experiments with monitoring information acquisition," *American Economic Review*, 2016, *106* (6), 1437–75.

Benjamin, Daniel J, "Errors in Probabilistic Reasoning and Judgmental Biases," in "Handbook of Behavioral Economics" 2019.

Bertrand, Marianne and Sendhil Mullainathan, "Are CEOs rewarded for luck? The ones without principals are," *The Quarterly Journal of Economics*, 2001, *116* (3), 901–932.

Bohren, Aislinn, "Informational herding with model misspecification," *Journal of Economic Theory*, 2016, *163*, 222–247.

_ **and Daniel Hauser**, "Social Learning with Model Misspecification: A Framework and a Robustness Result," *PIER Working Paper 18-017*, 2019.

Bondt, Werner FM De and Richard Thaler, "Does the stock market overreact?," *The Journal of finance*, 1985, *40* (3), 793–805.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Diagnostic expectations and credit cycles," *The Journal of Finance*, 2018, *73* (1), 199–227.

— , — , **and** — , "Memory, attention, and choice," *The Quarterly Journal of Economics*, 2020, *135* (3), 1399–1442.

— , — , **Rafael La Porta, and Andrei Shleifer**, "Diagnostic expectations and stock returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

— , — , **Yueran Ma, and Andrei Shleifer**, "Overreaction in Macroeconomic Expectations," *American Economic Review*, September 2020, *110* (9), 2748–82.

Brownback, Andy and Michael A Kuhn, "Understanding outcome bias," *Games and Economic Behavior*, 2019, *117*, 342–360.

Bushong, Benjamin and Tristan Gagnon-Bartsch, "Learning with Misattribution of Reference Dependence," *working paper*, 2018.

— , **Matthew Rabin, and Joshua Schwartzstein**, "A model of relative thinking," *Working paper*, 2019.

Caplin, Andrew and Daniel Martin, "A testable theory of imperfect perception," *The Economic Journal*, 2015, *125* (582), 184–202.

— **and Mark Dean**, "Revealed preference, rational inattention, and costly information acquisition," *The American Economic Review*, 2015, *105* (7), 2183–2203.

— , **Dániel Csaba, John Leahy, and Oded Nov**, "Rational inattention, competitive supply, and psychometrics," *The Quarterly Journal of Economics*, 2020, *135* (3), 1681–1724.

— , **Mark Dean, and Daniel Martin**, "Search and satisficing," *The American Economic Review*, 2011, *101* (7), 2899–2922.

Chauvin, Kyle, "A Misattribution Theory of Discrimination," Technical Report, Mimeo 2020.

Chen, Daniel L., Martin Schonger, and Chris Wickens, "oTree–An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Cheng, Patricia W, "From covariation to causation: A causal power theory.," *Psychological review*, 1997, *104* (2), 367.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and taxation: Theory and evidence," *The American Economic Review*, 2009, *99* (4), 1145.

Clark, Andy, "Whatever next? Predictive brains, situated agents, and the future of cognitive science," *Behavioral and brain sciences*, 2013, *36* (3), 181–204.

Coibion, Olivier and Yuriy Gorodnichenko, "What can survey forecasts tell us about information rigidities?," *Journal of Political Economy*, 2012, *120* (1), 116–159.

— **and** — , "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 2015, *105* (8), 2644–78.

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

Dean, Mark and Nathaniel Neligh, "Experimental tests of rational inattention," *Columbia University Working Paper*, 2019.

DellaVigna, Stefano and Devin Pope, "Predicting experimental results: who knows what?," *Journal of Political Economy*, 2018, *126* (6), 2410–2456.

Edwards, Ward, "Conservatism in human information processing," *Formal representation of human judgment*, 1968.

Enke, Benjamin, "What You See is All There is*," *The Quarterly Journal of Economics*, 05 2020. qjaa012.

_ and Florian Zimmermann, "Correlation Neglect in Belief Formation," *The Review of Economic Studies*, 2019, *86* (1), 313–332.

_ and Thomas Graeber, "Cognitive Noise and Intertemporal Choice," *Mimeo*, 2022.

_ and _ , "Cognitive Uncertainty," *Mimeo*, 2022.

Esponda, Ignacio and Demian Pouzo, "Berk–Nash equilibrium: A framework for modeling agents with misspecified models," *Econometrica*, 2016, *84* (3), 1093–1130.

_ and Emanuel Vespa, "Hypothetical thinking and information extraction in the laboratory," *American Economic Journal: Microeconomics*, 2014, *6* (4), 180–202.

_ and _ , "Contingent thinnking and the sure-thing principle: Revisiting classic anomalies in the laboratory," *Mimeo*, 2019.

Evans, Jonathan and Keith E. Stanovich, "Dual-process theories of higher cognition: Advancing the debate," *Perspectives on Psychological Science*, 2013, *8* (3), 223–241.

Eyster, Erik and Matthew Rabin, "Naive herding in rich-information settings," *American economic journal: microeconomics*, 2010, *2* (4), 221–43.

_ and _ , "Extensive imitation is irrational and harmful," *The Quarterly Journal of Economics*, 2014, *129* (4), 1861–1898.

_ , _ , and Georg Weizsäcker, "An Experiment On Social Mislearning," Technical Report, Discussion Paper 2018.

Falk, Armin, Anke Becker, Thomas J. Dohmen, David Huffman, and Uwe Sunde, "The preference survey module: A validated instrument for measuring risk, time, and social preferences," *IZA Discussionn Paper No. 9674*, 2016.

Fernbach, Philip M, Adam Darlow, and Steven A Sloman, "Neglect of alternative causes in predictive but not diagnostic reasoning," *Psychological Science*, 2010, *21* (3), 329–336.

_ and Bob Rehder, "Cognitive shortcuts in causal inference," *Argument & Computation*, 2013, *4* (1), 64–88.

Fiske, Susan T and Shelley E Taylor, *Social cognition: From brains to culture*, Sage, 2013.

Fodor, Jerry A and Zenon W Pylyshyn, "Connectionism and cognitive architecture: A critical analysis," *Cognition*, 1988, *28* (1-2), 3–71.

Frydman, Cary and Gideon Nave, "Extrapolative beliefs in perceptual and economic decisions: Evidence of a common mechanism," *Management Science*, 2016, *63* (7), 2340–2352.

_ **and Lawrence Jin**, "Efficient Coding and Risky Choice," *Working Paper*, 2019.

Fudenberg, Drew, "Advancing beyond advances in behavioral economics," *Journal of Economic Literature*, 2006, *44* (3), 694–711.

Gabaix, Xavier, "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ , "Behavioral inattention," in Douglas Bernheimn, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics-Foundations and Applications 2*, Elsevier, 2019, p. 261.

_ **and David Laibson**, "Myopia and discounting," Technical Report, National bureau of economic research 2017.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein, "Channeled attention and stable errors," 2019. Mimeo.

Gennaioli, Nicola and Andrei Shleifer, "What comes to mind," *The Quarterly journal of economics*, 2010, *125* (4), 1399–1433.

Gigerenzer, Gerd, "How to make cognitive illusions disappear: Beyond 'heuristics and biases'," *European Review of Social Psychology*, 1991, *2* (1), 83–115.

Greenwood, Robin and Andrei Shleifer, "Expectations of returns and expected returns," *The Review of Financial Studies*, 2014, *27* (3), 714–746.

Grether, David M, "Bayes rule as a descriptive model: The representativeness heuristic," *The Quarterly journal of economics*, 1980, *95* (3), 537–557.

Gurdal, Mehmet Y, Joshua B Miller, and Aldo Rustichini, "Why blame?," *Journal of Political Economy*, 2013, *121* (6), 1205–1247.

Haggag, Kareem, Devin G Pope, Kinsey B Bryant-Lees, and Maarten W Bos, "Attribution bias in consumer choice," *The Review of Economic Studies*, 2018, *86* (5), 2136–2183.

Handel, Benjamin and Joshua Schwartzstein, "Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care?," *Journal of Economic Perspectives*, 2018, *32* (1), 155–178.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein, "Learning through noticing: Theory and evidence from a field experiment," *The Quarterly Journal of Economics*, 2014, *129* (3), 1311–1353.

Hauser, David J. and Norbert Schwarz, "Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants," *Behavior Research Methods*, 2016, *48* (1), 400–407.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack, "Unrealistic expectations and misguided learning," *Econometrica*, 2018, *86* (4), 1159–1214.

Hellinger, Ernst, "Neue Begründung der Theorie quadratischer Formen von unendlich vielen Veränderlichen.," *Journal für die reine und angewandte Mathematik*, 1909, *136*, 210–271.

Henckel, Timo, Gordon Douglas Menzies, Peter G Moffatt, and Daniel John Zizzo, "Belief adjustment: A double hurdle model and experimental evidence," *CAMA Working Paper*, 2018.

Hestermann, Nina and Yves Le Yaouanq, "Experimentation with Self-Serving Attribution Biases," *American Economic Journal: Microeconomics*, 2020.

Hossain, Tanjim and Ryo Okui, "The binarized scoring rule," *The Review of Economic Studies*, 2013, *80* (3), 984–1001.

Jehiel, Philippe, "Analogy-based expectation equilibrium," *Journal of Economic theory*, 2005, *123* (2), 81–104.

Kahneman, Daniel, "Maps of bounded rationality: Psychology for behavioral economics," *The American Economic Review*, 2003, *93* (5), 1449–1475.

_ **and Amos Tversky**, "On the study of statistical intuitions," *Cognition*, 1982, *11* (2), 123–141.

Khaw, Mel Win, Ziang Li, and Michael Woodford, "Cognitive imprecision and small-stakes risk aversion," Technical Report 2019.

Landier, Augustin, Yueran Ma, and David Thesmar, "New experimental evidence on expectations formation," *CEPR Discussion Paper No. DP12527*, 2017.

Levy, Gilat and Ronny Razin, "Correlation neglect, voting behavior, and information aggregation," *American Economic Review*, 2015, *105* (4), 1634–45.

Malmendier, Ulrike and Young Han Lee, "The bidder's curse," *American Economic Review*, 2011, *101* (2), 749–87.

Mankiw, N Gregory and Ricardo Reis, "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve," *The Quarterly Journal of Economics*, 2002, *117* (4), 1295–1328.

Martínez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa, "Failures in Contingent Reasoning: The Role of Uncertainty," *American Economic Review*, 2019, *109*

(10), 3437–74.

Matejka, Filip and Alisdair McKay, "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *The American Economic Review*, 2014, *105* (1), 272–298.

Matějka, Filip and Alisdair McKay, "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 2015, *105* (1), 272–98.

Möbius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat, "Managing self-confidence," *NBER Working paper*, 2014.

Moore, Don A, Elizabeth R Tenney, and Uriel Haran, "Overprecision in judgment," *The Wiley Blackwell handbook of judgment and decision making*, 2015, pp. 182–209.

Morewedge, Carey K and Daniel Kahneman, "Associative processes in intuitive judgment," *Trends in cognitive sciences*, 2010, *14* (10), 435–440.

Newell, Allen and Herbert A. Simon, *Human Problem Solving*, Prentice-Hall, 1972.

Pearl, Judea, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Elsevier, 2014.

Pitt, David, "Mental Representation," in Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*, winter 2018 ed., Metaphysics Research Lab, Stanford University, 2018.

Rabin, Matthew, "Incorporating limited rationality into economics," *Journal of Economic Literature*, 2013, *51* (2), 528–43.

_ **and Joel L. Schrag**, "First impressions matter: A model of confirmatory bias," *The Quarterly Journal of Economics*, 1999, *114* (1), 37–82.

Robinson, Jonathan, Cheskie Rosenzweig, Aaron J Moss, and Leib Litman, "Tapped Out or Barely Tapped? Recommendations for How to Harness the Vast and Largely Unused Potential of the Mechanical Turk Participant Pool," *PsyArXiv*, 2019.

Ross, Lee, "The intuitive psychologist and his shortcomings: Distortions in the attribution process," in "Advances in experimental social psychology," Vol. 10, Elsevier, 1977, pp. 173–220.

Schwartzstein, Joshua, "Selective attention and learning," *Journal of the European Economic Association*, 2014, *12* (6), 1423–1452.

Shah, Anuj K and Daniel M Oppenheimer, "Heuristics made easy: an effort-reduction framework.," *Psychological bulletin*, 2008, *134* (2), 207.

Shiller, Robert, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 1981, *71* (3), 421–36.

Sims, Christopher, "Implications of rational inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

_ , “Rational inattention: Beyond the linear-quadratic case,” *The American economic review*, 2006, *96* (2), 158–163.

Sloman, Steven A and David Lagnado, “Causality in thought,” *Annual review of psychology*, 2015, *66*, 223–247.

Spiegler, Ran, “Bayesian networks and boundedly rational expectations,” *The Quarterly Journal of Economics*, 2016, *131* (3), 1243–1290.

_ , ““Data Monkeys”: A Procedural Model of Extrapolation from Partial Statistics,” *The Review of Economic Studies*, 2017, *84* (4), 1818–1841.

Stanovich, Keith E., *What intelligence tests miss: The psychology of rational thought*, Yale University Press, 2009.

_ **and Richard F. West**, “Individual differences in reasoning: Implications for the rationality debate?,” *Behavioral and Brain Sciences*, 2000, *23* (5), 645–665.

_ **and** _ , “On the relative independence of thinking biases and cognitive ability.,” *Journal of personality and social psychology*, 2008, *94* (4), 672.

Steiner, Jakub and Colin Stewart, “Perceiving prospects properly,” *American Economic Review*, 2016, *106* (7), 1601–31.

Taubinsky, Dmitry and Alex Rees-Jones, “Attention variation and welfare: theory and evidence from a tax salience experiment,” *The Review of Economic Studies*, 2018, *85* (4), 2462–2496.

Tversky, Amos and Daniel Kahneman, “Belief in the law of small numbers.,” *Psychological bulletin*, 1971, *76* (2), 105.

_ **and** _ , “Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment.,” *Psychological review*, 1983, *90* (4), 293.

Wiederholt, Mirko, “Rational inattention,” *The New Palgrave Dictionary of Economics (Online Edition ed.)*, 2010.

Wolfers, Justin, “Are voters rational?: Evidence from gubernatorial elections,” *Stanford Graduate School of Business Research Paper Series*, 2002.

Woodford, Michael, “Modeling Imprecision in Perception, Valuation and Choice,” *Working Paper*, 2019.

ONLINE APPENDIX

This appendix contains an overview of all treatment conditions (Section A), additional information and results on the baseline laboratory experiments (Section B), the robustness treatments (Section C) and the online experiments (Section D), additional details and results on the mechanisms treatments (Sections E and F) as well as all experimental instructions (Section G).

A Overview of Treatments

Table 4: Overview of laboratory treatments

Condition	Description	Covered in
Baseline experiment: Narrow and Broad	(Elements of baseline experiment in respective order below)	
Baseline Tasks	5 updating tasks in random order. X and Y follow independent discrete uniform distributions with outcome spaces smaller than 10. The information is the mean or the sum of the draws.	Appendix B
Robustness Tasks	5 updating tasks in random order. Data are correlated, drawn from a larger sample space, discretely normally distributed, or the information is outside of the range of X.	Appendix C.1
Bonus Task	1 surprise task with similar configuration to baseline. Within each condition, subjects are re-randomized and face either the same expected incentive size as before, or tenfold incentives.	Main text
Confidence	For each baseline and robustness problem, subjects indicate their valuation for their stated belief using a multiple-price list method.	Appendix E.2
Switch-role	2 tasks with similar configuration as baseline, but subjects face incentives from opposite treatment condition. That means *Narrow* is paid for X and Y, while *Broad* paid for X only.	Appendix E.3
Computation	Identical to *Narrow* baseline, except that a simple, task-varying algebraic calculation is added to the information structure (e.g., "the mean $+20-30$").	Appendix C.2
Simplification	Identical to *Narrow* baseline, but deciphering stage and all time limits removed.	Appendix C.4

Condition	Description	Covered in
Narrow with joint elicitation	Identical to *Narrow* baseline, but subjects indicate the joint distribution of X and Y (while only X is paid for).	Appendix C.3
Broad with sequential elicitation	Identical to *Broad* baseline, but subjects indicate the marginal distributions of X and Y in sequential order, such that the first screen is identical to *Narrow* baseline.	Appendix C.3
Hint	Identical to *Narrow* baseline, but subjects receive a reminder on the elicitation screen, stating "Also think about the role of Y".	Appendix E.1
Feedback	Identical to *Narrow* baseline, but subjects observe the actual draw of X after stating their guess.	Appendix F.4
Computation with feedback	Identical to *Computation*, but subjects observe the actual draw of X after stating their guess.	Appendix F.5
Computational feedback	Identical to *Computation with feedback*, except that the computation is added to the feedback instead of the information.	Appendix F.6
Imperfect feedback	Identical to *Feedback*, but subjects receive the true draw as feedback only with 80% probability, while seeing another value with 20% probability.	Appendix F.7

Table 5: Overview of online treatments

Condition	Description	Covered in
Baseline experiment (Narrow only)	5 updating tasks in random order (Table 8). X and Y follow independent distributions. Subjects only state a mean posterior belief about X. No deciphering stage.	Main text
Baseline for mechanisms experiments	5 updating tasks in random order (Table 10) followed by one confidence task in which subjects indicate their WTA for a stated belief. Otherwise identical to online baseline experiment.	Main text
High Stakes	As mechanism baseline with fivefold incentives.	Main text
Hint	As mechanism baseline, but additional hint: "Also think about the role of Y".	Main text
Hint and High Stakes	Combination of treatments Hint and High Stakes	Appendix
Deliberation time	As mechanism baseline, 30-second waiting time enforced before input forms activate and page can be submitted.	Main text
Deliberation and High Stakes	Combination of treatments Deliberation Time and High Stakes	Appendix
Form of Nuisance Neglect	10 updating tasks in random order (Table 9). Identical to baseline online experiment but different information structures to analyze different candidates for the belief formation rule under nuisance neglect.	Main text and Appendix D.1
Signal-to-Noise Ratio	7 updating tasks in random order. Identical to baseline online experiment but different information structures (Table 12). The signal-to-noise ratio is varied between tasks.	Main text
Directional Bias	5 updating tasks in random order. Identical to baseline online experiment but different information structures (Table 13). All elements of the information structure are kept fixed across tasks except the mean of Y.	Main text

B Laboratory Experiments

B.1 Sequence of Events in Updating Tasks

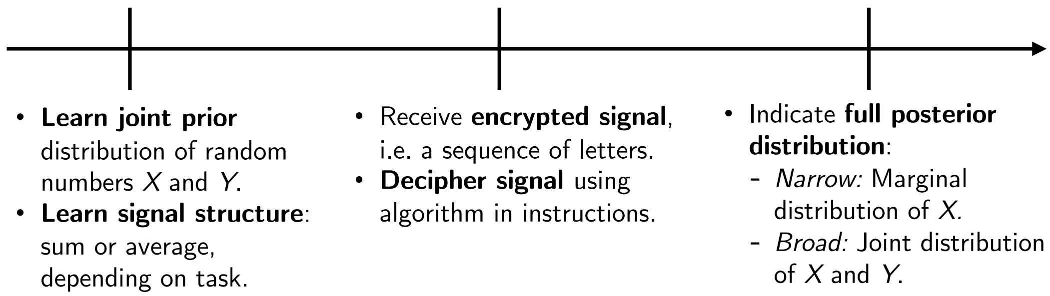


Figure 5: Timeline of updating task in laboratory experiment.

B.2 Consistency of Attention Across Tasks

In this section I examine how consistently inattentive or consistently Bayesian subjects behave across tasks. Figure 2 pools stated beliefs from all subjects. But do individuals exhibit a stable level of attention? Figure B.2 shows kernel density estimates of subject-level mean inattention. While there is a strong clustering of subjects in the *Broad* condition who always form beliefs that reflect average inattention of zero, there are no pronounced clusterings in the *Narrow* treatment – one at each end of the attention spectrum – as could be expected from Figure 2. Instead, there is a smaller peak at mean inattention values of between 0.8 and 1. Indeed, we find that many subjects in the *Narrow* condition form close to Bayesian beliefs in some tasks, and close to fully inattentive beliefs in other tasks. 15.5% of subjects in *Narrow* indicate both a fully Bayesian and a fully inattentive belief at least once. This may suggest that a subject's degree of attention to Y varies across situations to some extent, even for largely identical updating contexts.

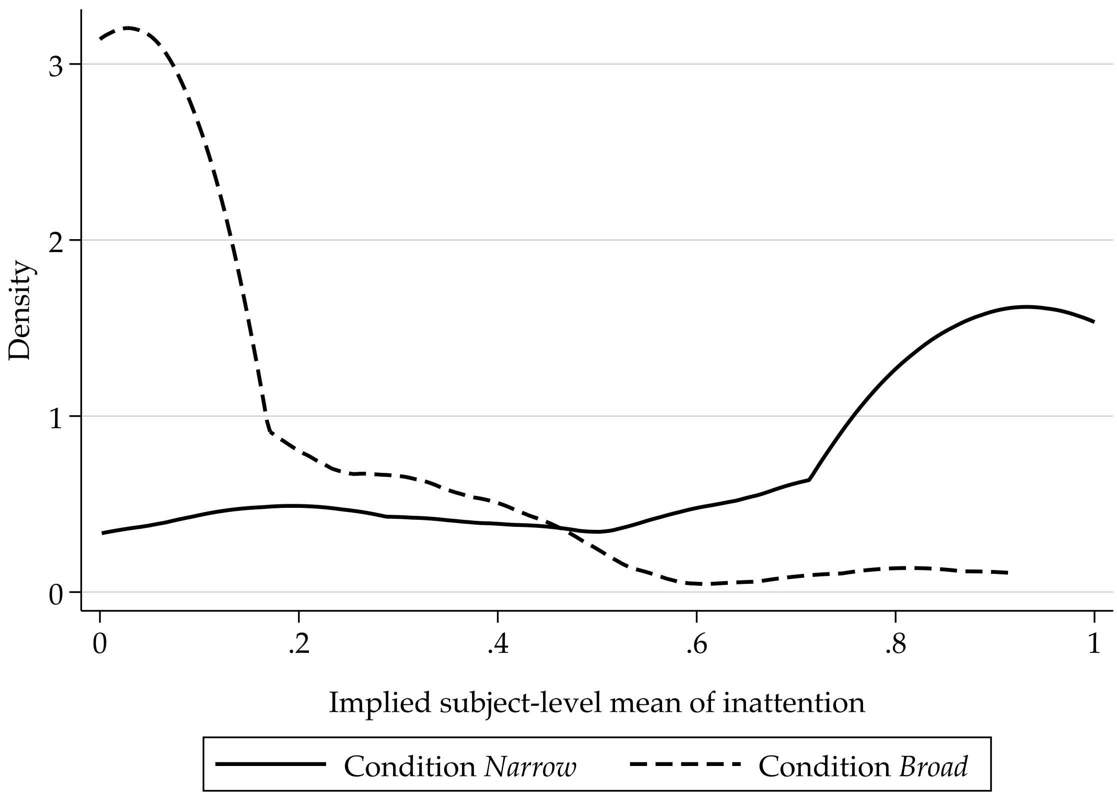


Figure 6: Subject-level mean of inattention to Y. N=144. For each subject I calculate the mean inattention in the five baseline tasks. The curves show kernel density estimates for each treatment (both N=72). A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention. Epanechnikov kernel with bandwidth 0.1.

C Robustness Treatments

C.1 Task Variations

First, moving toward a continuous data structure arguably increases the complexity of the inference problem and pushes the median subject in *Broad* away from the Bayesian benchmark. Second, normal instead of uniform data appears to have a similar effect of adding complexity in condition *Broad*. Third, a signal value outside of the range of X directly raises subjects' attention to the issue of nuisance neglect in *Narrow* and leads to an increased share of Bayesian beliefs. However, a substantial fraction of subjects instead jump to closest value in the support of X. Fourth, if X and Y are correlated instead of independent, the median subject in *Narrow* displays lower inattention. One explanation is that subjects accommodate the additional incentive to attend to Y that is induced by the correlation. At the same time, the presentational format of the distributions changes in this task to illustrate the correlation, which plausibly affects subjects' perception of the problem and so this task allows no definite conclusion.

Table 6: Overview of robustness tasks

Task	Sample space X	Sample space Y	Signal type	Signal value
Correlated data (r=0.7)	$\{95, 96, \ldots, 104, 105\}$	$\{-15, -14, \ldots, 14, 15\}$	$(X+Y) \div 2$	104
Larger sample space (> 10)	$\{190, 191, \ldots, 209, 210\}$	$\{180, 181, \ldots, 219, 220\}$	$(X+Y) \div 2$	208
Discrete normally distributed numbers	$\{170, 180, \ldots 220, 230\}$	$\{-50, -40, \ldots 40, 50\}$	$X+Y$	220
Signal out of X range	$\{240, 241, \ldots, 259, 260\}$	$\{-15, -14, \ldots, 14, 15\}$	$X+Y$	230

Notes: This table provides an overview of the four robustness belief tasks. The distributions of X and Y as well as the signal structure are identical in both treatment conditions. X and Y were independently drawn from two discrete uniform distribution, i.e., every indicated outcome was equally likely.

Table 7: Median inattention in robustness tasks

Task	Median inattention θ_{Mdn}		Mann-Whitney U test
	Narrow N=72	*Broad* N=72	(p-value)
Correlated data (r=0.7)	0.59	0.00	< 0.001
Larger sample space (> 10)	1.00	0.33	< 0.001
Discrete normally distributed numbers	0.44	0.27	< 0.001
Signal out of X range	0.49	0.17	< 0.001

Notes: This table displays group medians of implied inattention parameters by treatment condition for four additional belief formation tasks. Inattention is calculated as $\theta = \frac{H_B}{H_B + H_N}$, where H_B and H_N denote the Hellinger distance of the subjective distribution to the Bayesian posterior and the inattentive posterior distribution, respectively. Task order was randomized within each of the two blocks. 72 subjects participated in each condition.

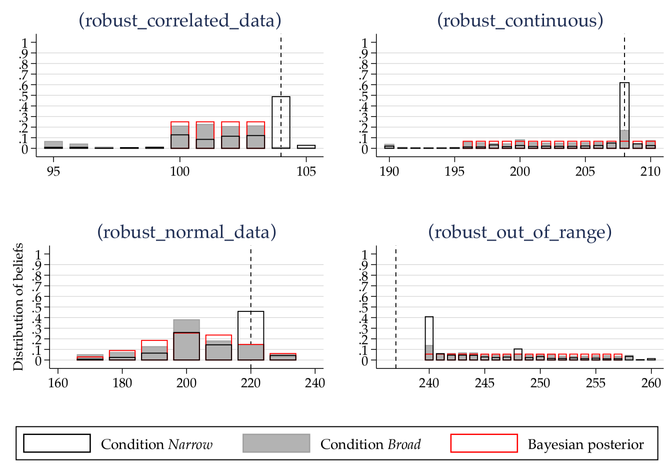


Figure 7: Distribution of elicited belief distributions about X in each one of four robustness tasks. N=72 for each condition in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the vertical dashed line. The conrresponding task configurations are shown in Table 6.

C.2 Face Value Heuristic and Anchoring

In treatment *Computation*, a simple algebraic computation is added on top of the signal structure. The resulting signals provided in the five baseline tasks are "average of X and Y minus $(3 \cdot 5)$ plus 35", "sum of X and Y plus $(2 \cdot 10)$ minus 30", "sum of X and Y plus 40 minus $(4 \cdot 5)$", "average of X and Y minus $(8 \cdot 5)$ plus 10", and "average of X and Y plus $(3 \cdot 5)$ plus 10".

The computations are chosen in such a way that anchoring on the signal value remains equally plausible. If subjects apply a simple face value heuristic, they should ignore both the the computation and the variation of Y. Figure 8 shows raw beliefs in condition *Computation*, including the initial value of the signal and the signal realization *after accounting for the computation*. There is limited evidence for anchoring on the signal value. Subjects do not appear to take the signal at face value per se: in condition *Computation*, they consistently account for the computation and then still neglect Y leading to nuisance neglect.

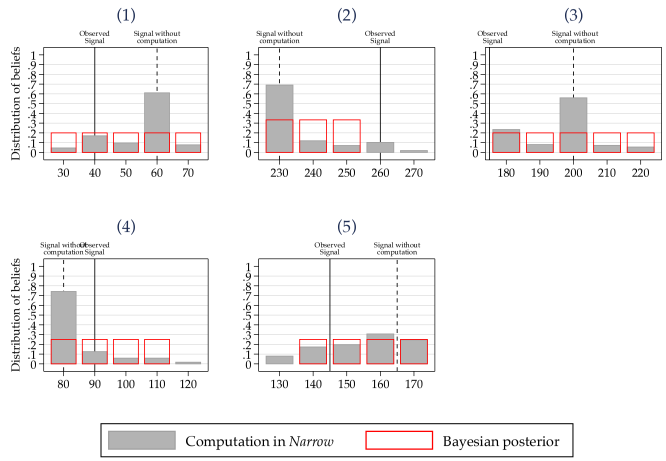


Figure 8: Distribution of elicited belief distributions about X in condition *Computation*. N=24 in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the solid dashed line, and the signal value after undoing the computation is shown by the dashed line. In all five tasks, X and Y follow independent discrete uniform distributions that were shown to subjects. Task order was randomized.

C.3 Elicitation Procedure

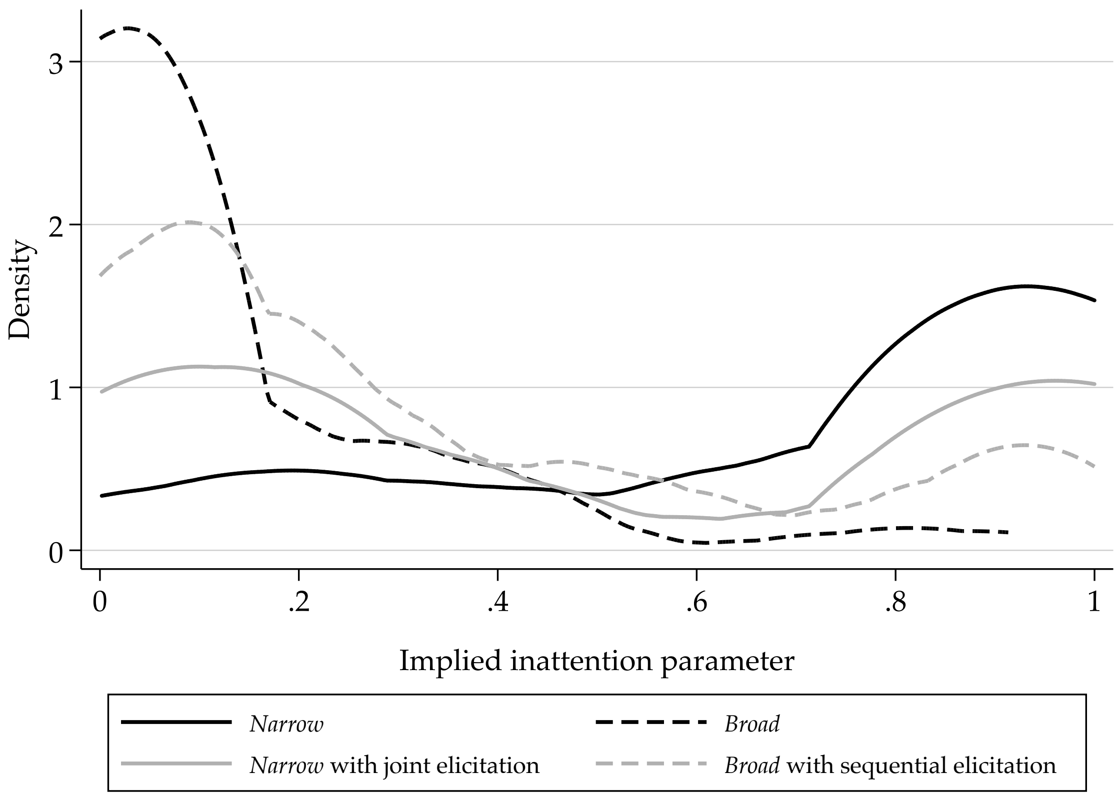


Figure 9: Subject-level mean of inattention to Y in four conditions. Based on N=216. For each subject I calculate the mean inattention in the five baseline tasks. The curves show kernel density estimates for each treatment (*Narrow* N=72, *Broad* N=72, *Narrow with joint elicitation* N=24, *Broad with sequential elicitation* N=48). A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention. Epanechnikov kernel with bandwidth 0.1.

In treatments *Narrow* and *Broad*, prediction incentives are different, but the elicitation method also differs. In *Narrow*, subjects only indicate the marginal distribution of X, while in Broad, subjects indicate the joint distribution of X and Y. To test whether treatment effects are driven by this difference in what is elicited, I design two additional treatments. In *Narrow with joint elicitation*, only the prediction of X is incentivized (as in *Narrow*) but the joint distribution is elicited exactly as in *Broad* (on a single screen). In *Broad with sequential elicitation*, guesses of both X and Y are incentivized (as in *Broad*), but now the subject first indicates the marginal of X, and then indicates the marginal of Y on a separate screen. This way, the first screen (for the marginal of X) is exactly identical to *Narrow*. Figure 9 plots kernel density estimates of the within-subject mean of inattention in the five belief tasks for all four treatments. Mean inattention in the four treatments is 0.25 (*Broad*), 0.34 (*Broad with sequential elicitation*), 0.47 (*Narrow with joint elicitation*), and 0.57 (*Narrow*). These findings imply that the treatment effect is not an artifact of different elicitation methods. Harmonizing the elicitation procedure somewhat reduces the effect in the predicted direction, but prediction incentives as such have a distinct effect.

C.4 Simplification

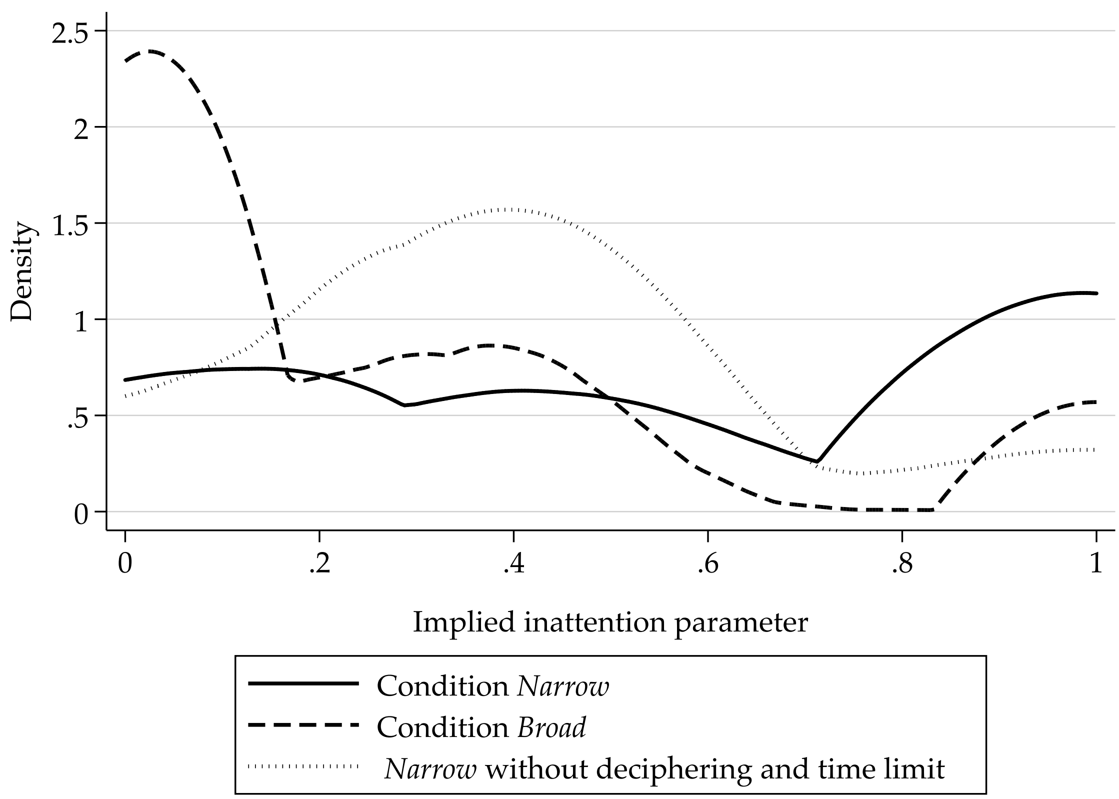


Figure 10: Implied inattention to Y in three conditions. Based on 1,944 stated beliefs. The curves show kernel density estimates for each treatment (*Narrow* N=864, *Broad* N=864, *Simplification* N=216). A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention. Epanechnikov kernel with bandwidth 0.1.

To study the role of complexity in the experimental setup, an additional condition drastically simplifies the experimental procedure by removing the deciphering stage as well as all time constraints. In this treatment, subjects are paid to predict X as in Narrow, but they do not have to decipher the signal and have unlimited time to indicate their guess. Effectively, they are given the distributions of X and Y, and immediately see the value of the signal. We find a statistically significant reduction in inattention relative to Narrow in this case ($p = 0.00$). At the same time, inattention remains far higher than in Broad ($p = 0.00$). Mean inattention is 0.57 in *Narrow*, 0.40 in *Simplification*, and 0.25 in *Broad*. Also, there is somewhat reduced bunching at fully inattentive and fully Bayesian beliefs. Considerable simplifications improve predictions, but do not eliminate the effect of Narrow incentives. Figure 10 plots kernel density estimates of the distribution of inattention parameters in *Simplification* together with *Narrow* and *Broad* for reference.

A possible explanation of nuisance neglect is that subjects use the heuristic of reporting back the signal value, akin to exact anchoring or taking the signal at face value. Treatment *Computation* tests the face value explanation, under which the observed bias does not reflect the specific neglect of Y, but a form of extreme simplification strategy. Treatment *Computation* is identical to *Narrow*, but adds a simple algebraic computation into the information structure, in a way that it remains equally plausible to anchor on the observed signal value. For example, instead of $S = \frac{X+Y}{2}$, subjects receive the modified signal $S = \frac{X+Y}{2} - (2 \cdot 10) + 30$. We find minimal evidence for anchoring on the observed signal. Instead, subjects are able and willing to invert the computations, but still do not account for Y.[27] Computed inattention scores for beliefs in *Computation* are indistinguishable from *Narrow* ($p = 0.37$, M-W U test), and significantly different from *Broad* ($p < 0.001$). This suggests that nuisance neglect reflects a specific error in probabilistic reasoning rather than mere anchoring on the signal value.

[27] Further treatment details, figures and results are relegated to Appendix C.2.

D Online Experiments

I complement the baseline evidence from the laboratory with online experiments for two reasons. First, I seek to replicate nuisance neglect in a large, more heterogeneous population and under less controlled choice conditions. While a lack of attention and motivation among online participants compared to laboratory subjects can be a matter of concern, online experiments can help to establish the robustness and generalizability of attention-related phenomena in more distracting and ecologically valid environments. All online experiments are conducted on Amazon Mechanical Turk (Mturk), which is widely used in recent work in experimental economics (DellaVigna and Pope, 2018; Martínez-Marquina et al., 2019). Second, the online platform allows to run multiple treatment variations with a large number of participants, which is difficult using laboratory samples (Robinson et al., 2019).

I implement four design modifications relative to the laboratory study. First, subjects do not have to indicate a full posterior distribution but are incentivized to state the mean of their posterior belief, substantially simplifying the elicitation procedure of the laboratory study. Second, X and Y are drawn from distributions with a larger sample space. The baseline task specifications are reported in Appendix Table 8. Third, there is no deciphering stage preceding the belief elicitation. Fourth, in the online study, X and Y are drawn by the computer at the individual level instead of jointly for all subjects. All design and procedural detail as well as detailed results are reported in Appendix D.

The main finding of substantial nuisance neglect replicates in the online study. Specifically, 53% of stated beliefs imply an attention parameter θ above 0.9. In addition, I document evidence for an additional updating mode, "signal neglect" or non-updating, a frequent finding in belief formation studies (Coutts, 2019; Henckel et al., 2018; Möbius et al., 2014).[28] The replication of nuisance neglect in the online study motivates that further experiments in Sections 4 and 4.3 are conducted online instead of in the laboratory.

[28] Non-updating could not be unambiguously identified in some tasks of the laboratory study due to the elicitation of belief distributions rather than point beliefs and the use of uniform data structures.

Procedures. I conduct incentivized experiments on Amazon Mechanical Turk (MTurk), an online labor marketplace frequently used for experimental economics research. To qualify for participation, MTurk workers have to be residents of the U.S. and of legal age, have an overall approval rating of more than 95%, and have successfully completed more than 100 assignments on MTurk. Workers are paid 0.5 dollars for participation and can earn up to 3 dollars for their performance on the guessing task. They play five rounds in randomized order. An example of the urn-based representation of distributions is reproduced in Appendix Figure 35. One round is randomly selected for payment in the end, and the payoff is determined based on a quadratic scoring rule.[29] In the online experiments, all subjects are paid to predict X only, analogous to condition *Narrow* in the laboratory experiments. 131 subjects participated in the online baseline experiments for an average payment of 1.7 dollars. Completion of the study took 13 minutes on average. It was implemented using oTree (Chen et al., 2016).

Table 8: Online baseline tasks

X	Y	I
$\mathcal{N}(100,100)$	$\mathcal{N}(0,100)$	$X+Y$
$\mathcal{N}(100,100)$	$\mathcal{N}(0,400)$	$X+Y$
$\mathcal{N}(100,400)$	$\mathcal{N}(0,100)$	$X+Y$
$\mathcal{U}[75,76,\dots,125]$	$\mathcal{U}[-25,-24,\dots,25]$	$X+Y$
$\mathcal{U}[75,76,\dots,125]$	$\mathcal{U}[90,91,\dots,110]$	$\frac{X+Y}{2}$

Notes: This table provides an overview of the five baseline belief tasks in the online experiment. For all normally distributed variables, the support was discretized to integers, truncated at $\mu-50$ and $\mu+50$ and then the distributions were scaled such that the they have unit probability mass.

Baseline Results: Online Study. Figure 11 shows all stated beliefs together with the signal realizations observed in the five baseline tasks. It further highlights which stated beliefs correspond to nuisance neglect, signal neglect and Bayesian updating.

[29]The monetary payoff (in US dollars) is calculated as:

$$\max\left\{0\ ,\ 3-0.2\cdot(\text{guess of } X-\text{draw of } X)^2\right\}$$

There is evidence for each of those three updating rules. In all tasks at least 60% of stated beliefs are exactly in line with these three updating modes. Among the three modes, Bayesian updating and nuisance neglect are observed with roughly similar frequency, while signal neglect occurs to a lesser extent.

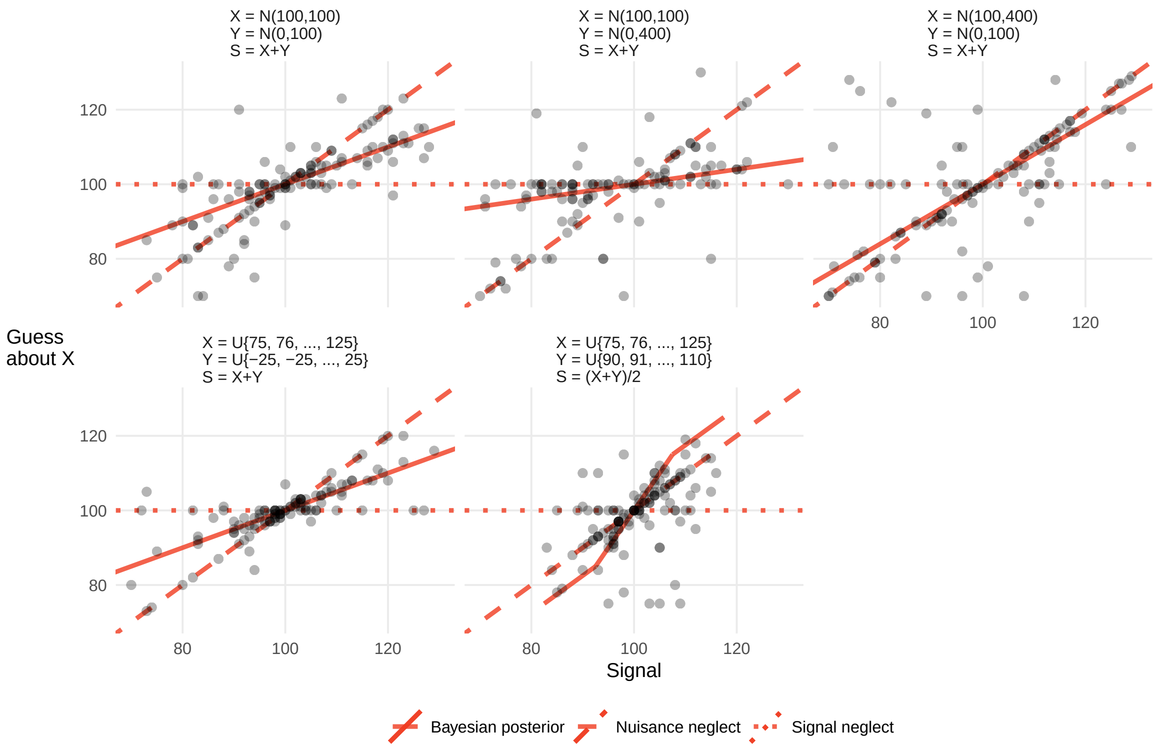


Figure 11: Beliefs in baseline tasks of online experiments. N=131 in each task. Each dot corresponds to one stated belief. The three red lines indicate the Bayesian benchmark, nuisance neglect, and signal neglect.

To illustrate the degree to which beliefs are clustered on these three updating modes, Figure 12 plots kernel density estimates for the task in the upper left corner of Figure 11. In this task, $X \sim \mathcal{N}(100, 100)$, $Y \sim \mathcal{N}(0, 100)$, and $S = X + Y$. The stated belief that corresponds to a Bayesian posterior in this case is $100 + \lambda \cdot (s - 100)$ where $\lambda = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2} = \frac{1}{2}$. Intuitively, since X and Y have equal variance, a normatively optimal guess of X attributes half of s's deviation from the expected value of 100 to X. Signal neglect, in turn, corresponds to a belief equal to the prior of X, $\mathbb{E}[X] = 100$. This is equivalent to assigning *none* of the deviation of s from its expected value to X. In fact, with m denoting a subject's stated guess, the empirical equivalent of λ can be backed out as $\hat{\lambda} = \frac{m-100}{s-100}$. In the case of signal neglect with $m = 100$, $\hat{\lambda} = 0$. Finally, if people commit nuisance neglect they state $m = s$, implying $\hat{\lambda} = 1$.

Figure 12 provides three insights. First, most of the probability mass is centered on the three updating modes. Second, nuisance neglect is the most frequent mode in this task, and signal neglect the least frequent one. Third, as indicated by the rug plot on the right, most people who neglect Y ($\hat{\lambda} = 1$) or the information ($\hat{\lambda} = 0$) do so *exactly*. By contrast, people are more dispersed around the Bayesian benchmark ($\hat{\lambda} \approx 0.5$), presumably because it is harder to compute the Bayesian posterior exactly.

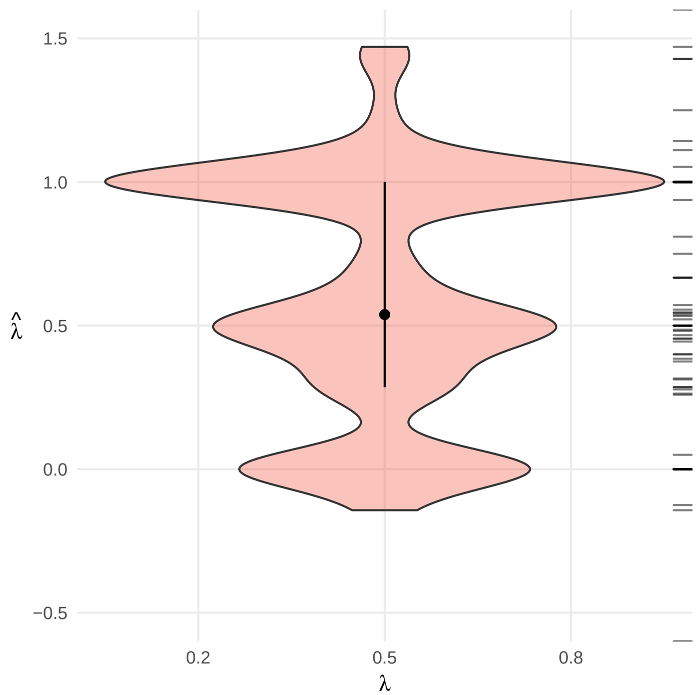


Figure 12: Kernel density plot for beliefs stated in a task where $X \sim \mathcal{N}(100, 100)$, $Y \sim \mathcal{N}(0, 100)$, and $S = X + Y$. In this task, the Bayesian belief corresponds to $100 + \lambda \cdot (s - 100)$ where $\lambda = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2} = \frac{1}{2}$. For each stated belief, the empirical counterpart of λ is calculated as $\hat{\lambda} = \frac{m-100}{s-100}$. The plot documents three distinct clusters at $\hat{\lambda} = 0$ (signal neglect), $\hat{\lambda} = 1$ (nuisance neglect) and around $\hat{\lambda} = \frac{1}{2}$ (Bayesian posterior.) Based on N=131. Epanechnikov kernel with bandwidth 0.07.

In the experimental settings studied in this paper, beliefs are clearly too heterogeneous to be adequately described by a single representative updating rule. Average beliefs mask the underlying structure. At the same time, there is little randomness in stated beliefs. Instead, most beliefs accord to a discrete set of three updating modes. They align exactly with one of these modes, and there is virtually no mixing between the modes, i.e., people do not seem to choose combinations of updating rules.

The main finding of substantial nuisance neglect in the laboratory replicates in the online study. In addition, I document evidence for an additional updating mode, signal neglect or non-updating, which is in line with typical findings in studies on belief formation. Consistent with the results reported on within-subject heterogeneity of updating rules in the laboratory experiment (Section 3.1.5), We find that 37% ob subjects report all of their belief in line with nuisance neglect, 12% always exhibit signal neglect and 8% are always close to Bayesian. Correspondingly, 47% switch at least once between updating modes. Again in line with the laboratory results, we find that a large fraction of beliefs (73%) stated following a belief that can be characterized as close-to-Bayesian was also close-to-Bayesian, indicating the relevance of "insight."

The baseline online study again builds on the working assumption from Section 3.1.2 that the form of nuisance neglect can be characterized as people (implicitly) using an alternative, simplified signal structure. I test this assumption in the following Section.

D.1 Characterizing the Form of Nuisance Neglect Across Information Structures

The baseline experiments implement the working assumption from Section 3.1.2 that inattention to Y may be characterized as people taking the signal as fully revealing about X, i.e., $X = s$. Note that some of the evidence so far can be explained in alternative ways: subjects might, for example, shrink the variance of Y in the updating process, or use a non-Bayesian updating rule that underweights the base rate (Bar-Hillel, 1980) or overweights the likelihood ratio as in *diagnostic* expectation formation (Bordalo et al., 2018, 2019). A characterization of the form of nuisance neglect across information structures is necessary, first, to incorporate it in formal models and distinguish between competing theories and, second, as an input to studying its underlying mechanisms. I characterize different candidate characterizations based on whether they correspond to the (implicit or explicit) use of (i) an alternative subjective signal structure S_i, (ii) a distorted distribution of Y, h_Y, or (iii) a non-Bayesian updating rule. These explanations need not be mutually exclusive. In an additional series of online experiments, subjects face various tasks that allow to distinguish between them. This evidence is reported in Appendix D.1.

I summarize three main findings. First, nuisance neglect is incompatible with likelihood-based models, specifically diagnostic expectations. People form diagnostic expectations if they overweight outcomes that become more likely upon arrival of new information (Bordalo et al., 2018). However, in my data people typically overweight outcomes of X that are close to s, even if these outcomes have become *less* likely under s. For example, consider two variables X and Y that are independent and uniformly distributed as $\mathscr{U}\{50, 51, ..., 150\}$, and signal structure $S = X + Y$.[30] Upon observing, e.g., $s = 145$, diagnostic expectations imply overweighting of small values of X below 100. In the experiment, however, subjects overweight outcomes of X *above* 100, as if trying to explain the signal solely through X (cf. Appendix Figure 14). Relatedly, empirical beliefs do not feature the *kernel of truth* property of diagnostic expectations, which implies that beliefs generally respond to news in a directionally correct, but excessive manner. In the experiment, subjects also respond to news that is fully uninformative about X.

Second, we find that nuisance neglect is not consistent with Bayesian updating under a distorted prior about Y. Specifically, in several tasks, stated beliefs about X are incompatible with any possible subjective belief about Y on the union of the actual support of Y, the mean, median and mode of Y, and a value of 0. This contradicts any rule that replaces Y by a single value in its support, its mean, etc., as well as any rule that shrinks the variance of Y.

Third, I document the following reduced-form patterns. If s is in the support of X, nuisance neglect corresponds to overweighting the outcome(s) closest to s. If s is not in the support of X but "sufficiently close", posterior beliefs allocate excessive mass to the outcomes in the support of X that are closest to s. If s is not in the support of X and sufficiently far from any value with positive likelihood, the share of nuisance neglect

[30] See Appendix Table 9.

substantially decreases.[31]

The upshot of this analysis is that, for the data and signal structures analyzed here, nuisance neglect is best characterized as a strong form of ignorance about the existence of Y: people seem to apply a modified signal structure S_i that excludes Y.

Table 9: Online experiment on form of nuisance neglect

X	Y	Info structure I	Observed info i
$\mathscr{U}\{75,76,\ldots,125\}$	$\mathscr{U}\{90,91,\ldots,110\}$	$\frac{X+Y}{2}$	Individual draw
$\mathscr{U}\{75,76,\ldots,125\}$	$\mathscr{U}\{-25,-24,\ldots,25\}$	$X+Y$	Individual draw
$\mathscr{N}(100,400)$	$\mathscr{N}(50,100)$	$\frac{X+Y}{2}$	Individual draw
$\mathscr{N}(100,400)$	$\mathscr{N}(100,100)$	$\frac{X+Y}{2}$	Individual draw
$\mathscr{N}(100,400)$	$\mathscr{N}(100,100)$	$X+Y$	Individual draw
$\mathscr{U}\{50,51,\ldots,150\}$	$\mathscr{U}\{50,51,\ldots,150\}$	$X+Y$	145
$\mathscr{U}\{75,76,\ldots,125\}$	$\mathscr{U}\{90,91,\ldots,110\}$	$\frac{X+Y}{2}$	116
$\mathscr{N}(100,400)$	$\mathscr{N}(100,100)$	$2\cdot X+2\cdot Y$	412
$\mathscr{N}(100,400)$	$\mathscr{N}(100,100)$	$2\cdot X+Y$	266
$\mathscr{N}(100,400)$	$\mathscr{N}(100,100)$	$X+Y$	110

Notes: This table provides an overview of the ten belief tasks in the online experiment on the form of nuisance neglect. Note that for all normally distributed variables, the support was discretized to integers, truncated at $\mu-50$ and $\mu+50$ and then the distributions were scaled such that the they have unit probability mass.

Table 9 displays the ten tasks used in an online experiment on the form of nuisance neglect with 79 subjects recruited from Mturk. In five of those tasks, information values are drawn individually for each subject, while in the remaining tasks one information value is drawn jointly for all subjects to obtain higher power for a specific realization.

Figures 13 and 14 illustrate the corresponding results. In each of the tasks in Figure 13, the solid reference line corresponds to Bayesian posteriors while the dashed line indicates

[31]Note that these results reflect the specific experimental design – that is, algebraic signal structures in which X and Y are combined additively – and should be viewed as such without implying any generality. In practice, information environments rarely have these features. Uncertainty about the exact form of the information structure may make people even more susceptible to nuisance neglect.

reference beliefs under nuisance neglect.

Figure 14 demonstrates that the form of nuisance neglect is not generally in line with people using a modified distribution of Y. To see this, the green line indicates a corresponding threshold: all beliefs on the opposite side of the Bayesian posterior are not compatible with *any* possible implied distribution of Y on the actual support of Y. At the same time, these tasks indicate that nuisance neglect is not easily reconciled with oversensitivity to the likelihood (or neglect of base rates), as would be in line with, e.g., diagnostic expectations (Bordalo et al., 2018). Consider for example the task displayed in the upper right corner of Figure 14, where $X \sim \mathscr{U}\{50, 51, \ldots, 150\}$, $Y \sim \mathscr{U}\{50, 51, \ldots, 150\}$, $I = X + Y$ and $i = 145$. Here, an information value of 145 indicates that a relatively small value of X, i.e., $x < 100$, has been drawn, and the likelihood increase is greatest for values of X below 100. However, people predominantly choose values above 100, close to 145.

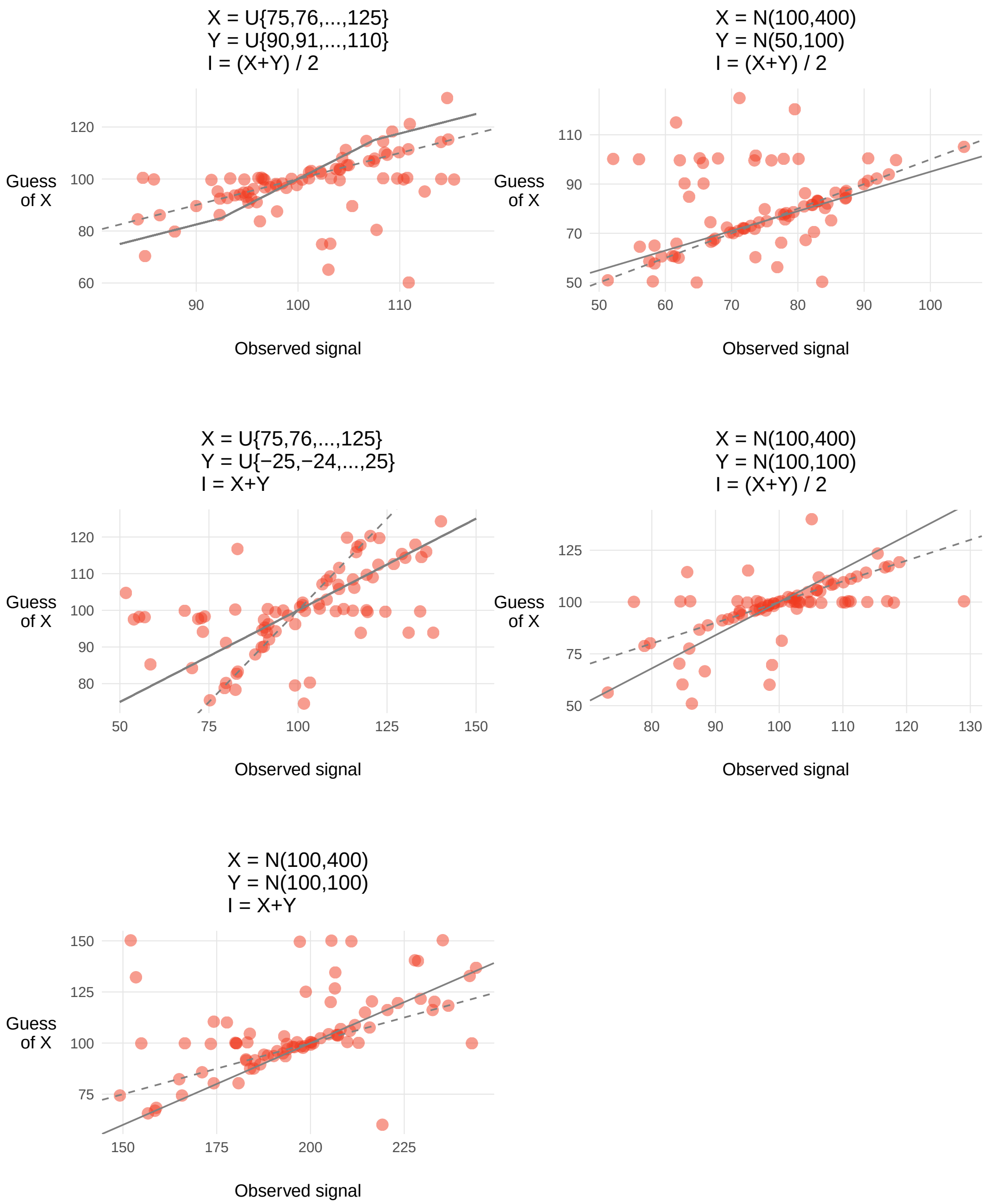


Figure 13: Raw beliefs in online experiment on the form of nuisance neglect. The solid reference line indicates the Bayesian posterior, the dashed line shows nuisance neglect. N = 79 in each task. Displayed are the five out of ten tasks in which the information value was individually drawn for each subject. The task order (of all ten tasks) was randomized at the individual level.

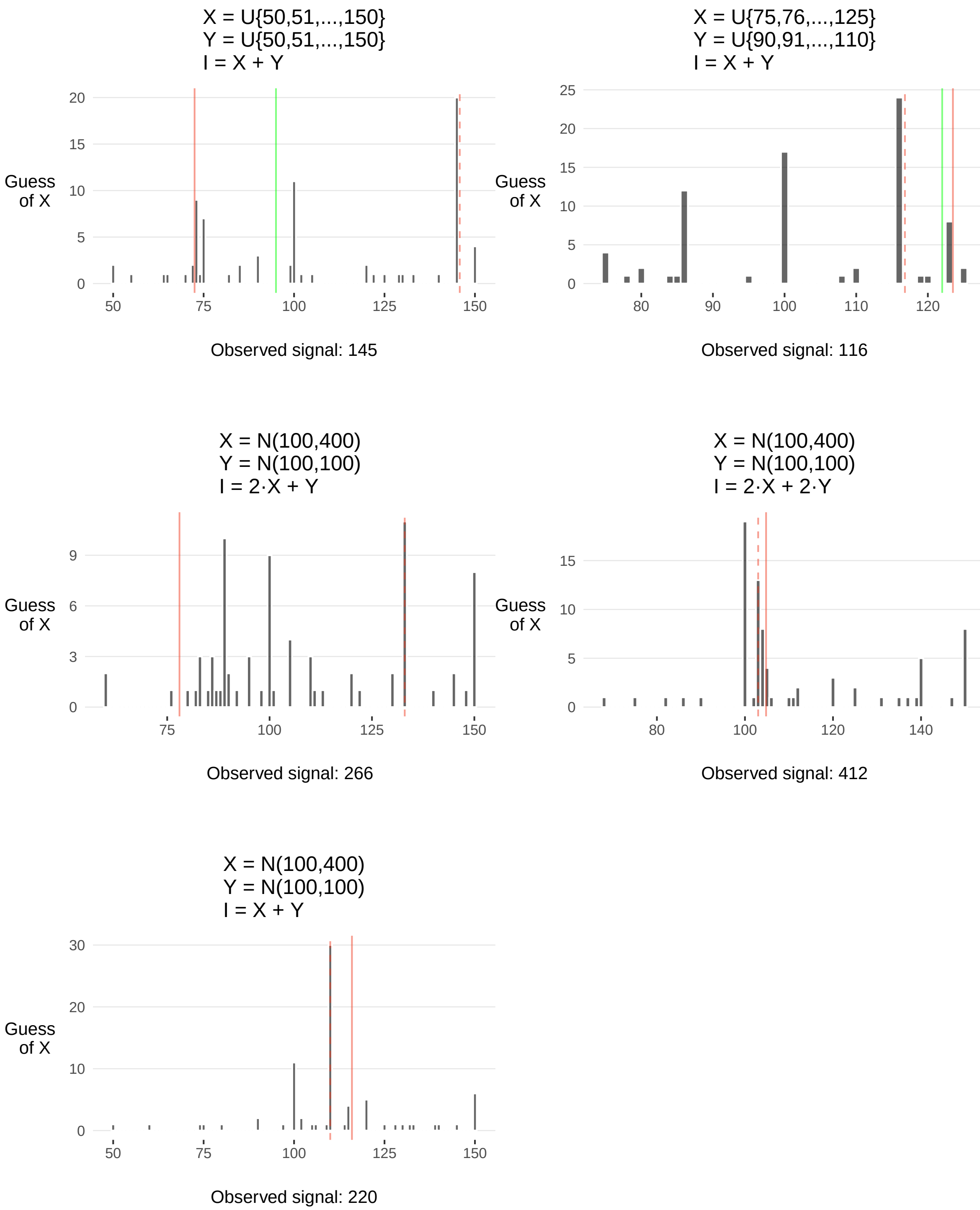


Figure 14: Raw beliefs in online experiment on the form of nuisance neglect. The solid red reference line indicates the Bayesian posterior, the dashed red line shows nuisance neglect. The green line indicates a threshold. All belief on the opposite side of the Bayesian posterior are not compatible with any possible implied distribution of Y on the actual support of Y. These tasks therefore provide evidence against the idea that nuisance neglect is in line with people using a modified distribution of Y. N = 79 in each task. Displayed are the five out of ten tasks in which all subjects observed the same information value. The task order (of all ten tasks) was randomized at the individual level.

E Mechanism Treatments

For all online mechanism experiments the five tasks displayed in Table 10 were used.

Table 10: Online tasks in mechanism treatments

X	Y	I
$\mathcal{N}(100,100)$	$\mathcal{N}(0,100)$	$X+Y$
$\mathcal{N}(100,100)$	$\mathcal{N}(0,100)$	$X+Y$
$\mathcal{N}(100,100)$	$\mathcal{N}(0,400)$	$X+Y$
$\mathcal{U}[75,76,\ldots,125]$	$\mathcal{U}[-25,-24,\ldots,25]$	$X+Y$
$\mathcal{U}[75,76,\ldots,125]$	$\mathcal{U}[90,91,\ldots,110]$	$\frac{X+Y}{2}$

Notes: This table provides an overview of the five baseline belief tasks in the online mechanism experiments. Note that for all normally distributed variables, the support was discretized to integers, truncated at $\mu-50$ and $\mu+50$ and then the distributions were scaled to have unit probability mass.

E.1 Hint Treatment

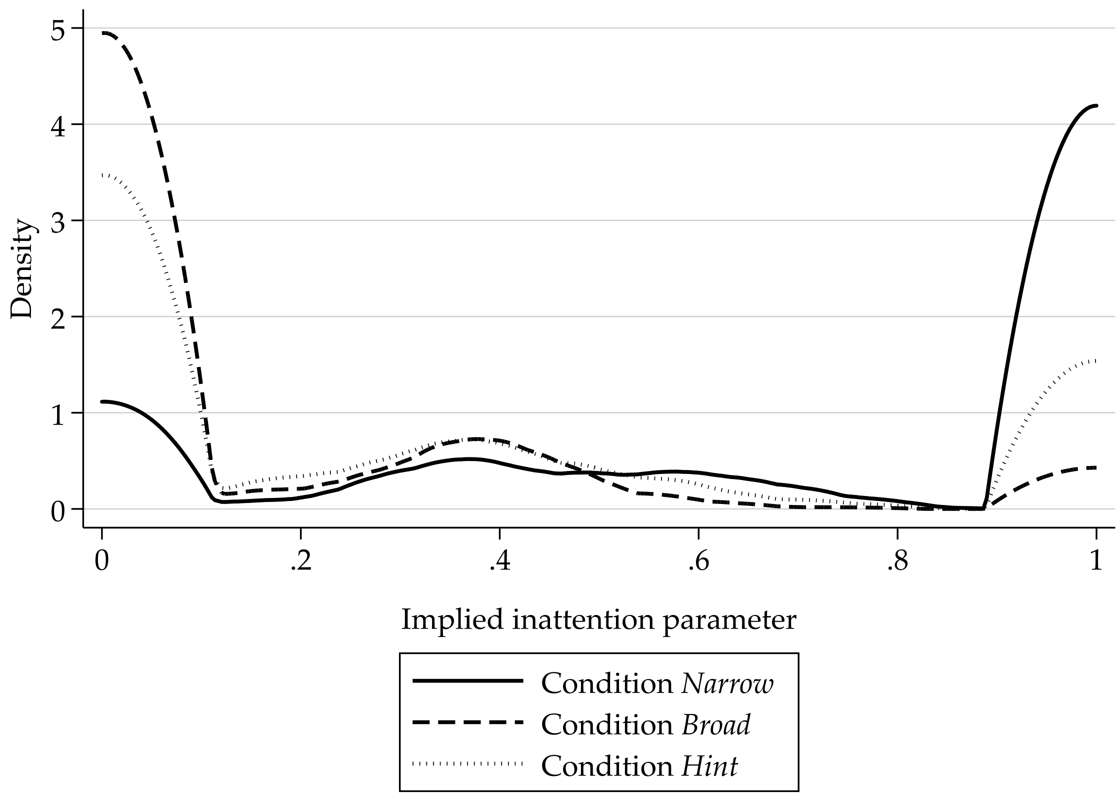


Figure 15: Implied inattention to Y in three conditions. Based on 950 stated beliefs. The curves show kernel density estimates for each treatment (*Narrow* N=360, *Broad* N=360, *Hint* N=230). A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete nuisance neglect. Epanechnikov kernel with bandwidth 0.05.

E.2 Confidence Ratings

After finishing the baseline, robustness and bonus belief tasks in the laboratory, each of the tasks was again presented successively including all previously shown information as well as the subject's stated guess. In a list with fixed monetary amounts from 0 euros to 6.25 euros in steps of 0.25 euros, subjects then indicated whether they prefer to be paid out for their stated belief, or receive this fixed amount, in case this belief task would be selected to count. Single switching was enforced. Figure 16 shows that implied inattention of the belief and stated valuations for the belief are virtually unrelated.

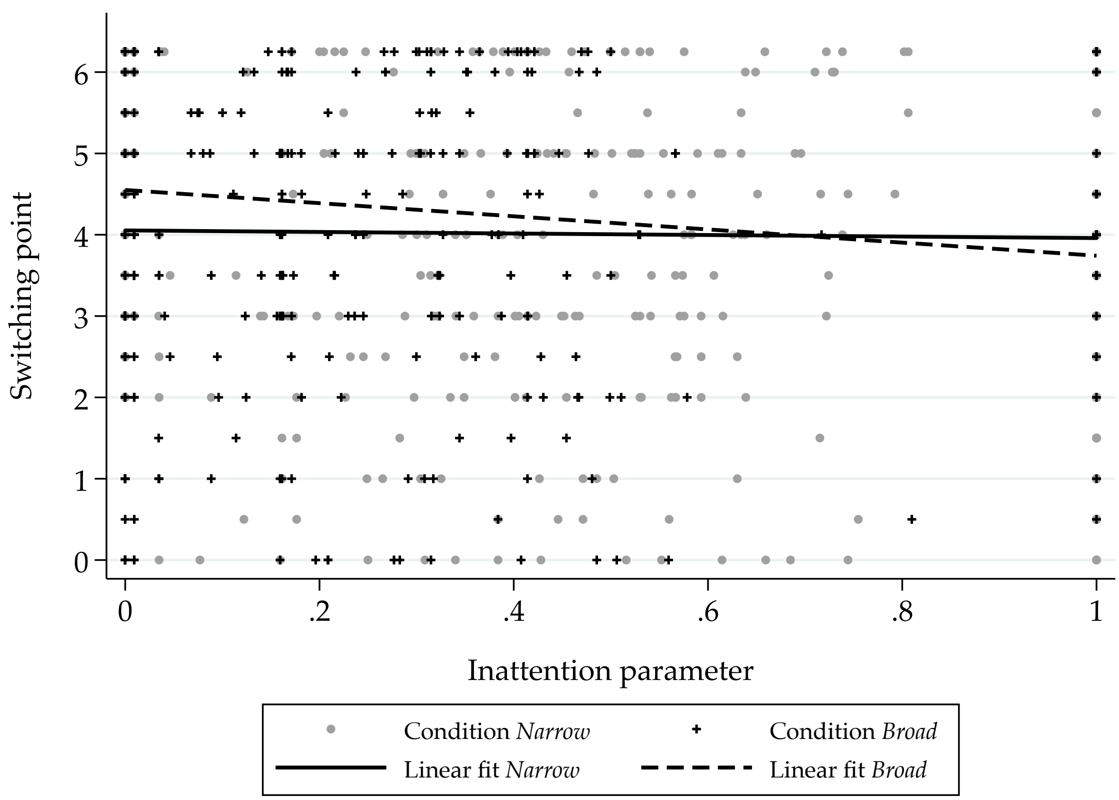


Figure 16: Scatterplot and linear regression fits for valuations of stated beliefs and implied inattention by condition. Based on N=360 each for condition *Narrow* and condition *Broad*.

E.3 Switch-Role Tasks

As the last part of the main baseline experiment, i.e., following the confidence tasks, subjects were (unexpectedly) presented with two additional tasks in which roles were switched with the respective other condition. The switch-role task configurations were comparable to those of the baseline tasks. Figure 17 displays group means of inattention for each of the blocks of tasks by condition. Having previously predicted X and Y in condition *Broad* makes subjects somewhat less inattentive than in the *Narrow* baseline, but not by much. A highly significant reverse treatment effect persists in teh switch-role tasks.

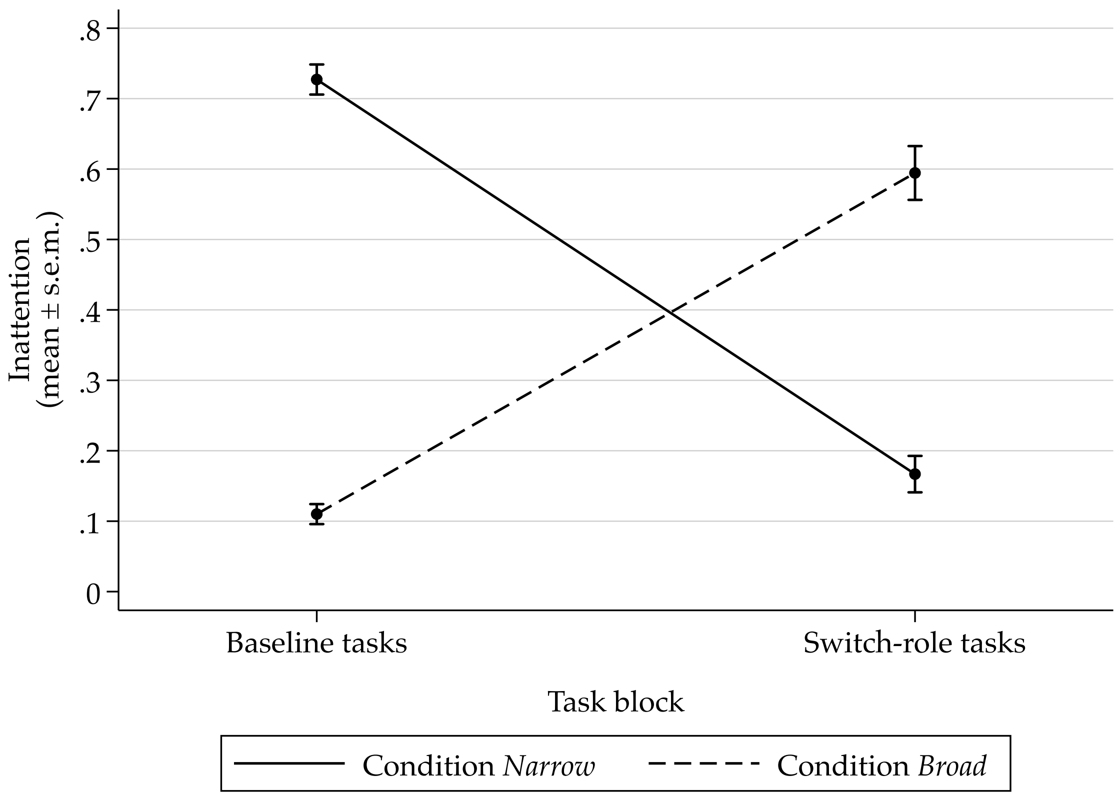


Figure 17: Group means of inattention by task block and condition. Based on N=360 baseline beliefs and N=144 switch-role beliefs each for condition *Narrow* and condition *Broad* .

In Table 3 I analyze inattention scores as defined in equation (6) on the pooled sample of beliefs from the baseline and *Switch-role* tasks. We find that *(i)* unsurprisingly, *Narrow* subjects almost immediately improve when facing the broad setup, and display a similar level of inattention as *Broad* subjects in the baseline ($p > 0.7$, see footer of Table 3), *(ii)* *Broad* subjects do transfer their experience in forming Bayesian beliefs, as indicated by a significant improvement relative to *Narrow* subjects in baseline ($p < 0.05$), *(iii)* this transfer, however, is far from perfect and a significant treatment effect between *Narrow* and *Broad* persists in the *Switch-role* tasks, albeit now with the reverse sign. In fact, mean inattention in *Broad* is 0.59 in *Switch-role*, compared to baseline means of 0.11 in *Broad* and 0.73 in *Narrow*. Put differently, the improvement in *Broad* is small and subjects effectively commit inattentive inference to a roughly similar extent as if they had not made the baseline experience.

E.4 The Role of Effort: Manipulation of Stake Size

In treatment *High Stakes*, the possible prize is raised five-fold relative to the baseline online experiment. Under higher incentives, we find that the prevalence of Bayesian updating increases statistically significantly, but fully at the expense of signal neglect, see Figure 4. Strikingly, the share of nuisance neglect remains roughly constant. Effort as measured by response times increases significantly, both overall and within each subgroup (pairwise t tests, all $p < 0.001$). This means, given higher incentives, subjects try harder, but that only affects non-updating, reducing the fraction of subjects that ignore the signal altogether. On average, higher effort does not reduce nuisance neglect, however. Compellingly, a tenfold increase of the stake size in the laboratory experiment leads to a similar pattern, see Appendix Section E.4. This indicates that psychic costs, cognitive miserliness, laziness or effort reduction may explain non-updating, but have limited explanatory power for nuisance neglect.

Note that these findings square with the result from Section 4.2.2 on the effect of enforced deliberation time when considering their differential effect on attention. By disabling the belief elicitation tool, treatment *Enforced Deliberation* may nudge subjects into specifically re-considering their solution strategy, which leads some to figure out the role of Y. Treatment *High Stakes* does not direct effort in this way. Higher effort per se is insufficient to reduce nuisance neglect, perhaps because it is targeted at the computational elements rather than subjects' mental representation of the task.

Table 11: Inattentive inference and effort

Tasks:	Bonus round (variation of stakes)		
Conditions:	*Narrow* and *Broad*		*Narrow*
Dependent variable:	**Response time**	**Inattention θ**	
	(1)	(2)	(3)
High stakes in bonus task	19.778**	-0.017	0.001
	(8.711)	(0.054)	(0.100)
0 if *Broad*, 1 if *Narrow*	-32.444***	0.502***	
	(7.538)	(0.084)	
Treatment dummy * High stakes	-3.830	0.019	
	(11.795)	(0.113)	
Constant	66.111***	0.101**	0.603***
	(5.671)	(0.042)	(0.073)
R^2	0.24	0.37	0.00
# Observations	144	144	72

Notes: OLS regressions. In the bonus round I randomly vary within each treatment whether incentives are 1 euro or 10 euros. Response time is the duration in seconds the subject spent on the belief elicitation page. Inattention is calculated as $\theta = \frac{H_B}{H_B+H_N}$, where H_B and H_N denote the Hellinger distance of the subjective distribution to the Bayesian posterior and the inattentive posterior distribution, respectively. Robust standard errors clustered at participant level in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

E.5 The Effect of the Specific Cost of Nuisance Neglect

The stake size manipulation changes the benefit of forming optimal Bayesian beliefs, which means it simultaneously increases the cost of *all* deviations from optimality. In the following, I investigate whether the *specific* cost of nuisance neglect affects its prevalence. In economic models of rational belief formation, the likelihood of committing a specific error depends on its expected cost in utility terms (Caplin and Dean, 2015; Gabaix, 2014; Wiederholt, 2010). On that account, the prevalence of nuisance neglect should vary systematically with its expected accuracy in a given information environment.

Two remarks about this exercise are in order. First, one way of conceptualizing the role of Y is in terms of the amount of variance and the amount of bias that Y introduces into the signal structure relative to a fully revealing signal structure $S = X$. As the signal-to-noise ratio $\frac{\sigma_X^2}{\sigma_Y^2}$ increases, a neglect of Y induces fewer distortions. Similarly, as the directional bias $\left|\mathbb{E}[S] - \mathbb{E}[X]\right|$ increases due to a shift in the mean of Y, neglecting Y leads to increasingly distorted beliefs about X. Therefore, I run two sets of experiments that fix the signal structure while separately varying the amount of variance and bias introduced by Y.

Second, the objective of this Section is to test the dependence of the attribution error on its costliness (the friction explanation) against the alternative, mental gap account of nuisance neglect. The variation of the cost of nuisance neglect should therefore not in itself direct attention to Y and close a potential mental gap in this way. Note, however, that an information structure in which nuisance neglect leads to larger distortions is associated with signal realizations that are more likely to lead to implausible (or ex-ante surprising) beliefs. A signal realization with low likelihood under nuisance neglect may in itself act as a cue that triggers a reconsideration of the problem interpretation. Take the following example, in which X and Y are both independently drawn from a (discretized) $\mathcal{N}(100, 10)$ distribution and the subject receives a signal $S = \frac{x+y}{2} = 107$. Intuitively, after observing $S = 107$ and believing that $S = X$, the agent allocates the entire probability mass to $X = 107$. The *plausibility* refers to how likely this specific posterior belief was prior to receiving the signal (interpreted under the flawed mental model). The plausibility equals the prior probability of X in this case, i.e., $P(X = 107)$. A low plausibility of the belief under nuisance neglect can in itself act as a hint and nudge subjects to realize their neglect of Y. Next, consider the otherwise identical situation where Y (but not X) is instead drawn from a (discretized) $\mathcal{N}(100, 50)$ distribution, so that Y introduces substantially more variation into the signal structure. Compared to the previous case, a low-plausibility posterior belief (under nuisance neglect) becomes more likely due to the high variance of Y, implying that even a mental gap explanation can rationalize reduced nuisance neglect on average in this task. At the same time, *for any given signal realization*, ex-post plausibility remains the same, even though the expected cost of neglecting Y increases across tasks. This means that only the friction explanation predicts an effect of the expected distortion *for a given signal realization*. In the following analysis, I will distinguish between the effect of the expected size of the distortion associated with nuisance neglect in a problem and that same effect conditional on the signal realization.

E.5.1 Treatment *Signal-to-Noise Ratio*

Treatment *Signal-to-noise ratio* varies the ratio between the variance of X, σ_X^2, and the variance of Y, σ_Y^2, across tasks while fixing all other parameters of the information structure at $\mu_X = 100$, $\mu_Y = 0$ and $S = X + Y$. Denote $\lambda = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}$ the fraction of the signal variation coming from X, which is a re-parameterization of the signal-to-noise ratio that is bounded by 0 and 1. All seven task specifications are listed in Appendix Table 12, which shows that λ varies between 0.015 to 0.985 across tasks. I conduct additional online experiments with a sample of $N = 209$ subjects. The task order is again randomized and one task is randomly incentivized with a prize of 3 dollars.

I define the empirical analogue of λ for the signal information structure studied here as $\hat{\lambda} = \frac{m-100}{s-100}$, such that $\hat{\lambda} = 1$ indicates nuisance neglect. Appendix figure 18 plots the estimated kernel densities of $\hat{\lambda}$ by task. Mirroring earlier findings, there are three pronounced empirical modes in each task, corresponding to Bayesian updating, nuisance neglect and signal neglect, i.e., non-updating.

I report two main findings. First, the share of nuisance neglect strongly increases in λ, which is in line both with people responding to the decreasing cost of nuisance neglect and a mental gap that is closed by low plausibility signals under nuisance neglect, as discussed above.[32] Second, as discussed above, signal realizations further away from 100 become more likely as the signal-to-noise ratio decreases, decreasing the ex-post plausibility $P(X = s)$. Using additional regression analyses we find that the effect of λ on the prevalence of nuisance neglect turns insignificant upon adding $P(X = s)$ as a control. This means, comparing two tasks with identically distributed X and identical signal realization s, but different variance of Y, there is no statistically significant difference in the propensity to commit nuisance neglect. Both findings are in line with the mental gaps explanation, but the latter is inconsistent with the friction explanation.

Appendix figure 18 documents the results by plotting estimated kernel densities of $\hat{\lambda} = \frac{m-100}{s-100}$ by task. In line with the previous results, there are three empirical modes in each task, corresponding to Bayesian updating, nuisance neglect and information neglect, i.e., non-updating. Note that the value of λ in line with Bayesian beliefs changes across tasks, as indicated by the dashed diagonal line. To support the visual analysis, I perform non-parametric test on the distributions of $\hat{\lambda}$. First, the summed share of beliefs in line with either one of the three updating modes (defined as being within $[\lambda-0.05, \lambda+0.05]$) does not significantly differ across tasks ($p > 0.1$ for all pairwise comparisons in χ^2 tests). Second, for each task with $\lambda > 0.75$, the share of beliefs in line with nuisance neglect (again, defined as being within $[0.95, 1.05]$), is significantly higher than in all tasks with $\lambda < 0.75$ (all $p < 0.05$, pairwise χ^2 tests). Third, for each task with $\lambda < 0.25$, the share of beliefs in line with signal neglect (defined as being within $[-0.05, 0.05]$), is significantly higher than in all tasks with $\lambda > 0.25$ (all $p < 0.01$, pairwise χ^2 tests). All this means is that, in line with the prediction, the share of nuisance neglect increases with increasing signal-to-noise ratio λ.

[32]For each task with $\lambda > 0.75$, the share of beliefs in line with nuisance neglect (defined as $\hat{\lambda} \in [0.95, 1.05]$), is significantly higher than in all tasks with $\lambda < 0.75$ ($p < 0.05$ in all pairwise χ^2 tests).

Note, however, that as the signal-to-noise ratio decreases, subjects are also more likely to observe signal realizations further away from 100, which is the mean of the normally distributed X. This means the subjective likelihood of a signal $P(S_i = s)$ decreases. In additional analyses reported in Appendix E.5.1, I show that while nuisance neglect increases in λ, this effects becomes insignificant once I *control for the subjective likelihood of observed signal values under nuisance neglect*. That means, in two tasks with identically distributed X and an identical observed information value s, there is no statistically significant difference in the propensity to commit nuisance neglect.

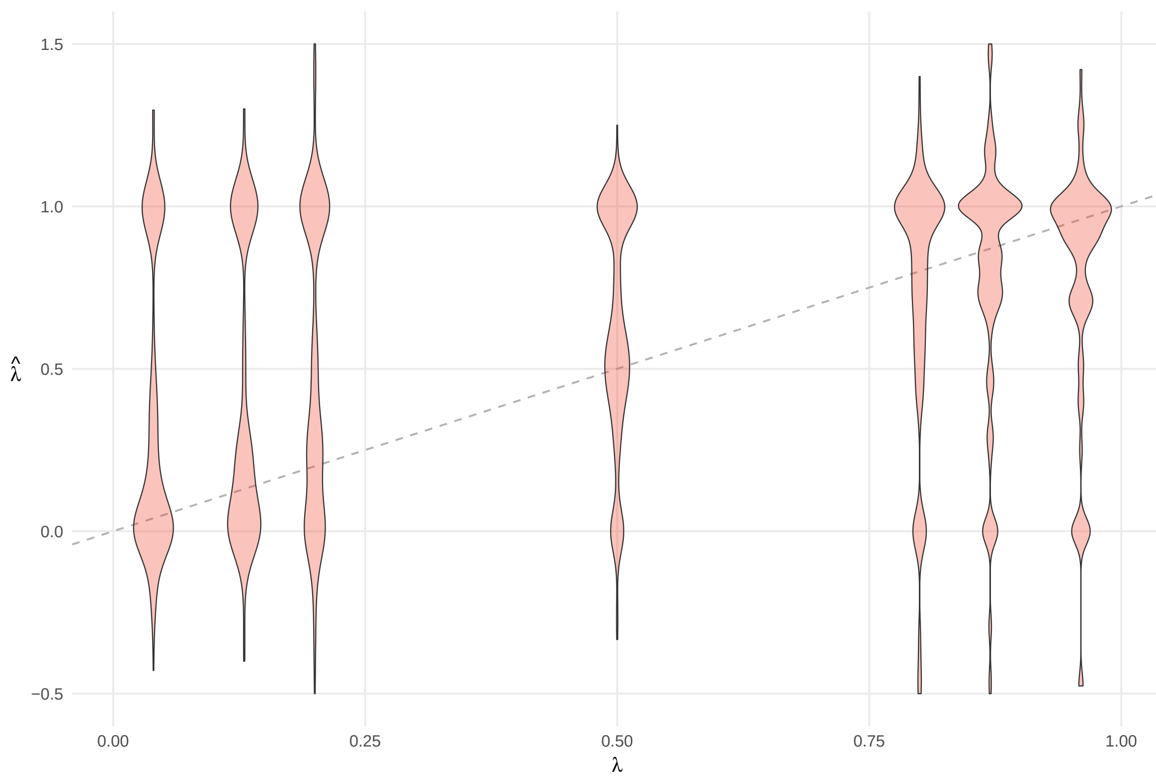


Figure 18: Kernel density estimates for seven different tasks (see Table 12) in an online experiment testing the effect of the signal-to-noise ratio on the prevalence of different updating modes. The horizontal axis indicates $\lambda = \frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}$ of the task. The vertical axis shows the empirical equivalent derived from subjects guesses as $\hat{\lambda} = \frac{m-100}{s-100}$. Note that $\hat{\lambda} = 0$ indicates signal neglect, $\hat{\lambda} = 1$ indicates nuisance neglect, and the dashed line indicates the $\hat{\lambda}$ that corresponds to Bayesian updating. Based on N=207 in each task. Epanechnikov kernel with bandwidth 0.07.

E.5.2 Treatment *Directional Bias*

The second treatment, *Directional Bias*, tests the effect of directional bias in the information structure while holding constant the signal-to-noise ratio. For an information structure $S = X + Y$, the expected bias of beliefs that neglect Y by assuming $S = X$ increases in $|\mu_S - \mu_X|$.

I conduct an additional online experiment ($N = 112$) using the five task configurations displayed in Appendix Table 13. This time, only μ_Y varies across tasks while all other parameters of the information structure are fixed at $\mu_X = 100$, $\sigma_X^2 = \sigma_Y^2 = 100$ and $S = X + Y$.[33] The mean value of the signal, μ_S, varies with μ_Y. To optimally learn from S, subjects need to account for the fact that observed values of S were on average higher or lower than the corresponding draws of X whenever $\mu_X \neq \mu_S$, i.e., they are directionally biased. This also sets this experiment apart from the baseline study, where S is an unbiased estimator of the mean of X.

Raw beliefs for each task are plotted in Appendix Figure 19. I report two main findings. First, in non-parametric tests we find that subjects are less likely to commit nuisance neglect as the directional bias of the signal increases, i.e., the greater the absolute value of μ_Y ($p < 0.05$ in all pairwise χ^2 tests, see Appendix E.5.1 for details). Second, with increasing bias subjects are again more likely to observe signal realizations that lead to beliefs with low ex-post plausibility under nuisance neglect, $P(X = s)$. Signal realizations are more likely to make subjects aware of their nuisance neglect. In regression analyses reported in Appendix E.5.1, I again find that the decrease of nuisance neglect with increasing directional bias turns insignificant upon adding $P(X = s)$ as a control.

Taken together, the prevalence of nuisance neglect decreases as its expected cost increases in both experiments. However, an analysis that pools all signal realizations ignores the fact that information structures with higher cost generate more extreme signal realizations in this setting, and that more extreme signal realizations can act as attentional

[33] Note that $\lambda = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}$ is held constant across tasks, setting this experiment apart from treatment *Signal-to-noise ratio*.

cues in themselves. We find that conditional on the signal realization there is no effect of the expected cost of ignoring Y. This is consistent with a mental gaps explanation but not a friction explanation as defined here. Specifically, the prevalence of nuisance neglect is insensitive to its costliness across tasks of the same experiment.

Table 12: Online tasks: Experiment on signal-to-noise ratio

X	Y	S	$\lambda = \frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}$
$\mathcal{N}(100,25)$	$\mathcal{N}(0,1600)$	$X+Y$	0.015
$\mathcal{N}(100,100)$	$\mathcal{N}(0,1600)$	$X+Y$	0.059
$\mathcal{N}(100,25)$	$\mathcal{N}(0,100)$	$X+Y$	0.25
$\mathcal{N}(100,100)$	$\mathcal{N}(0,100)$	$X+Y$	0.5
$\mathcal{N}(100,100)$	$\mathcal{N}(0,25)$	$X+Y$	0.75
$\mathcal{N}(100,1600)$	$\mathcal{N}(0,100)$	$X+Y$	0.941
$\mathcal{N}(100,1600)$	$\mathcal{N}(0,25)$	$X+Y$	0.985

Notes: This table provides an overview of the five tasks in the online experiment on the effect of the signal-to-noise ratio. Note that for all normally distributed variables, the support was discretized to integers, truncated at $\mu-50$ and $\mu+50$ and then the distributions were scaled such that the they have unit probability mass.

Table 13: Online tasks: Experiment on directional bias in information

X	Y	S	μ_S
$\mathcal{N}(100,100)$	$\mathcal{N}(0,100)$	$X+Y$	100
$\mathcal{N}(100,100)$	$\mathcal{N}(-25,100)$	$X+Y$	75
$\mathcal{N}(100,100)$	$\mathcal{N}(-50,100)$	$X+Y$	50
$\mathcal{N}(100,100)$	$\mathcal{N}(25,100)$	$X+Y$	125
$\mathcal{N}(100,100)$	$\mathcal{N}(50,100)$	$X+Y$	150

Notes: This table provides an overview of the five tasks in the online experiment on the effect of the directional bias in signals. Note that for all normally distributed variables, the support was discretized to integers, truncated at $\mu-50$ and $\mu+50$ and then the distributions were scaled such that the they have unit probability mass.

Raw beliefs for each task are plotted in Appendix figure 19. The figure indicates that subjects were less likely to commit nuisance neglect as the directional bias of the signal increased, i.e., the greater the distance of μ_Y from 0.

This observation is supported by non-parametric tests. The share of beliefs in line with

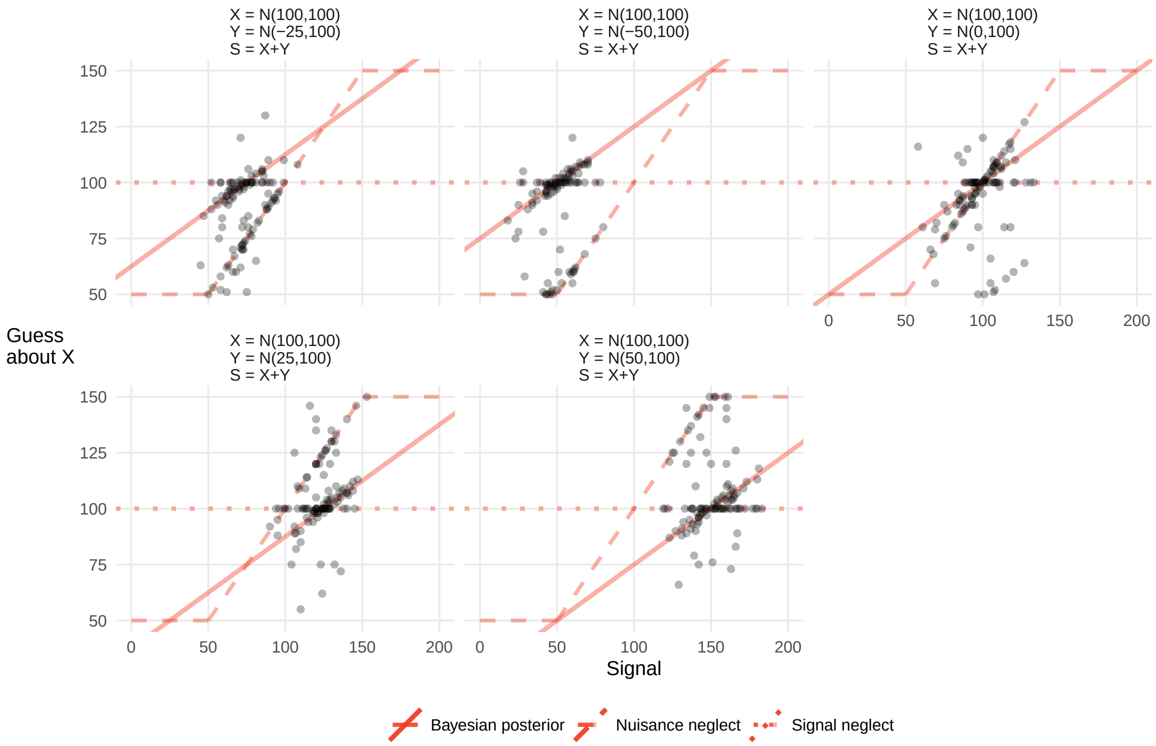


Figure 19: Beliefs in baseline tasks of online experiments. N=112 in each task. Each dot corresponds to one stated belief. The three red line indicate the Bayesian benchmark, nuisance neglect, and information neglect.

nuisance neglect significantly decreased as μ_S moved away from 100.[34] Notably, I found that this decrease went hand in hand with an increase in Bayesian beliefs, rather than signal neglect.[35]

Similar to the case of treatment *Signal-to-noise ratio*, however, as directional bias in a signal structure increases, subjects are more likely to observe signal values with low subjective likelihood under nuisance neglect. This means the plausibility of observed signals

[34]That is, the share of nuisance neglect decreased in both directions away from 100 for adjacent tasks, e.g., both for $\mu_S = 100$ vs. $\mu_S = 75$ and $\mu_S = 75$ vs. $\mu_S = 50$. Nuisance neglect can be defined in different ways. I either define it as any guess falling within a margin of 5 units around the hypothetical nuisance neglect guess, or based on $d_{FN}^{rel} = \frac{d_{FN}}{d_{FN}+d_B+d_{SN}}$ falling within 0.05 to either side of 0, where d is the distance of a stated belief m to the respective benchmark belief for each of three updating modes. That means, e.g., $d_B = \left|m - m_B\right|$ is the distance to the Bayesian belief. Hence, d_{FN}^{rel} is the distance of a belief to a hypothetical belief under nuisance neglect, *relative* to the summed distances of the elicited belief to all three updating modes. $p < 0.05$ in all pairwise χ^2 tests.

[35]The share of Bayesian beliefs as defined above significantly increases with the distance of μ_S from 100, $p < 0.1$ in all pairwise χ^2 tests.

decreases under the subjective signal structure associated with the default representation, and hence a signal is more likely to be a cue. In the regression analyses reported in Appendix E.5.1, I again find that the decrease of nuisance neglect with increasing directional bias vanishes *after controlling for the subjective likelihood (assuming nuisance neglect) of observed signal values*.

Table 14: Directional bias

Dependent variable:	Rel. distance from nuisance neglect	1 if rel. distance from nuisance neglect < 0.1
	(1)	(2)
$Y \sim \mathcal{N}(-50, 100)$	0.088**	-0.128**
	(0.044)	(0.061)
$Y \sim \mathcal{N}(-25, 100)$	0.142**	-0.176**
	(0.058)	(0.084)
$Y \sim \mathcal{N}(25, 100)$	0.154***	-0.211***
	(0.043)	(0.064)
$Y \sim \mathcal{N}(50, 100)$	0.169***	-0.238***
	(0.056)	(0.084)
Absolute difference of signal from mean	0.005***	-0.006***
	(0.001)	(0.002)
Constant	0.262***	0.744***
	(0.031)	(0.051)
R^2	0.15	0.12
# Observations	548	549

Notes: The dependent variables are computed based on $\frac{d_{NN}}{d_{NN}+d_B+d_{IN}}$ where $d_{\cdot}$ is the distance of a stated belief m to the respective benchmark belief for each of three updating modes, e.g., $d_B = \left|m - m_B\right|$ is the difference to the Bayesian belief. Hence the dependent variable in (1) is the distance of a belief to a hypothetical belief under nuisance neglect, *relative* to the summed distances of the elicited belief to all three updating modes. The dependent variable in (2) is a dummy for whether this relative distance is smaller than 0.1, such that a belief is plausibly classified as nuisance neglect. In all tasks, $X \sim \mathcal{N}(100, 100)$ and $I = X + Y$. OLS regressions. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Taken together, the prevalence of nuisance neglect decreases as its expected cost increases in both experiments. However, an analysis that pools all signal realizations ignores the fact that information structures with higher cost generate more extreme signal realizations in this setting, and that more extreme signal realizations can act as attentional cues in themselves. We find that conditional on the signal realization there is no effect of the expected cost of ignoring Y. This is consistent with a mental gaps explanation but not a friction explanation as defined here. Specifically, the prevalence of nuisance neglect is insensitive to its costliness across tasks of the same experiment.

F On Complexity and the Persistence of Errors

F.1 Overview

Evidence for systematic errors in seemingly simple tasks may be surprising given people's ability to learn from feedback. In practice, people frequently receive performance feedback and have ample opportunities to learn. Yet, the persistence of errors even in the presence of performance feedback is widely documented (Gigerenzer, 1991; Stanovich and West, 2000). At the same time, our understanding of its sources remains limited.

The findings on the behavioral mechanism underlying nuisance neglect point to a potential explanation: when a solution strategy comprises several elements, people may fail to learn about the actual source of an error even in the presence of feedback. This motivates the following investigation of the origins of persistence of errors in the case of nuisance neglect.

I highlight three implications of the previous results for learning. First, we find that interventions that specifically direct attention to the neglected part of the problem reduce the bias (e.g., treatments *Hint* and *Enforced Deliberation*), but experimental manipulations that unspecifically increase the overall motivation do not (e.g., *High Stakes*). Simple performance feedback is typically unspecific: people are either informed that they made a mistake and/or they learn the optimal action. This way, unspecific performance feedback leaves open what exactly led to a suboptimal response.

Second, a person is less likely to identify the source of an error if it is an element of the solution strategy that they do not deliberately execute. The previous evidence from, e.g., confidence statements, suggests that people neglect nuisance variables without being aware of the neglect.

Third, a more complex solution strategy with many steps may divert attention from the specific source of error. The propensity to learn from unspecific feedback may thus be reduced by more computational steps required in the solution process.

Prediction 2. *Learning from feedback is limited by subjects' failure to identify the specific* *source of their error.*

(a) *Nuisance neglect is not eliminated by performance feedback.*

(b) *An increase in the computational complexity of the updating problem reduces learning from feedback because subjects (falsely) attribute mistakes to deliberately executed computations.*

F.2 Design

In three additional laboratory experiments I test these predictions about persistence and learning.

Treatment *Feedback*. This treatment is identical to the baseline treatment *Narrow*, except that it provides standard performance feedback: in each of the five baseline tasks, after guessing X, subjects learn the true value of X.

Treatment *Computation with Feedback*. To manipulate the computational complexity of the updating task, condition *Computation with Feedback* adds a simple algebraic computation to the signal structure. These computations are extremely simple, e.g., "+20 − 30". Identical computations used in treatment *Computation* show that they do not affect stated beliefs. This treatment tests the prediction that while added computations do not affect stated beliefs, they affect learning form feedback.[36] Specifically, the computation provides an obvious source of error, so that subjects who commit nuisance neglect may misattribute an error to the computation and fail to notice their neglect of Y.

Treatment *Computational Feedback*. A concern in comparing behavior in the previous two treatments is that added computations increase the complexity of both the updating stage and the feedback stage. A reduction of learning might result from the latter, whereas the hypothesis concerns the former. To address this issue, a third treatment holds the overall complexity of the combined updating and feedback problem constant.

[36]*Computation with Feedback* is identical to condition *Computation*, except for the feedback; and it is identical to condition *Feedback*, except for the additional simple computation.

In *Computational Feedback*, subjects receive a signal on X and Y without additional computations, as in *Feedback*. However, a computation is added at the feedback stage. Instead of observing the true value of X, subjects see a different value, e.g., $X + 20 - 30$, and need to undo the simple computation to learn the true value. Note that the added computation enters after stating a guess about X and thus cannot possibly affect inference.

This treatment tests the prediction that upon learning about a mistake after deciphering the true X, subjects cannot attribute this updating error to the computation, which occurred *after* stating the guess. This increases the likelihood of reflecting on other sources of error including the role of Y. If learning in this treatment is similar to *Feedback* but less that in *Computation with Feedback*, one may conclude that learning is impaired by computational complexity that is specific to the updating task.

F.3 Results

In the first round – before receiving feedback for the first time –, nuisance neglect as measured by the inattention score θ (equation (6)) is indistinguishable across the feedback treatments (Appendix Figure 21), as should be the case.[37] The following analyses show data from the last round only, since learning effects should be highest after several rounds of feedback.[38] Figure 20 displays mean inattention by treatment condition. All statistical analyses, however, are based on empirical distributions of inattention.[39] Figure 20 shows the three feedback conditions (below the dashed line) alongside three no-feedback conditions (above the dashed line, discussed in Section 3.1.4) for comparison. I report three main findings. First, the provision of feedback alone (condition *Feedback*) significantly decreases inattention relative to treatment *Narrow* ($p < 0.001$, M-W U test). At the same time, learning is far from perfect, as indicated by a remaining treatment effect between *Feedback* and the *Broad* treatment without feedback ($p < 0.001$).

[37]This validates the previous result that the computation in itself does not affect stated beliefs.

[38]I obtain similar findings when pooling beliefs from round two to round five. See further results in Appendix F.

[39]See Appendix F for distribution plots.

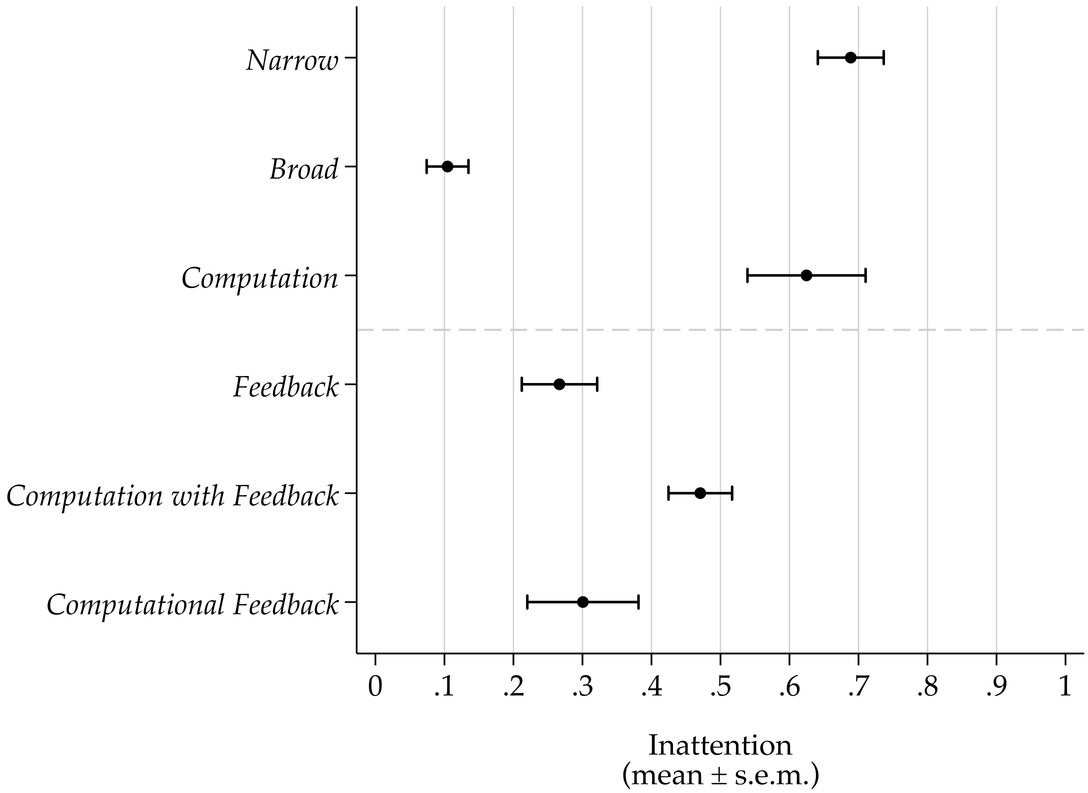


Figure 20: Treatment means of inattention to Y. Inattention scores are calculated following Section 3.1.5. The three treatments above the dashed line show conditions without feedback for reference (see Section 3). Subjects receive feedback is about the actually drawn value of X. Displayed are implied inattention scores in the final baseline round, after having received feedback in the four preceding rounds. Sample sizes are N = 72 in both *Narrow* and *Broad*, N = 48 in *Feedback*, and N = 24 in each of *Computation*, *Computation with Feedback* and *Computational Feedback*.

Second, additional computations in the inference process impede learning. Inattention in *Feedback* and *Computation with Feedback* differs significantly ($p = 0.008$). Added computations presumably reduce subjects' propensity to notice that there are parts of the problem that they have failed to attend to, i.e., the role of Y. Notably, the added computation virtually eliminates learning. Inattention in *Computation with Feedback* is not significantly lower than in *Computation* ($p = 0.21$).

Third, the elimination of all learning effects in *Computation with Feedback* is not driven by the increase in overall complexity relative to *Feedback*. In fact, inattention in *Computational Feedback* is indistinguishable from *Feedback* ($p = 0.48$), but significantly differs from *Computation with Feedback* ($p = 0.005$).

Taking stock, the data from three feedback experiments support the prediction that attribution errors persist because subjects fail to identify the specific source of their updating error, which is exacerbated by higher computational complexity.

In an attempt to gather more direct (correlational) evidence for this hypothesis, the experiments include an additional choice in all feedback treatments. On the feedback screen that informed about the actual draw, subjects could choose to be reminded of up to exactly one piece of the preceding belief task: the distribution of X, the distribution of Y, or the signal structure. Revealing such details can help subjects figure out the source of their erroneous guess. In the first round, i.e., upon receiving feedback for the first time, subjects are indeed more likely to reveal the distribution of Y in *Feedback* than in *Computation with Feedback* ($p = 0.044$), indicating that they are more likely to notice a potential role of Y in the absence of additional computations. This effect, however, is not robust and loses significance when pooling all rounds. Procedural details and further results are relegated to Appendix F.[40]

Result 3.

(a) Performance feedback decreases, but does not eliminate nuisance neglect.

(b) An increase in computational complexity eliminates learning from performance feedback.

In the first baseline round, i.e., before receiving feedback for the first time, inattention scores do not significantly differ between the four learning treatments, as expected.

[40]Finally, note that performance feedback in practice is often noisy. The possibility that observed feedback is imprecise provides another obvious way for subjects to resolve the conflict between their subjective belief and surprising feedback, again reducing learning about their neglect. In condition *Imperfect Feedback*, subjects receive feedback about the true X that is correct only with 80% probability, but see a value of X which is not true with 20% probability. Again, learning is reduced to a similar extent as by adding the computation, see Appendix F.7.

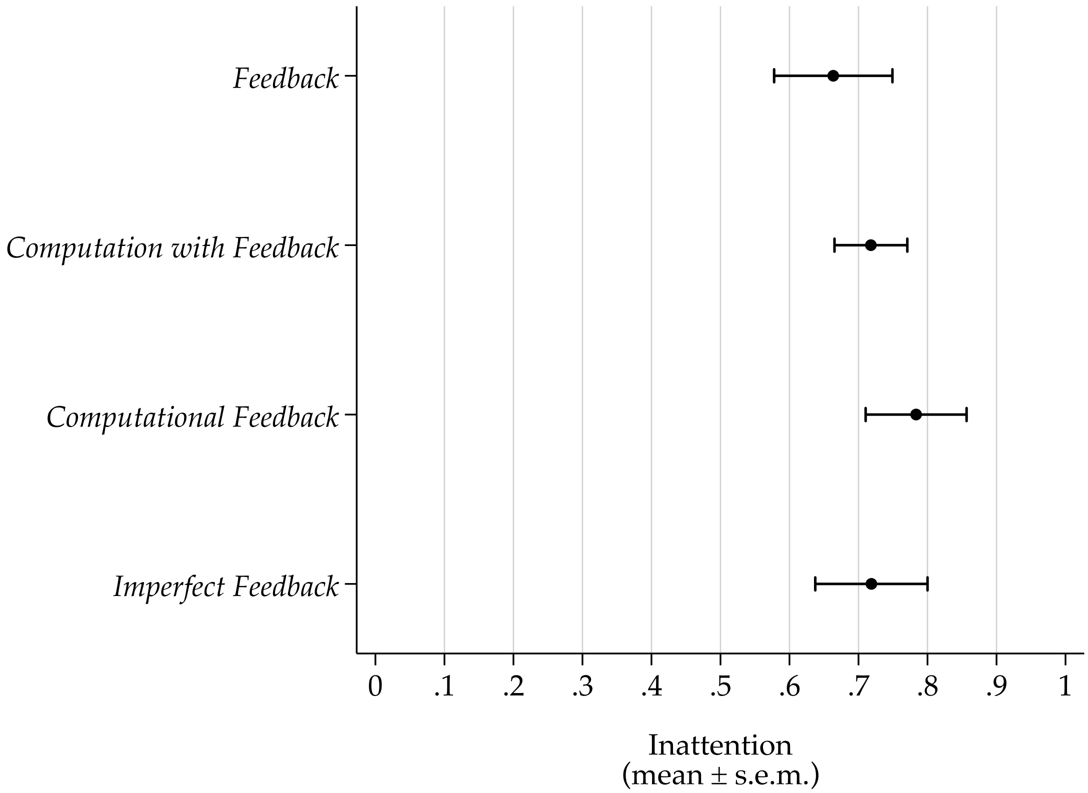


Figure 21: Treatment means of inattention to Y in the first round. Displayed are implied inattention scores in the initial baseline round. Subjects have not previously received feedback when stating these guesses. Sample sizes are N = 48 in both *Feedback* N = 24 each in all other three conditions.

F.4 Feedback

From the initial experiments we know that the neglect of Y is typically confident and occurs outside subjects' awareness. The key hypothesis motivating the feedback treatments is that people fail to reflect on steps of their solution strategy that are not available to introspection or recall, interfering with learning even in the presence of surprising feedback. Condition *Feedback* is akin to *Narrow*, but also shows the actually drawn number of X after guessing it. Relative to the no-feedback benchmark (condition *Narrow*), there is marginally significant learning after receiving feedback for the first time (p=0.06, in a regression of inattention in the second round on a treatment dummy and including task-fixed effects). After having received feedback four times, mean inattention is .27 as compared to .69 in the no-feedback baseline. Despite this sizable improvement, inattention is still significantly greater than in the fifth round of the no-feedback setting with Broad incentives (mean inattention 0.10, p=0.00). Figure 22 shows a histogram of inattention parameters, and Figure 23 histograms of the raw beliefs in condition *Feedback*.

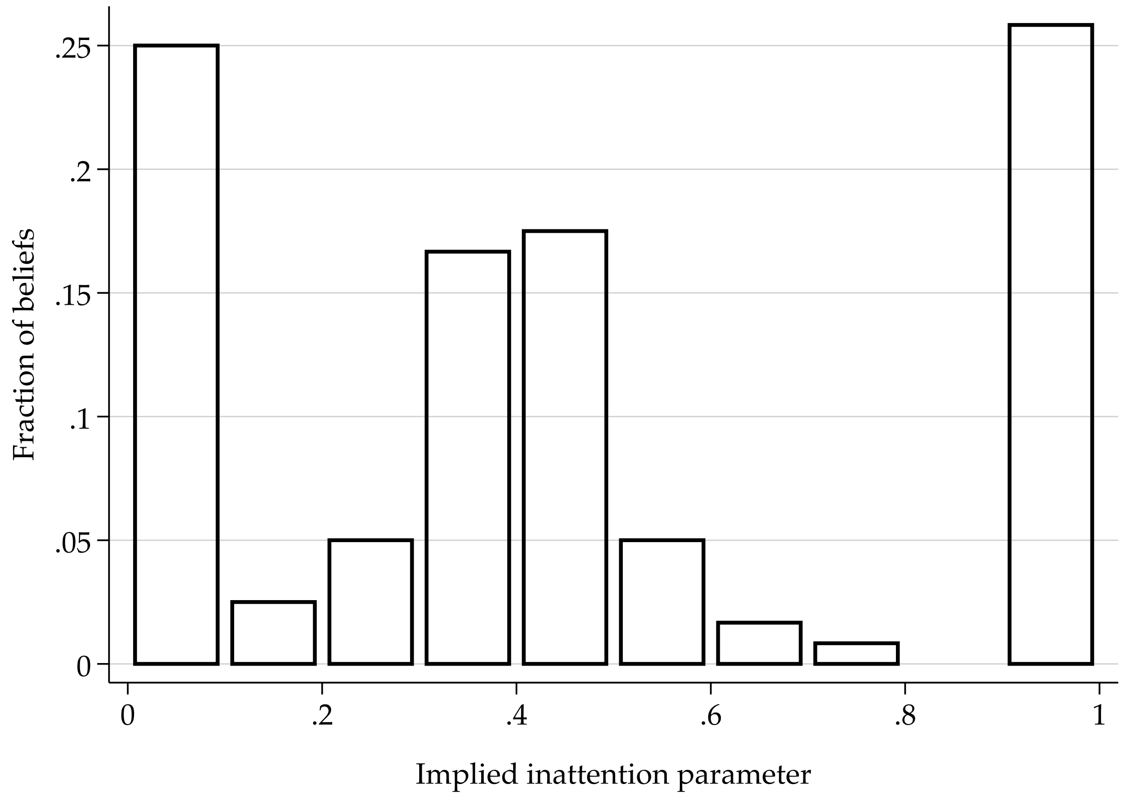


Figure 22: Histogram of implied inattention to Y in condition *Feedback*. Based on 216 stated beliefs. A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention.

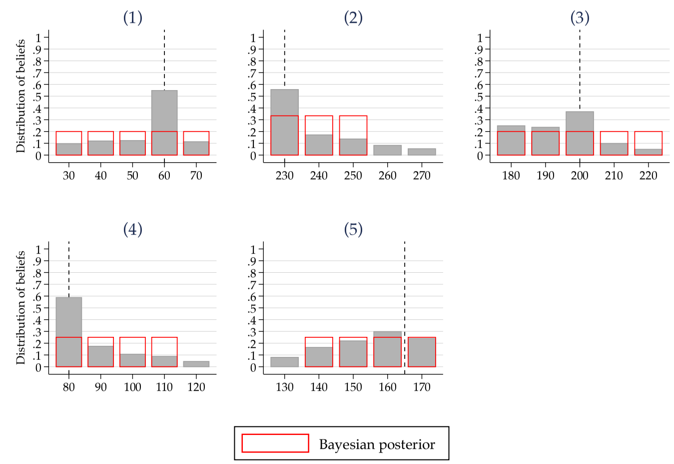


Figure 23: Distribution of elicited belief distributions about X in each one of five baseline tasks of condition *Feedback*. N=24 for each condition in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the vertical dashed line. Task order was randomized.

F.5 Computation with Feedback

To directly test the hypothesis that people fail to reflect on the non-accessible elements of their solution strategy, *Computation with Feedback* provides feedback that is identical to *Feedback*, but the initial signal about X and Y is modified by a simple algebraic computation. This condition is identical to the anchoring treatment *Computation*, but including the feedback stage. As found in the *Computation* condition and confirmed here, the additional computation is inconsequential for the guesses about X that subjects submit (see also Figure 21). Virtually every subject correctly accounts for the computation but then tends to forget about Y. Presented with surprising feedback about the actually drawn number, however, subjects might now first remember the conscious part of their inference strategy, i.e., undoing the calculations. The computations provide them with "a place to hang their coat" in the sense of an obvious – albeit unlikely – source of error. This is what we find: Adding the computation virtually eliminates learning. Figure 24 shows a histogram of inattention parameters, and Figure 25 histograms of the raw beliefs in condition *Computation with Feedback*.

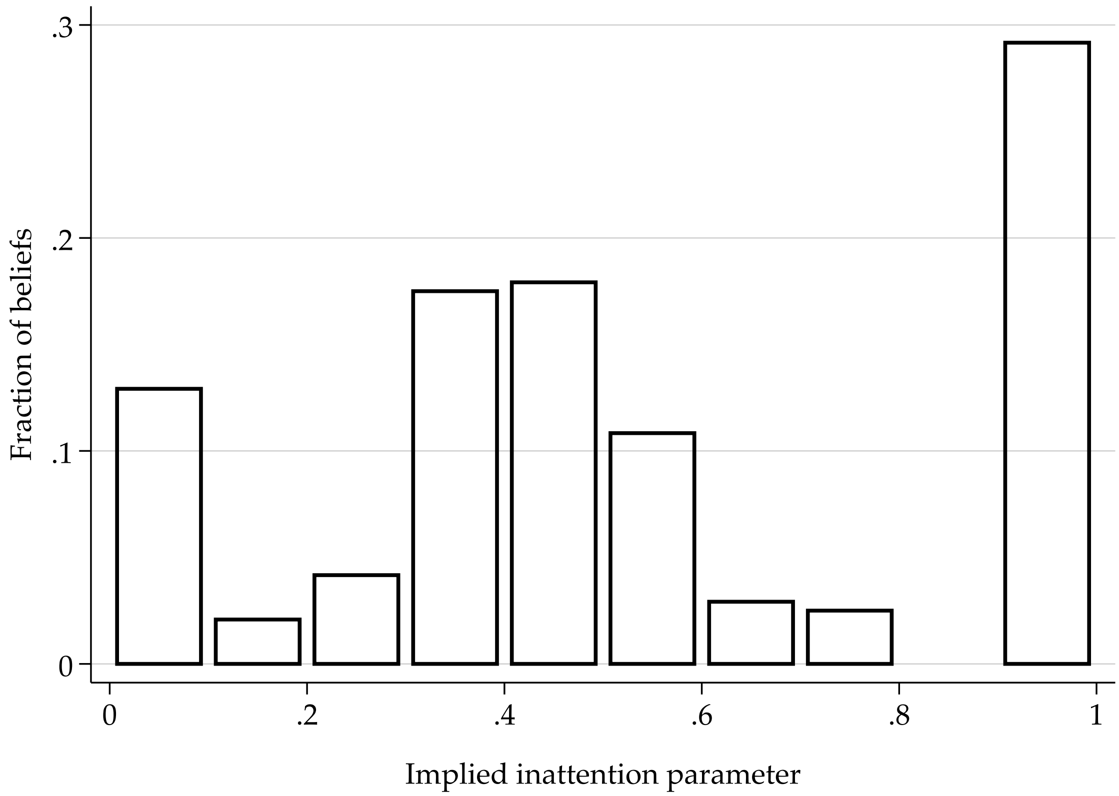


Figure 24: Histogram of implied inattention to Y in condition *Computation with Feedback*. Based on 216 stated beliefs. A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention.

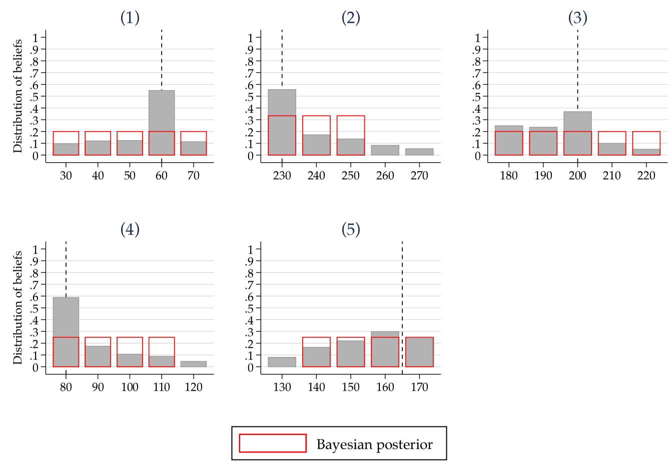


Figure 25: Distribution of elicited belief distributions about X in each one of five baseline tasks of condition *Computation with Feedback*. N=24 for each condition in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the vertical dashed line. Task order was randomized.

F.6 Computational Feedback

Reduced learning when algebra is added could result from increased complexity. In condition *Computation Feedback*, therefore, subjects have narrow incentives and receive a signal on X and Y without additional computations, i.e., the mean or sum as before. This time however, the same computations as in *Computation with Feedback* are added at the feedback stage. That means, instead of seeing the true value of X, subjects see a different value on which they first perform the computations and then arrive at the true value of X. The results suggest it is not computational complexity of a problem per se that reduces learning form feedback. Instead, it is precisely the consciously accessible steps of reasoning performed *when doing inference* that interfere with reflecting on the role of Y. Figure 26 shows a histogram of inattention parameters, and Figure 27 histograms of the raw beliefs in condition *Computational Feedback*.

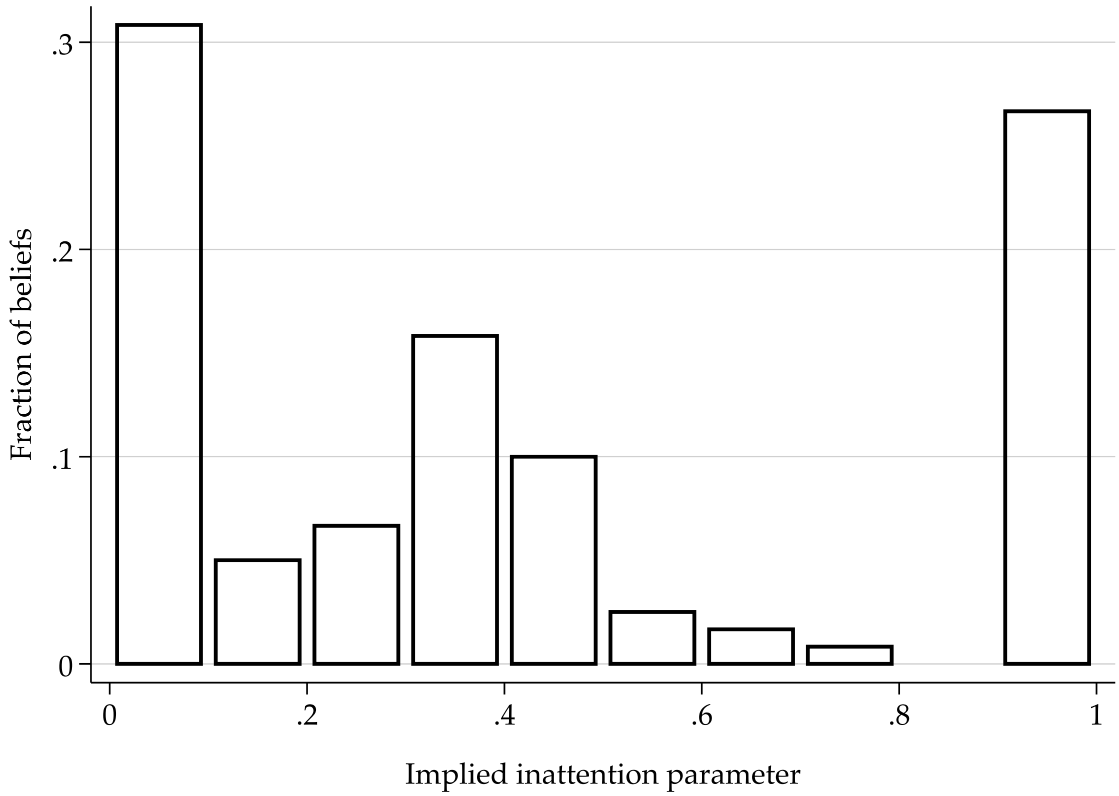


Figure 26: Histogram of implied inattention to Y in condition *Computational Feedback*. Based on 216 stated beliefs. A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention.

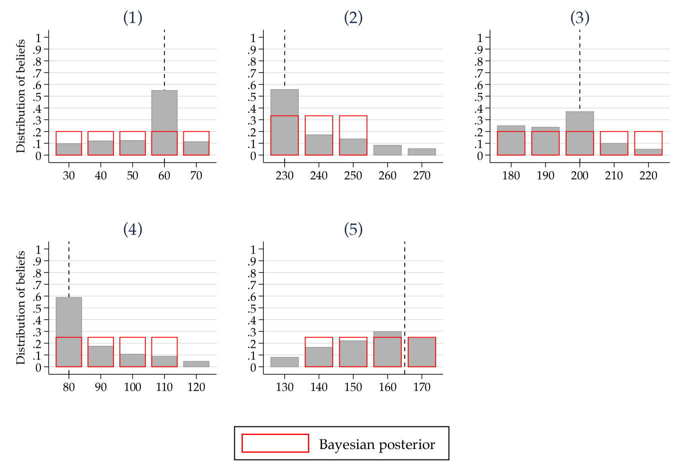


Figure 27: Distribution of elicited belief distributions about X in each one of five baseline tasks of condition *Computational Feedback*. N=24 for each condition in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the vertical dashed line. Task order was randomized.

F.7 Imperfect Feedback

Learning in practice is often based on imprecise signals. The possibility that observed feedback is not exactly right might provide another obvious way for subjects to explain a conflict between their stated belief and received feedback, reducing learning. In an additional treatment, feedback about the true X was only correct with 80% probability, and the remaining 20% subjects would see a value of X which is not the true one. Pooling beliefs following the first four rounds of feedback, there is only a small and marginally significant positive effect of receiving this feedback on inattention relative to receiving no feedback at all ($p = 0.09$). As predicted, simple solutions for why beliefs conflict with the feedback compromise the ability to reflect on the role of Y. Figure 28 shows a histogram of inattention parameters, and Figure 29 histograms of the raw beliefs in condition *Imperfect Feedback*.

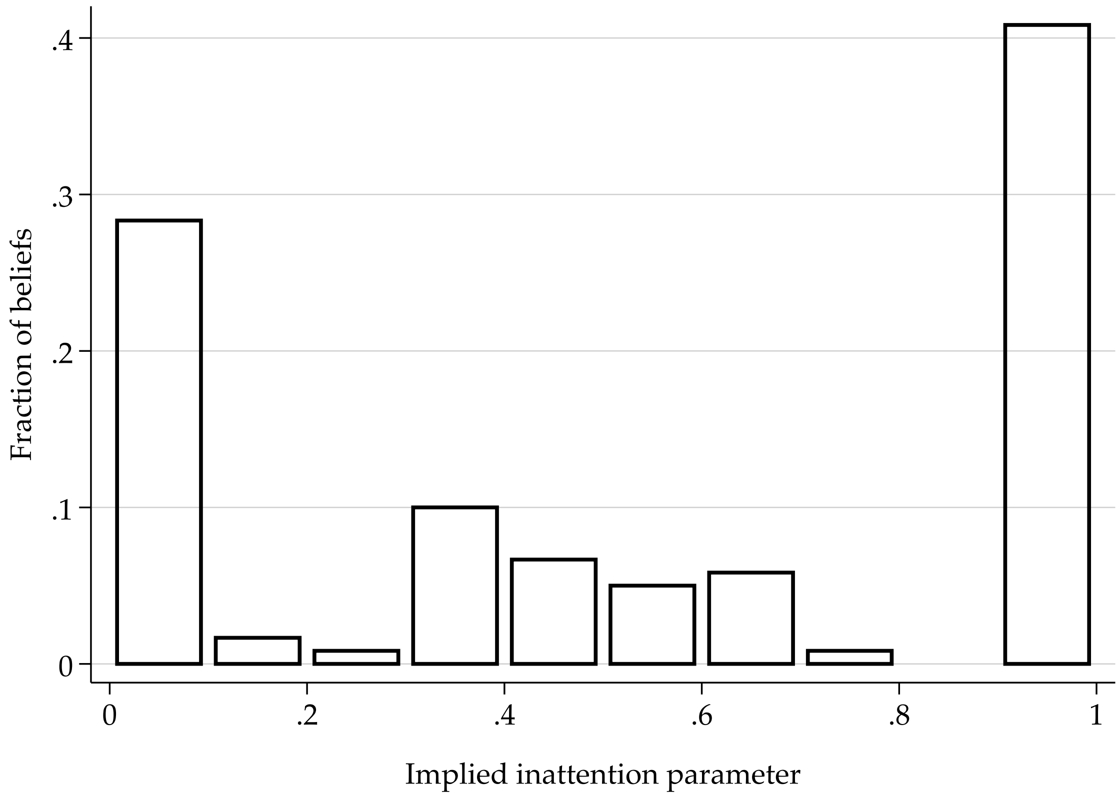


Figure 28: Histogram of implied inattention to Y in condition *Imperfect Feedback*. Based on 216 stated beliefs. A parameter of $\theta = 0$ is consistent with Bayesian updating. $\theta = 1$ means complete inattention.

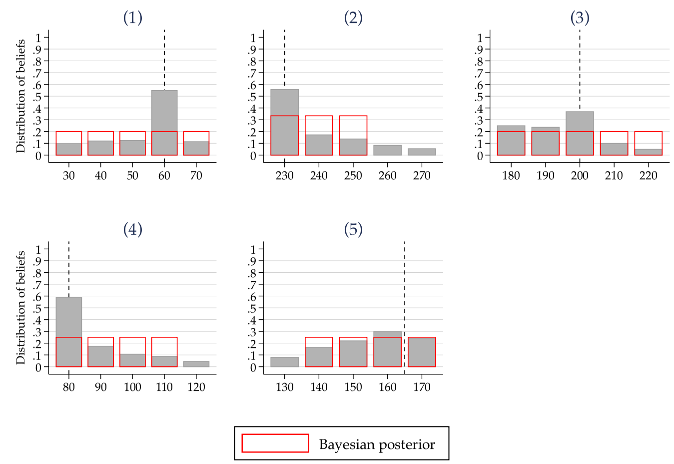


Figure 29: Distribution of elicited belief distributions about X in each one of five baseline tasks of condition *Imperfect Feedback*. N=24 for each condition in each task. The horizontal axis shows possible outcomes of X. The Bayesian posterior belief is provided for reference. The observed signal is indicated by the vertical dashed line. Task order was randomized.

G Experimental Instructions

G.1 Main Instructions in *Narrow* and *Broad*

All instructions were computerized. Translated from German into English.

Welcome. For your participation you will receive a fixed payment of 5.00 € , which will be paid to you in cash at the end. In this study you will take decisions on the computer. Depending on how you decide you can earn additional money. **During the study it is not allowed to communicate with other participants. Note also that the curtain of your cubicle must be closed during the entire study.** Please turn off your mobile phone now, so that other participants will not be disturbed. Please only use the designated functions on the computer and make your entries using the keyboard and the mouse. If you have questions, please make a hand signal. Your question will be answered at your seat. To proceed click "Next".

Your Task

You will successively receive 9 different guessing tasks. The guessing tasks are about guessing numbers that are randomly drawn. The better your guess, the more money you can earn. In each guessing task there is a random number X. The computer randomly picks X from a range of possible numbers. You will receive an encrypted hint about which number was actually drawn, and you can then indicate your guess about X. There are 9 rounds in total. In each round you receive a new guessing task. That means, in each round the computer again determines a number X independently of the other rounds. Your payoff depends on how precisely you guess, that means how accurate your guess is. At the end of the study, one of the 9 rounds is picked at random and you will be paid according to the precision of your guess in that round.

The Guessing Tasks

Example. Imagine there are exactly 3 balls. These 3 balls have the following numbers on them: 10, 20, 30. In this example, the number X is determined as follows: The computer randomly draws one of these three balls. Each ball is drawn with equal probability. It is equally likely that the "10" will be drawn, that the "20" will be drawn, or that the "30" will be drawn. The number X is then the number of the ball that was randomly drawn by the computer. However, you will not be told which number X was drawn. Instead you receive an additional hint. You can look at this hint, before you guess the number X. Please note:

- For each guessing you will be informed about which numbers can be drawn. In different guessing tasks, different numbers can be drawn. Sometimes the numbers repeatedly occur across rounds. However, the draws in these rounds are completely independent of one another.
- The additional hint can give you different types of information in different rounds. In each round you will learn anew, what the additional hint means. Therefore you should pay attention in every new guessing task to which information the hint indicates.

Your guess. You can state your guess by allocating 100 percentage points to the different numbers. The *more certain* you are, that a particular number was drawn, the *more* points you should allocated to this numbers. Similarly, the more certain you are, that a particular was *not* drawn, the *fewer* points you should allocate to this numbers. The sum of your allocated points must be exactly 100. In the example above, if after receiving the additional note you are, for example, sure that X = 30, then you should allocate 100 points to the number 30, and 0 points to both the numbers 10 and 20. In the example above, if after receiving the additional note you are, for example, sure that X = 20, then you should allocate 100 points to the number 20, and 0 points to both the numbers 10 and 30. In the example above, if after receiving the additional note you think, for example, that the number 30 have definitely not been drawn, but the 10 and 20 have been drawn with equal probability, then you should allocate 50 points each to the number 10 and 20, and 0 points to the number 30. You can arbitrarily allocate the points. However you can only allocate full points, that means for example that you cannot allocate half points. For instance, you could allocate 21 points to number 30, 47 points to number 20, and 32 points to number 10. **The more points you assign to the number that was actually drawn, the more money you can earn. Similarly, the fewer points you allocate to those numbers, that are not equal to X, the more money you can earn.** The calculation of your payoff will be explained in greater detail in the following section.

Your payment

In addition to your show up fee you will be paid based on how precisely you guessed. To this end one of the 9 rounds will randomly be picked and you will be paid according to the precision of your guess in that round. This means for you that *each one* of your guesses is potentially relevant for your payment and accordingly you should carefully think through every guess. You can either earn and additional 10 € or 0 € from your guessing task. While the following explanation might look difficult, the basic principle is very simple: **the better your guess, that is the more percentage points your guess assigns to the actually drawn number and the fewer percentage points it allocates to every wrong number, the more likely it is that you receive the 10€ .** Concretely this means the following: In expectation you will earn most money if you allocate your points according to how probable you find it that the respective numbers was drawn (with 1 point = 1 percent). If you have understood this, it is not necessary for the maximization of your earnings to read the following section on the details of the calculation of your additional payment. You can then directly click on "Next."

For your information: Details on the calculation of your additional earnings. After you have stated your guess, the computer will randomly draw another number kj This number is between 0 and 20,000. (More precisely, this numbers is drawn from a discrete uniform distribution on the interval from 0 to 20,000.) You will then receive the 10 € if the sum S is smaller or equal to k. S is the sum of the following elements:

- The squared deviation between the number of points that you allocated to the actually drawn numbers X, and 100 points.
- For *each* possible number, that has not been drawn (i.e., every other number than X): The squared deviation between 0 points and the number of points that you allocated to this numbers.

An exact mathematical formula of the sum S is displayed in the footnote.[41] If the sum S is bigger than k you will receive 0 € . Accordingly, the payoff rule is as follows:

Payment = 10.00 € , if $S \leq k$

Payment = 0.00 € , if $S > k$

This means the following: If the sum of the squared deviations exceeds a particular value k, you will receive 0 € . If, however, the sum of the squared deviations is smaller than k, you will receive 10 € in addition. You can notice here that it should be your goal a) to keep the difference between the points allocated to X and 100 points as low as possible, that is to allocate as many points as possible to X, and b) to allocate as few points as possible to ever other number than X. An *example*: Let us assume that the computer has randomly drawn the number X = 30, while the numbers 10, 20 and 30 could have been drawn with equal probability. Also the number $k = 5,000$ For the following guesses you would receive the indicated payments.

[41] *Footnote text:* Exact mathematical formulation: There are N possible number from which X is drawn. In the example, $N = 3$. The number of points that you allocate to the ith of the N numbers is p_i. The indicator function $\mathbb{1}_i$ takes the value 1, if X is the ith number, and 0 otherwise. The sum S is calculated as follows: $S = \sum_{i=1}^{N}(\mathbb{1}_i - p_i)^2$. The expected payoff amount is maximized by indicating the probability distribution of the numbers after receiving the additional hint.

Your guess			Sum of squared deviations	Comparison to	Your payment
10	20	30	(X = 30)	k = 5,000	
100 Points	0 Points	0 Points	$100^2 + 0^2 + 100^2 = 20{,}000$	> k	0,00 €
0 Points	100 Points	0 Points	$0^2 + 100^2 + 100^2 = 20{,}000$	> k	0,00 €
0 Points	0 Points	100 Points	$0^2 + 0^2 + 0^2 = 0$	≤ k	10,00 €
10 Points	10 Points	80 Points	$10^2 + 10^2 + 20^2 = 600$	≤ k	10,00 €
25 Points	25 Points	50 Points	$25^2 + 25^2 + 50^2 = 3{,}750$	≤ k	10,00 €
0 Points	50 Points	50 Points	$0^2 + 50^2 + 50^2 = 5{,}000$	≤ k	10,00 €
33 Points	33 Points	34 Points	$33^2 + 33^2 + 66^2 = 6{,}534$	> k	0,00 €
45 Points	45 Points	10 Points	$45^2 + 45^2 + 90^2 = 12{,}150$	> k	0,00 €
90 Points	0 Points	10 Points	$90^2 + 0^2 + 90^2 = 18{,}100$	> k	0,00 €

In particular this means the following: If you allocate all 100 points to the right number X, you will received the 10 € **in any case**. However, you will also receive 10 € in many cases in which you allocate less than 100 points to X. The more points you allocate to the right number X, the more likely it is, that you receive the 10 € . **In expectation, you will earn the most money if you allocate the points according to how probable you think it is that the respective number was drawn.** Please note:

- It is not necessary, to allocate 100 points to the number that you think is most likely. As you can see in the examples of the table, you can also win 10 € if you have allocated less than 100 points to the right number X. Your earnings depend on the randomly drawn number *k*.

- Your guess in *one* randomly picked round will be paid. The guessing task that is payoff relevant for you is determined by the computer at the end of the study. Therefore you should indicate your best guess in each guessing task, independent of all other guessing tasks.

Summary

In each round it is your task to state a guess about the number that was randomly drawn by the computer. Before this, you will get a computer-generated, **encrypted hint**. For each guessing task you will see this additional hint and you can subsequently indicate your guess. Which hint you will receive, and how this hint is encrypted will be explained in the following. For the deciphering of the hint and your subsequent guess there is a time limit. You will previously be informed about how much time you have. The remaining time will be displayed while working on the tasks.

Encryption of Hints

You receive additional hints that have been encrypted by an **encryption device**. The encryption device transforms each hint (a number) into a letter code. You first need to decrypt the letter code back into a number in order to use the hint.

Decryption of the additional hint. When you get an encrypted sequence of letters as hint, you can decipher this hint by following these steps:

a. Transform the sequence of letters into a number using the code table.

1	2	3	4	5	6	7	8	9	0
A	B	C	D	E	F	G	H	I	J

b. Add 20 to the number

Before every guessing task you will receive an encrypted hint that you can decipher before stating your own guess. Whenever you receive a hint, you will see the code table as well as the decryption instructions. **That means you don not have to remember the decryption procedure.** You will soon get the opportunity to practice the decryption on an example hint.

Control Questions

Please notify one of the experiments now if you have questions about the instructions so far. If there is something that is unclear to you, please re-read the respective information carefully. You can return to the previous pages by clicking “Back”. If you click on “Proceed to control questions”, you will receive several control questions, which ensure your understanding of the instructions. You will not get paid for the control questions. However, you have to correctly answer all control questions to proceed to the guessing tasks. After you have correctly answered all control questions, you will be presented with the first guessing task.

G.2 Control Questions

Control Question 1 of 9

What is your main task in this study?

- There are several number from which X can be drawn. I need to add these numbers up to a sum.
- I guess the drawn number X.

Control Question 2 of 9

The numbers from which X is drawn vary across rounds. Sometimes the numbers occur in different rounds. For example, it could be that in two different rounds, the number X is randomly drawn from the number “10”, “20”, and “30”. Please evaluate the following statement: “In both round, each of the 3 numbers is drawn with equal probability.”

- **Wrong.** If, for example, the “10” was drawn in the first round, it is more probable that “10” will not be drawn in the next round.
- **Correct.** Both rounds are completely independent. The draw in the first round has no influence on which number is drawn in the second round.

Control Question 3 of 9

In guessing X, how can you make most money?

- By allocating the points to the numbers as precisely as possible based on how certain I am, that the respective number is X.
- By varying my guess and allocating by instinct sometimes more points to high numbers and sometimes more points to low numbers.

Control Question 4 of 9

After you have read the description of the guessing task and received the additional hint, you think that the number "20" is the most likely drawn number among the numbers. However you are not certain that it is the "20". Assess the following statement: "To maximize my payoff I have to put all 100 points on the number "10"."

- **Correct.** It is only this way that I can earn the 10 euros.
- **Wrong.** While I should put more points on the "20" than on all other numbers, I should not putt all points on the "20" , because I am not certain. If for example i am 60% sure that X = 20, I should put exactly 60 points on the "20", and allocate the remaining 40 points to the other numbers. This way it is most probably that I earn the 10 euros.

Control Question 5 of 9

Which of your guess is payoff-relevant?

- Every guess is paid out.
- No guess is paid out.
- A randomly picked guess is paid out.

Control Question 6 of 9

Imagine the number X is drawn with equal probability from the following four numbers: 50, 60, 70, 80. You have no additional information. Please indicate how in this case you should allocate the 100 points to the four numbers such that you make winning the 10 euros as likely as possible. Start by picking a number in the selection box to the left and assign a number of percentage points in the input field to the right. Use further input rows if you want to assign percentage points to other numbers.

Control Question 7 of 9

As before the number X is drawn with equal probability from the following four numbers: 50, 60, 70, 80. Please imagine now that after deciphering the hint you are certain that the "70" was drawn. Please indicate how in this case you should allocate the 100 points to the four numbers such that you make winning the 10 euros as likely as possible. If you want to allocate 0 percentage points to a number then you do not have to enter this into an extra row, but you can simply skip this number (0 points will automatically be allocated).

Control Question 8 of 9

Imagine you receive the hint: AJ. Please decipher the hint and enter your result below.

a. Transform the sequence of letters into a number using the code table.

1	2	3	4	5	6	7	8	9	0
A	B	C	D	E	F	G	H	I	J

b. Add 20 to the number

The decrypted hint reads:

Control Question 9 of 9

Imagine now you receive the hint: ACJ. Please decipher the hint and enter your result below.

a. Transform the sequence of letters into a number using the code table.

1	2	3	4	5	6	7	8	9	0
A	B	C	D	E	F	G	H	I	J

b. Add 20 to the number

The decrypted hint reads:

G.3 Task Instructions

Next to X another number was drawn by the computer, Y. Whether a participant has to guess Y as well was randomly determined at the beginning of the study and has no impact on the size of possible earnings.

[*Treatment Narrow:* **To you applies the following**: You indicate a guess **only about X** and will be paid for your guess of X as described.]

[*Treatment Broad:* **To you applies the following**: You guess **both drawn numbers**, X and Y. One of the numbers will later be picked and you will be paid for your guess of this number as described.]

[**The following description varies by task**]
X was randomly drawn from the following 5 numbers between 80 and 120, where each number was equally likely: 80, 90, 100, 110, 120.
Y was randomly drawn from the following 7 numbers between -30 and 30, were each number was equally likely: -30, -20, -10, 0, 10, 20, 30.
X and Y were drawn independently.

[*Treatment Broad:* You will guess X and Y *simultaneously*, that means in each entry row

you have to pick both a number for X and a number for Y and indicate a percentage along-side, which is your guess that these two numbers were drawn together.]
When you click "Next", you will first receive your *additional hint* on the following page. You have 5 minutes time to decipher the hint. Then you have another 5 minutes of time to indicate you guess. The remaining time will be displayed on the upper right corner of the pages.

Your Additional Hint

Your additional hint for the guess of X [X and Y] is: FJ. The completely decrypted hint indicates the sum of the 2 drawn numbers, i.e., X + Y.
Decryption Instructions.

a. Transform the sequence of letters into a number using the code table.

1	2	3	4	5	6	7	8	9	0
A	B	C	D	E	F	G	H	I	J

b. Add 20 to the number

[*Calculator provided.*] On the next page you will see the entry fields for your guess. You can now enter your decrypted additional hint below, then it will be displayed again on the next page.**Your deciphered additional hint reads: ...** Once you click on "Next" you have 5 minutes time to indicate your guess.

G.4 Decision Screens: Laboratory, Baseline Study

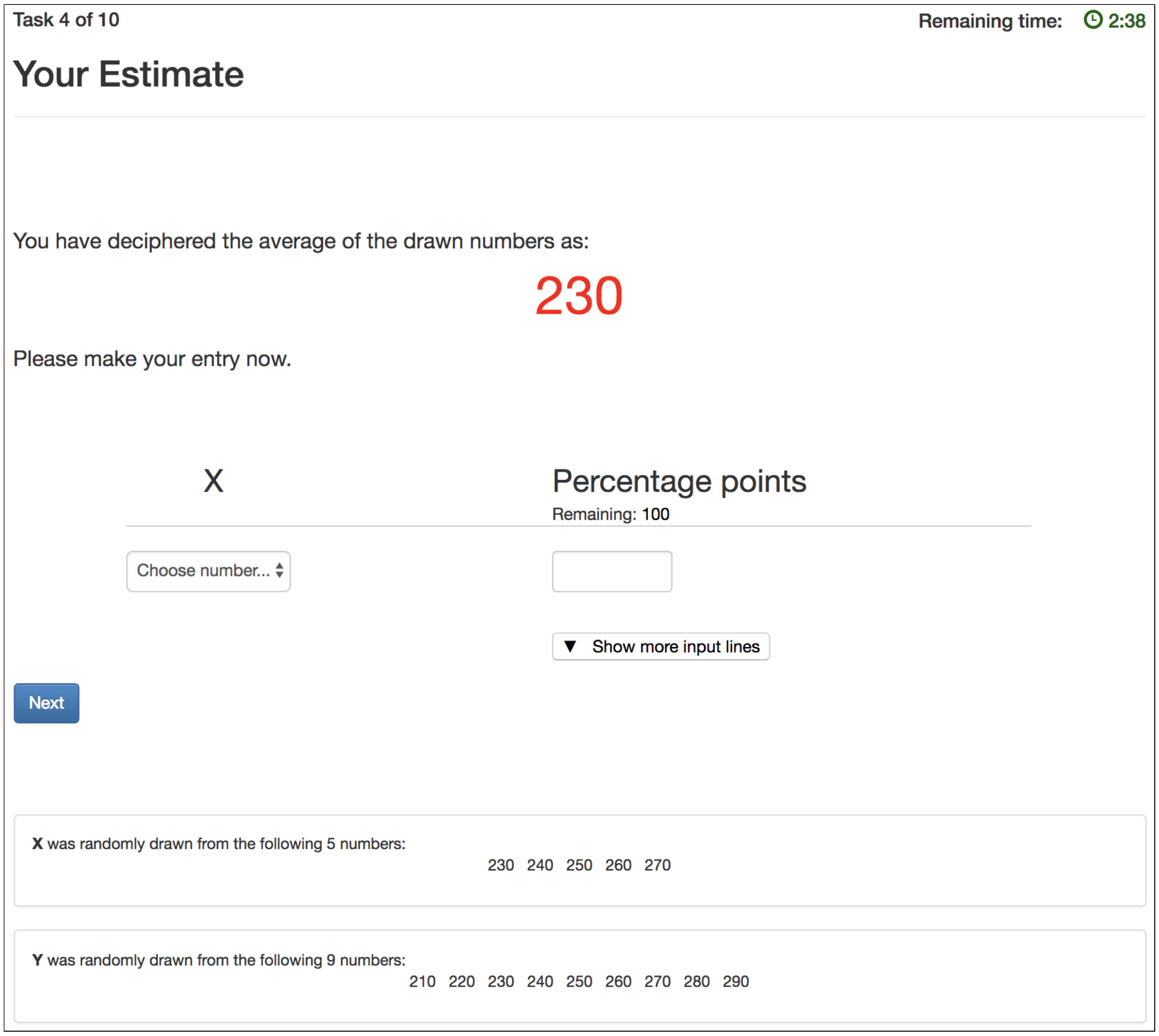


Figure 30: Exemplary decision screen in condition *Narrow* (translated from German). The number 230 indicates the average of X and Y. Subjects state their belief by indicating a full posterior distribution for X. They have to select values for X using the dropdown menu and enter a number of percentage points in the fields on the right. They can use arbitrary many entry lines. The current sum of percentage points is indicated and has to equal exactly 100 before one can proceed. On the bottom of the screen, the distributions of X and Y are indicated as a reminder.

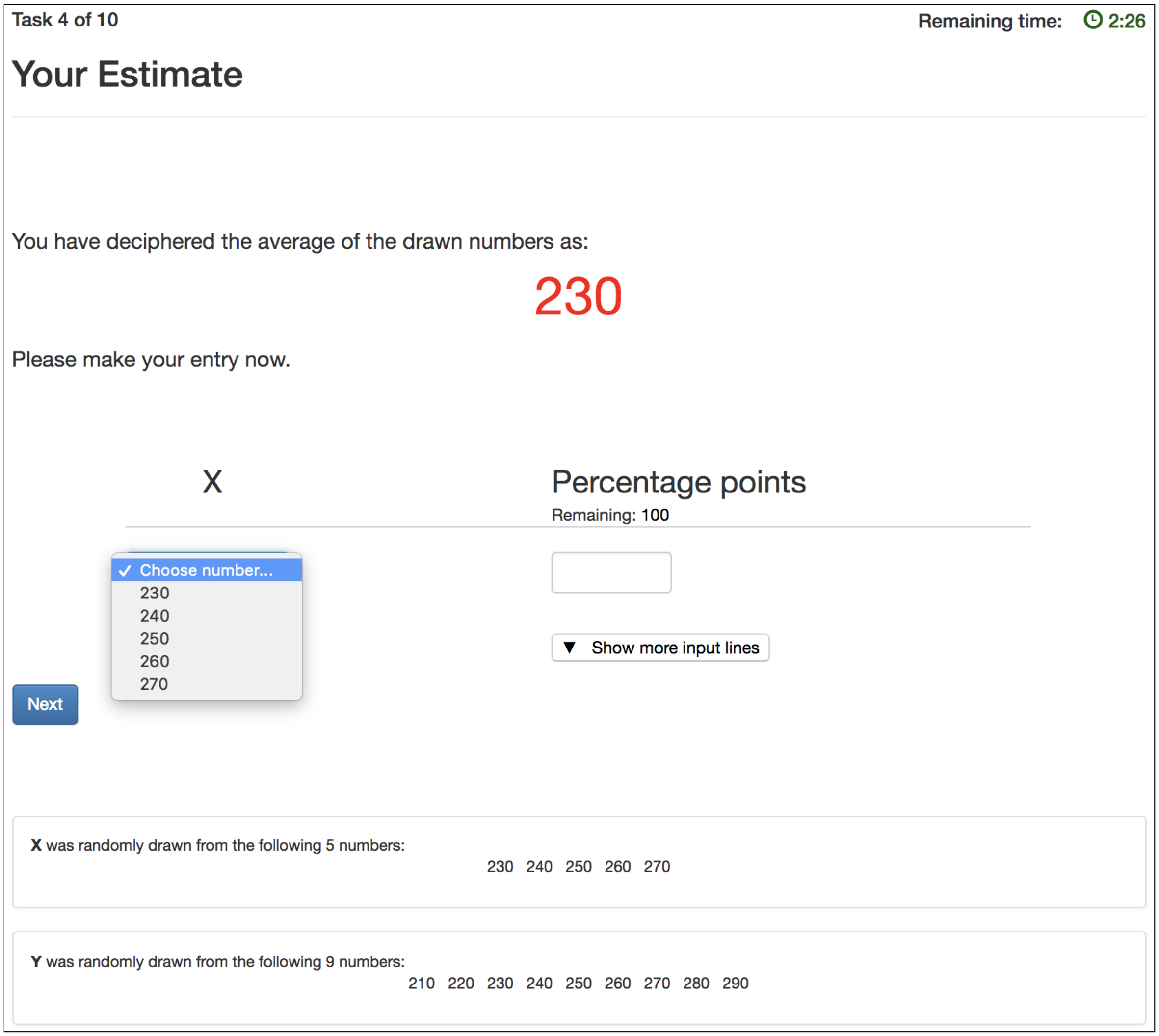


Figure 31: Exemplary decision screen in condition *Narrow* (translated from German). Use of dropdown menu.

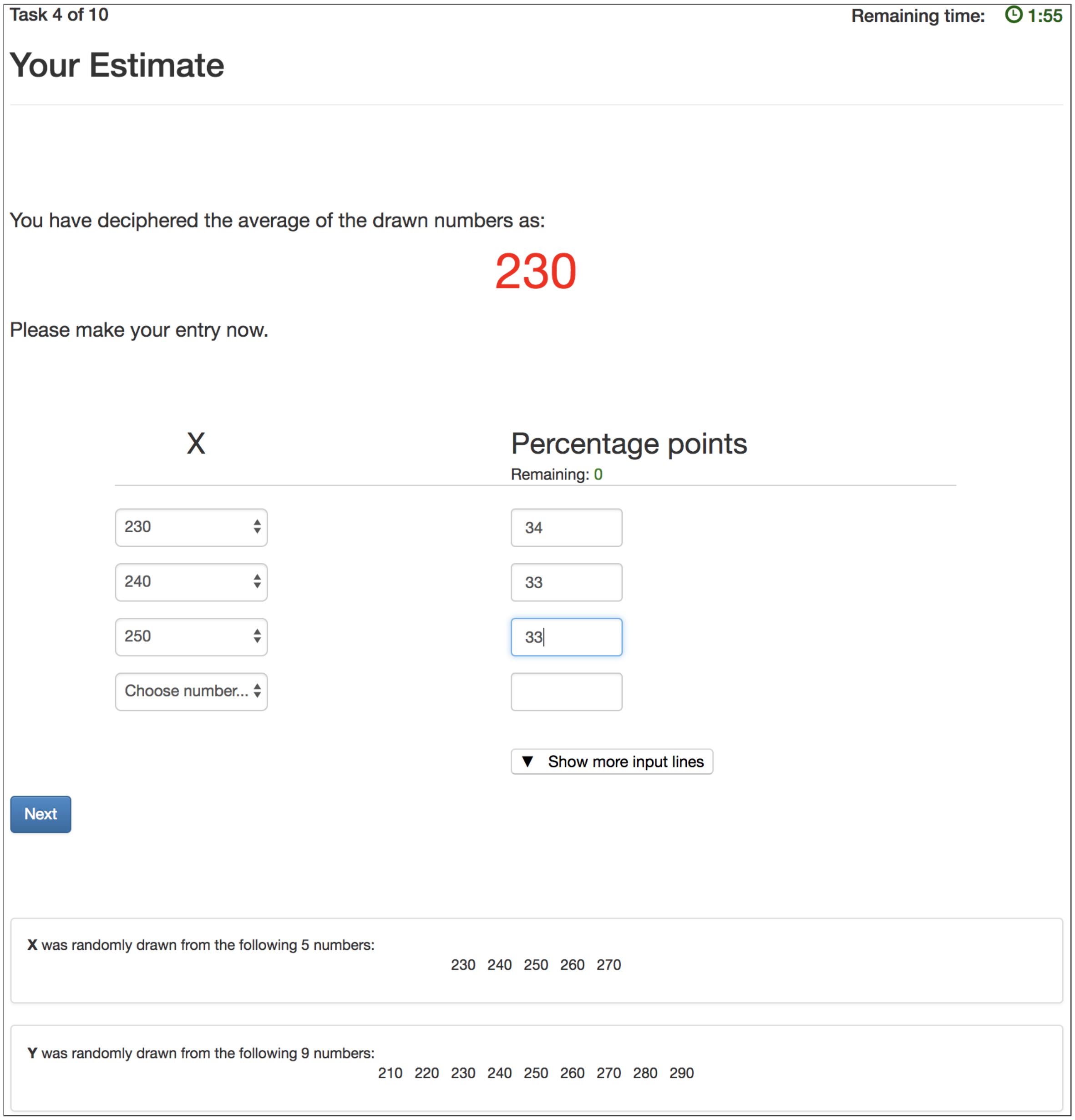


Figure 32: Exemplary decision screen in condition *Narrow* (translated from German). Use of multiple entry rows to indicate the full subjective distribution.

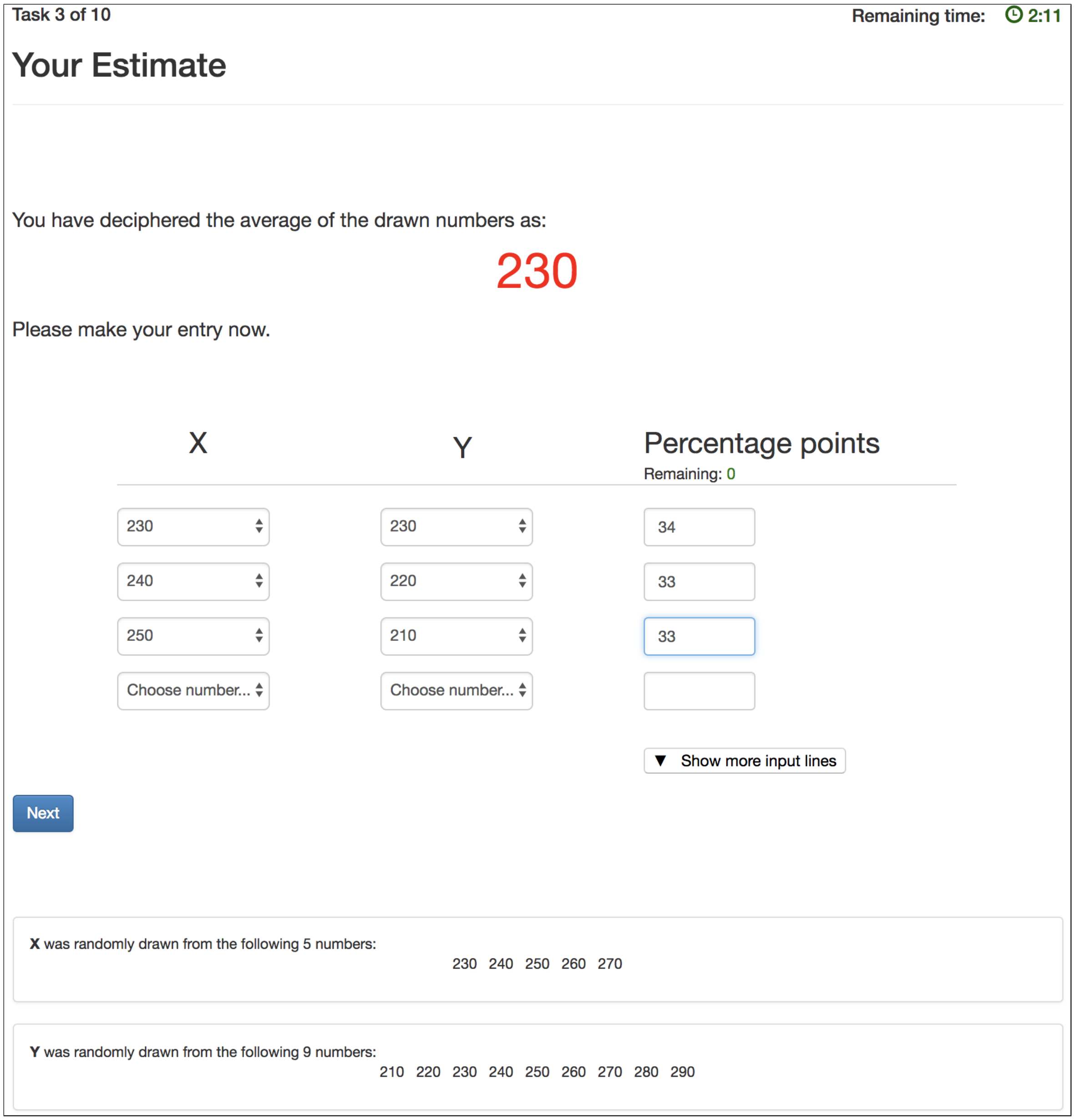


Figure 33: Exemplary decision screen in condition *Broad* (translated from German). The number 230 indicates the average of X and Y. Subjects state their belief by indicating a full posterior distribution for X and Y. They have to select values using the dropdown menu and enter a number of percentage points in the fields on the right. They can use arbitrary many entry lines. The current sum of percentage points is indicated and has to equal exactly 100 before one can proceed. On the bottom of the screen, the distributions of X and Y are indicated as a reminder.

G.5 Decision Screens: Online, Baseline Study

Instructions

You will play 9 guessing games. In each game, you have to guess a randomly drawn number, called X, that is unknown to you. The closer your guess is to the drawn number X, the higher your bonus payment will be.
In every round, **a new number X is drawn** from an urn. The **composition of the urn may differ between rounds**.

Your bonus payment will be determined as follows:

- At the end of the study, one of the 9 rounds will be randomly chosen to count.
- Your bonus depends on how close your guess is to the drawn number in that round. If you have guessed correctly, you will receive $3.00 as bonus payment for this part of the study.[1]

The 9 rounds are fully independent of one another. To maximize your bonus, you should think carefully in each round. A typical participant spends around 1 minute per round.

On the next page, you can play a trial round. The trial round does not affect your bonus payment.

Next

[1]Your payoff is calculated using the following formula: $3 - 0.2 \times (\text{your guess} - \text{drawn number})^2$.
For example, if your guess was 6, and the drawn number was 8, you will receive
$3 - 0.2 \times (6 - 8)^2 = \$2.20$.

Figure 34: Instructions screen in online experiment.

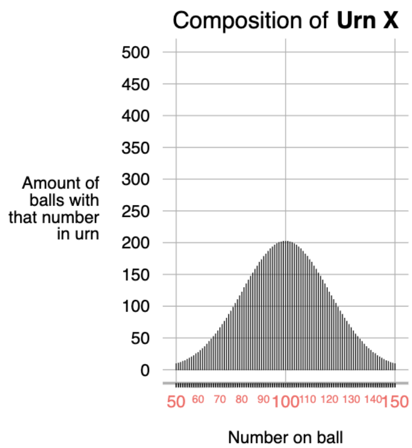
Game 1 of 9

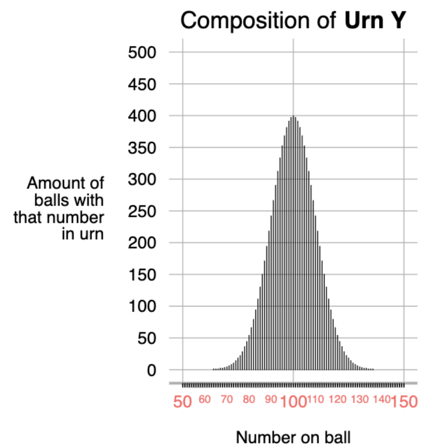
Instructions

Imagine there are **two urns, Urn X and Urn Y**. Both urns contain balls.

Every ball has a whole number on it. In Urn X, this number can be any whole number between 50 and 150, that means 50, or 51, or 52, and so on up to 150. In Urn Y, this number can be any whole number between 50 and 150, that means 50, or 51, or 52, and so on up to 150.
Each urn **contains exactly 10,000 balls**. The picture below shows how many times each number occurs in an urn. A higher bar means that more balls with that number are contained in the urn.
Ball X was **randomly drawn from Urn X** and **ball Y was randomly drawn from Urn Y**.

Reminder: Your are only paid to guess the number on ball X.





The following hint for guessing X is **the sum of the two drawn numbers, X + Y.**

Your hint is **147**.

What is the number X?

Next

Figure 35: Screenshot of example task in online experiment.

G.6 Decision Screens: Online, Vignette Studies

Welcome

Thank you for participating in this study. This study will take **approximately 3 to 5 minutes** to complete.

You will earn a **reward of $0.80 for completing** the study in its entirety. To complete the study, you will need to read all instructions carefully.

Instructions

In what follows, you will receive information about a real-world scenario and will be asked to make a decision or an assessment based on this scenario.

Important: To receive your completion payment, please **carefully read** the scenario.

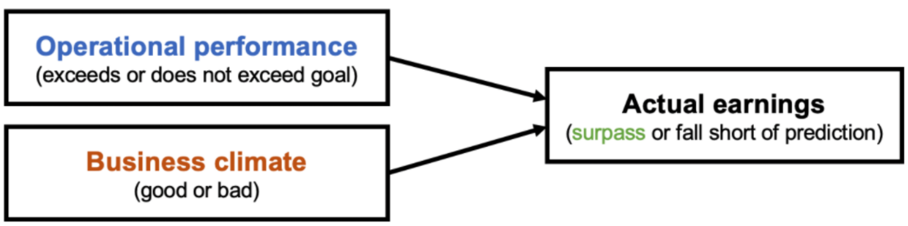
In total, you will be presented with two separate scenarios. Please evaluate these independently of one another.

Figure 36: Screenshot of instructions in vignette study.

The hypothetical company "The Manufacturist" makes quarterly earnings announcements. Actual earnings either surpass or fall short of the analysts' earnings prediction. Actual earnings are determined by 2 factors: (1) the company's operational performance and (2) the business climate in the industry.

(1) Whenever the company exceeds its operational performance goal, this will **cause** realized earnings to surpass analyst predictions **with 70% probability**, independent of whether the business climate is also good or normal.
The company exceeds its operational performance goal 50% of the time.

(2) Whenever the business climate is good, this will **cause** realized earnings to surpass analyst expectations **with 90% probability**, independent of whether the company also exceeds its operational performance goal or not.
The business climate is good 50% of the time.



You now learn that **realized earnings surpassed the analysts' earnings goal**.

Given that realized earnings surpassed the analysts' earnings goal, **what do you think is the likelihood (percentage chance) that...**

(enter two numbers between 0 and 100)

... the business climate was good?

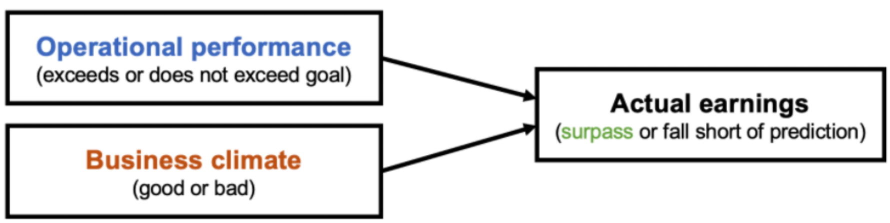
... the company exceeded its operational performance goal?

Figure 37: Screenshot of *Earnings* vignette, treatment *Broad*, outcome *Belief Probabilistic*.

The hypothetical company "The Manufacturist" makes quarterly earnings announcements. Actual earnings either surpass or fall short of the analysts' earnings prediction. Actual earnings are determined by 2 factors: (1) the company's operational performance and (2) the business climate in the industry.

(1) Whenever the company exceeds its operational performance goal, this will **cause** realized earnings to surpass analyst predictions **with 70% probability**, independent of whether the business climate is also good or normal.
The company exceeds its operational performance goal 50% of the time.

(2) Whenever the business climate is good, this will **cause** realized earnings to surpass analyst expectations **with 90% probability**, independent of whether the company also exceeds its operational performance goal or not.
The business climate is good 50% of the time.



You now learn that **realized earnings surpassed the analysts' earnings goal**.

Given that realized earnings surpassed the analysts' earnings goal, **what do you think is the likelihood (percentage chance) that...**

(enter a number between 0 and 100)

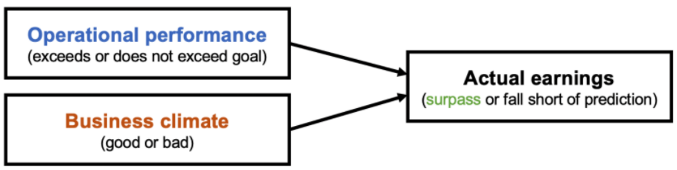
... the company exceeded its operational performance goal?

Figure 38: Screenshot of *Earnings* vignette, treatment *Narrow*, outcome *Belief Probabilistic*.

The hypothetical company "The Manufacturist" makes quarterly earnings announcements. Actual earnings either surpass or fall short of the analysts' earnings prediction. Actual earnings are determined by 2 factors: (1) the company's operational performance and (2) the business climate in the industry.

(1) Whenever the company exceeds its operational performance goal, this will **cause** realized earnings to surpass analyst predictions **with 70% probability**, independent of whether the business climate is also good or normal.
The company exceeds its operational performance goal 50% of the time.

(2) Whenever the business climate is good, this will **cause** realized earnings to surpass analyst expectations **with 90% probability**, independent of whether the company also exceeds its operational performance goal or not.
The business climate is good 50% of the time.



You now learn that **realized earnings surpassed the analysts' earnings goal**.
You have $2. You can either keep this money or make up to two separate bets.

- If you bet $1 on the business climate and the business climate WAS GOOD, you **win an additional $2 for a total of $3.**
- If you bet $1 on the business climate and the business climate WAS **NOT** GOOD, you **lose this bet** and don't receive any money.
- If you bet $1 on operational performance and the company DID EXCEED its operational performance goal, you **win an additional $2 for a total of $3.**
- If you bet $1 on operational performance and the company DID **NOT** EXCEED its normal operational performance goal, you **lose this bet** and don't receive any money.

Given that **realized earnings surpassed the analysts' earnings goal, do you want to...**

Keep $1	Bet $1 on good industry climate?

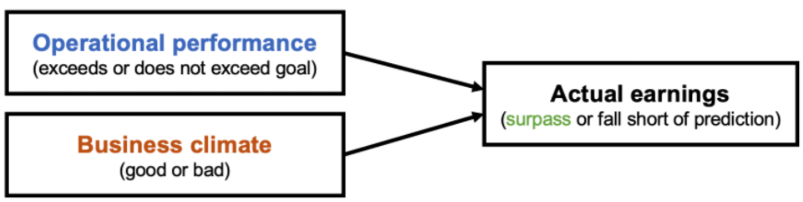
Keep $1	Bet $1 on the company having exceeded the operational performance goal

Figure 39: Screenshot of *Earnings* vignette, treatment *Broad*, outcome *Action*.

The hypothetical company "The Manufacturist" makes quarterly earnings announcements. Actual earnings either surpass or fall short of the analysts' earnings prediction. Actual earnings are determined by 2 factors: (1) the company's operational performance and (2) the business climate in the industry.

(1) Whenever the company exceeds its operational performance goal, this will **cause** realized earnings to surpass analyst predictions **with 70% probability**, independent of whether the business climate is also good or normal.
The company exceeds its operational performance goal 50% of the time.

(2) Whenever the business climate is good, this will **cause** realized earnings to surpass analyst expectations **with 90% probability**, independent of whether the company also exceeds its operational performance goal or not.
The business climate is good 50% of the time.



You now learn that **realized earnings surpassed the analysts' earnings goal**.

You have $1. You can either keep or bet this $1 on the company's operational performance.

- If you bet your money and the company DID EXCEED its operational performance goal, you **win an additional $2 for a total of $3.**
- If you bet your money and the company DID **NOT** EXCEED its normal operational performance goal, you **lose your bet** and don't receive any money.

Given that **realized earnings surpassed the analysts' earnings goal, do you want to...**

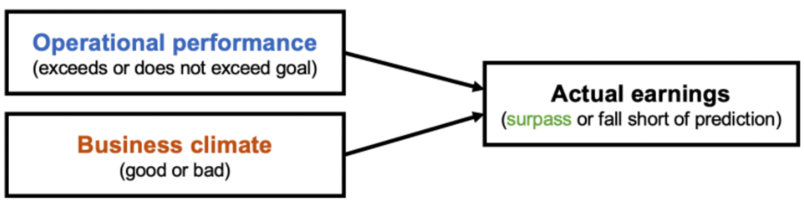
Keep $1	**Bet $1 on the company having exceed the operational performance goal**

Figure 40: Screenshot of *Earnings* vignette, treatment *Narrow*, outcome *Action*.

The hypothetical company "The Manufacturist" makes quarterly earnings announcements. Actual earnings either surpass or fall short of the analysts' earnings prediction. Actual earnings are determined by 2 factors: (1) the company's operational performance and (2) the business climate in the industry.

(1) Whenever the company exceeds its operational performance goal, this will **cause** realized earnings to surpass analyst predictions **with moderately high probability**, independent of whether the business climate is also good or normal.
It is equally likely that the company exceeds or does not exceed its operational performance.

(2) Whenever the business climate is good, this will **cause** realized earnings to surpass analyst expectations **with very high probability**, independent of whether the company also exceeds its operational performance goal or not.
It is equally likely that the business climate is good or normal.



You now learn that **realized earnings surpassed the analysts' earnings goal**.

Given that realized earnings surpassed the analysts' earnings goal, **what do you think was...**

... the business climate?

Good	**Normal**

... the operational performance?

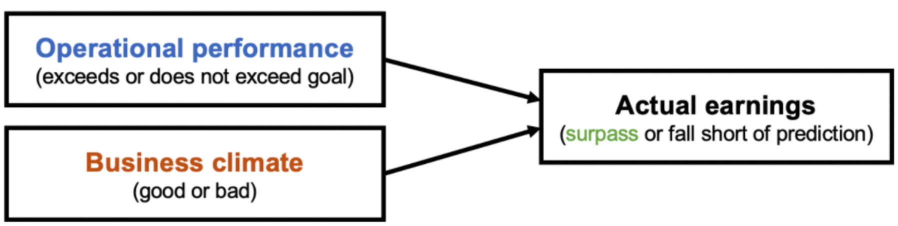
Exceeded performance goal	**Did NOT exceed** performance goal

Figure 41: Screenshot of *Earnings* vignette, treatment *Broad*, outcome *Belief Simple*.

The hypothetical company "The Manufacturist" makes quarterly earnings announcements. Actual earnings either surpass or fall short of the analysts' earnings prediction. Actual earnings are determined by 2 factors: (1) the company's operational performance and (2) the business climate in the industry.

(1) Whenever the company exceeds its operational performance goal, this will **cause** realized earnings to surpass analyst predictions **with moderately high probability**, independent of whether the business climate is also good or normal.
It is equally likely that the company exceeds or does not exceed its operational performance.

(2) Whenever the business climate is good, this will **cause** realized earnings to surpass analyst expectations **with very high probability**, independent of whether the company also exceeds its operational performance goal or not.
It is equally likely that the business climate is good or normal.



You now learn that **realized earnings surpassed the analysts' earnings goal**.

Given that realized earnings surpassed the analysts' earnings goal, **what do you think was...**

... the operational performance?

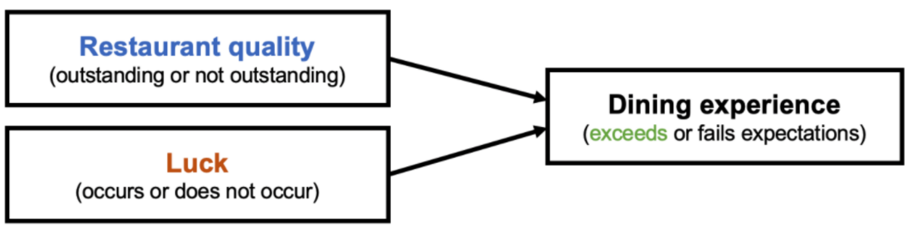
Exceeded performance goal	**Did NOT exceed** performance goal

Figure 42: Screenshot of *Earnings* vignette, treatment *Narrow*, outcome *Belief Simple*.

Assume that you have dinner, for the first time, at a new restaurant called "The Dinnery". Assume further that your dining experience either exceeds or fails your expectations based on similar restaurants. Your dining experience is determined by 2 factors: (A) the restaurant's actual quality and (B) luck that is unrelated to the restaurant itself, e.g., you may have come on a sunny day, you were in an exceptionally good mood or dined with particularly enjoyable company.

(A) Whenever the actual restaurant quality is outstanding, it will cause your dining experience to exceed your expectations **with 95% probability**, independent of whether good luck is also involved or not.
The actual restaurant quality is outstanding 50% of the time.

(B) Whenever good luck is involved, it will cause your dining experience to exceed your expectations **with 80% probability,** independent of whether actual restaurant quality is outstanding or not.
Good luck occurs 50% of the time.



Assume that your **actual dining experience exceeded your expectations**.

Given that your dining experience exceeded your expectations, **what do you think is the likelihood (percentage chance) that...**

(enter two numbers between 0 and 100)

... good luck was involved?

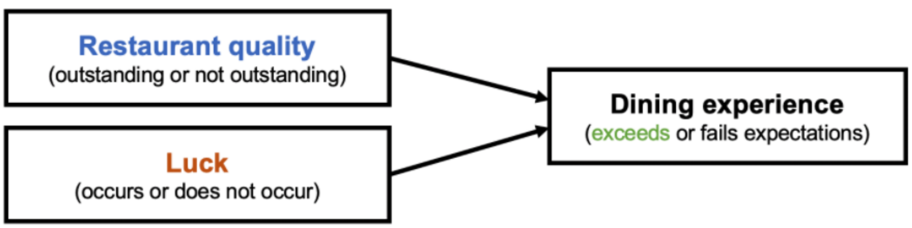
... the actual restaurant quality is outstanding?

Figure 43: Screenshot of *Restaurant* vignette, treatment *Broad*, outcome *Belief Probabilistic*.

Assume that you have dinner, for the first time, at a new restaurant called "The Dinnery". Assume further that your dining experience either exceeds or fails your expectations based on similar restaurants. Your dining experience is determined by 2 factors: (A) the restaurant's actual quality and (B) luck that is unrelated to the restaurant itself, e.g., you may have come on a sunny day, you were in an exceptionally good mood or dined with particularly enjoyable company.

(A) Whenever the actual restaurant quality is outstanding, it will cause your dining experience to exceed your expectations **with 95% probability**, independent of whether good luck is also involved or not.
The actual restaurant quality is outstanding 50% of the time.

(B) Whenever good luck is involved, it will cause your dining experience to exceed your expectations **with 80% probability,** independent of whether actual restaurant quality is outstanding or not.
Good luck occurs 50% of the time.



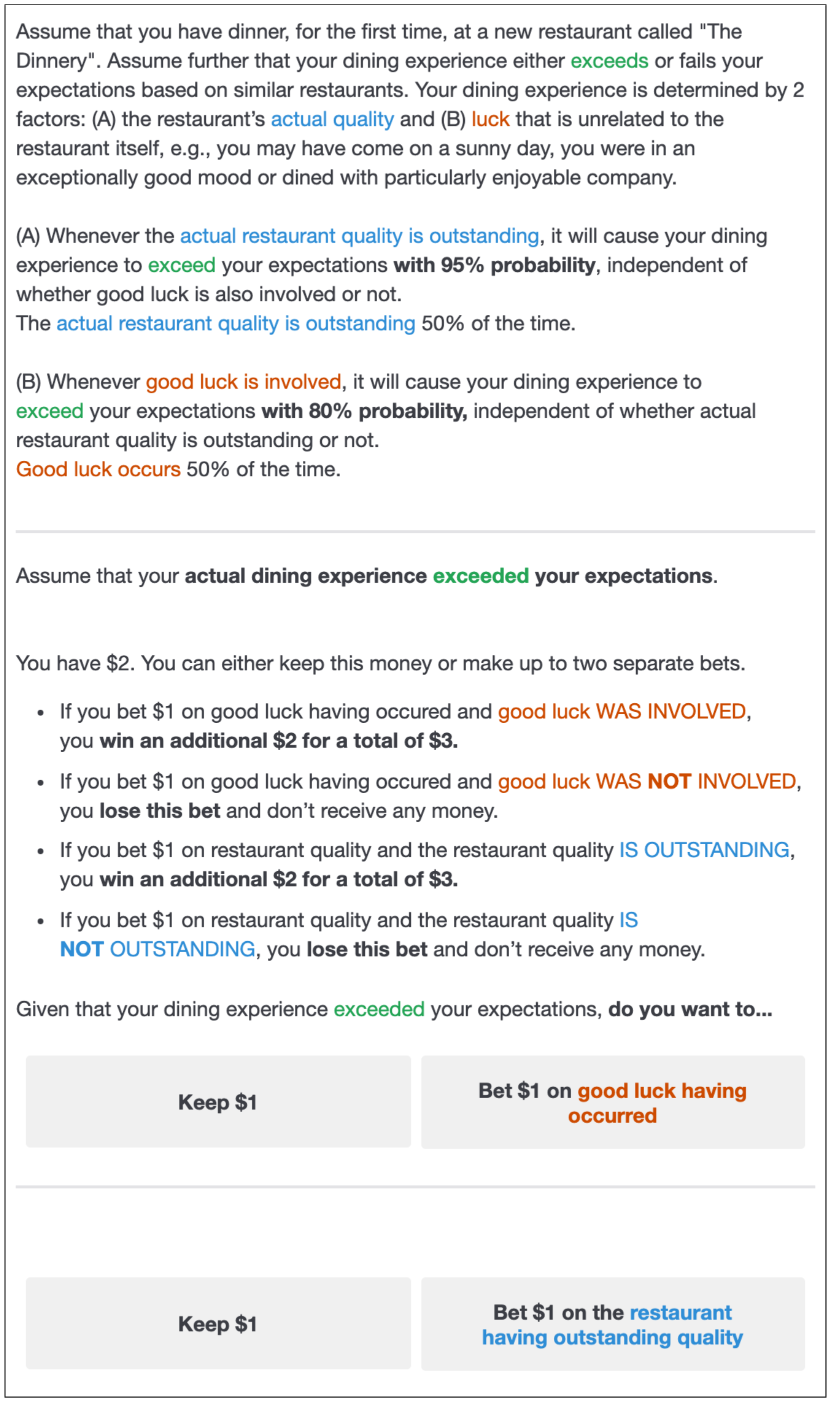
Assume that your **actual dining experience exceeded your expectations**.

Given that your dining experience exceeded your expectations, **what do you think is the likelihood (percentage chance) that...**

(enter a number between 0 and 100)

... the actual restaurant quality is outstanding?

Figure 44: Screenshot of *Restaurant* vignette, treatment *Narrow*, outcome *Belief Probabilistic*.



Assume that you have dinner, for the first time, at a new restaurant called "The Dinnery". Assume further that your dining experience either exceeds or fails your expectations based on similar restaurants. Your dining experience is determined by 2 factors: (A) the restaurant's actual quality and (B) luck that is unrelated to the restaurant itself, e.g., you may have come on a sunny day, you were in an exceptionally good mood or dined with particularly enjoyable company.

(A) Whenever the actual restaurant quality is outstanding, it will cause your dining experience to exceed your expectations **with 95% probability**, independent of whether good luck is also involved or not.
The actual restaurant quality is outstanding 50% of the time.

(B) Whenever good luck is involved, it will cause your dining experience to exceed your expectations **with 80% probability,** independent of whether actual restaurant quality is outstanding or not.
Good luck occurs 50% of the time.

Assume that your **actual dining experience exceeded your expectations**.

You have $2. You can either keep this money or make up to two separate bets.

- If you bet $1 on good luck having occured and good luck WAS INVOLVED, you **win an additional $2 for a total of $3.**
- If you bet $1 on good luck having occured and good luck WAS **NOT** INVOLVED, you **lose this bet** and don't receive any money.
- If you bet $1 on restaurant quality and the restaurant quality IS OUTSTANDING, you **win an additional $2 for a total of $3.**
- If you bet $1 on restaurant quality and the restaurant quality IS **NOT** OUTSTANDING, you **lose this bet** and don't receive any money.

Given that your dining experience exceeded your expectations, **do you want to...**

Keep $1 | **Bet $1 on good luck having occurred**

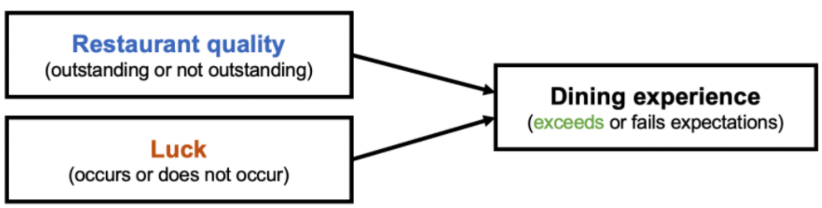
Keep $1 | **Bet $1 on the restaurant having outstanding quality**

Figure 45: Screenshot of *Restaurant* vignette, treatment *Broad*, outcome *Action*.

Assume that you have dinner, for the first time, at a new restaurant called "The Dinnery". Assume further that your dining experience either exceeds or fails your expectations based on similar restaurants. Your dining experience is determined by 2 factors: (A) the restaurant's actual quality and (B) luck that is unrelated to the restaurant itself, e.g., you may have come on a sunny day, you were in an exceptionally good mood or dined with particularly enjoyable company.

(A) Whenever the actual restaurant quality is outstanding, it will cause your dining experience to exceed your expectations **with 95% probability**, independent of whether good luck is also involved or not.
The actual restaurant quality is outstanding 50% of the time.

(B) Whenever good luck is involved, it will cause your dining experience to exceed your expectations **with 80% probability,** independent of whether actual restaurant quality is outstanding or not.
Good luck occurs 50% of the time.



Assume that your **actual dining experience exceeded your expectations**.

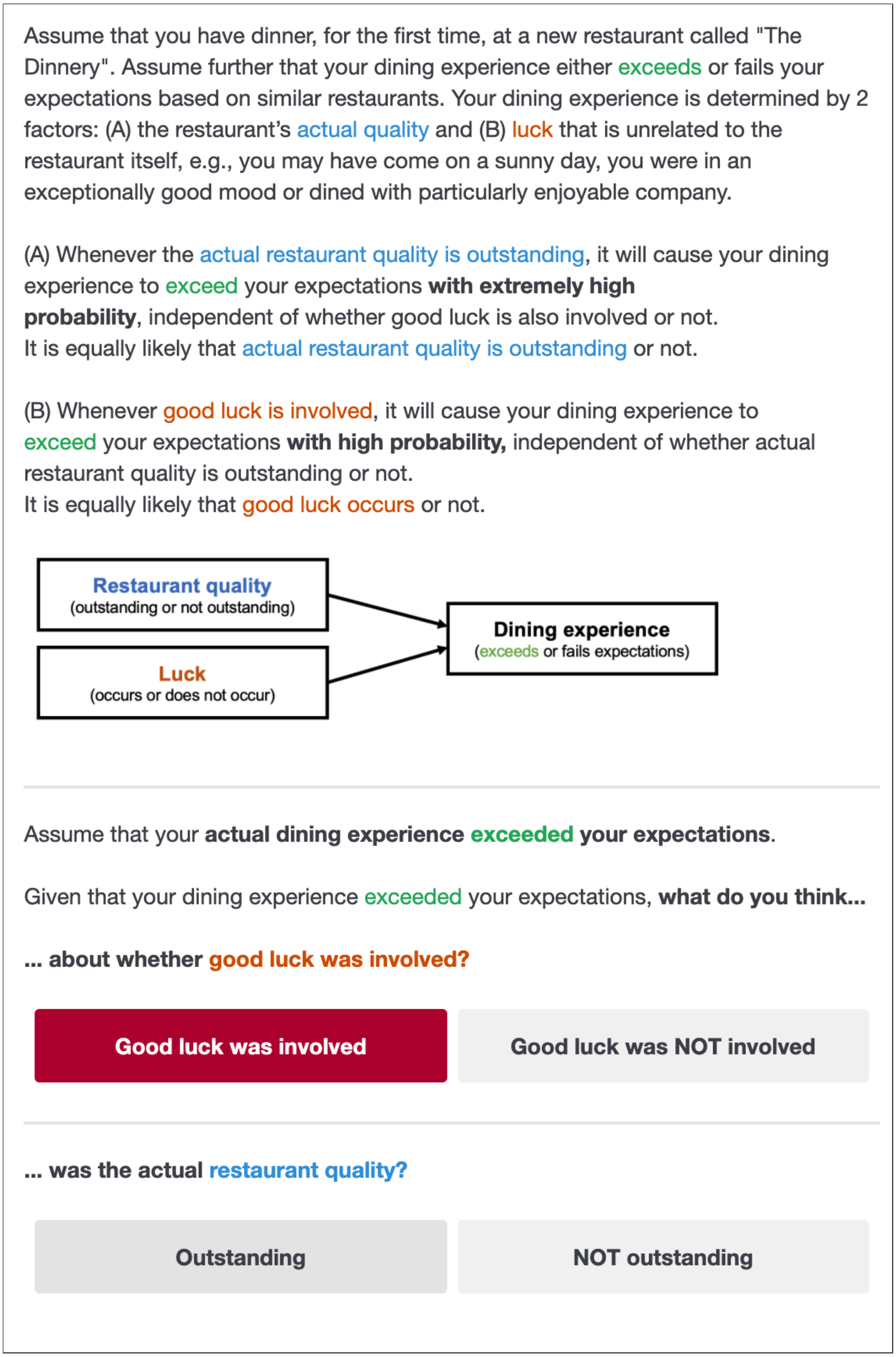
You have $1. You can either keep or bet this $1 on the restaurant quality.

- If you bet your money and the restaurant quality IS OUTSTANDING, you **win an additional $2 for a total of $3.**
- If you bet your money and the restaurant quality IS **NOT** OUTSTANDING, you **lose your bet** and don't receive any money.

Given that your dining experience exceeded your expectations, **do you want to...**

Keep $1	Bet $1 on the restaurant having outstanding quality

Figure 46: Screenshot of *Restaurant* vignette, treatment *Narrow*, outcome *Action*.



Assume that you have dinner, for the first time, at a new restaurant called "The Dinnery". Assume further that your dining experience either exceeds or fails your expectations based on similar restaurants. Your dining experience is determined by 2 factors: (A) the restaurant's actual quality and (B) luck that is unrelated to the restaurant itself, e.g., you may have come on a sunny day, you were in an exceptionally good mood or dined with particularly enjoyable company.

(A) Whenever the actual restaurant quality is outstanding, it will cause your dining experience to exceed your expectations **with extremely high probability**, independent of whether good luck is also involved or not.
It is equally likely that actual restaurant quality is outstanding or not.

(B) Whenever good luck is involved, it will cause your dining experience to exceed your expectations **with high probability,** independent of whether actual restaurant quality is outstanding or not.
It is equally likely that good luck occurs or not.

Assume that your **actual dining experience exceeded your expectations**.

Given that your dining experience exceeded your expectations, **what do you think...**

... about whether good luck was involved?

Good luck was involved | Good luck was NOT involved

... was the actual restaurant quality?

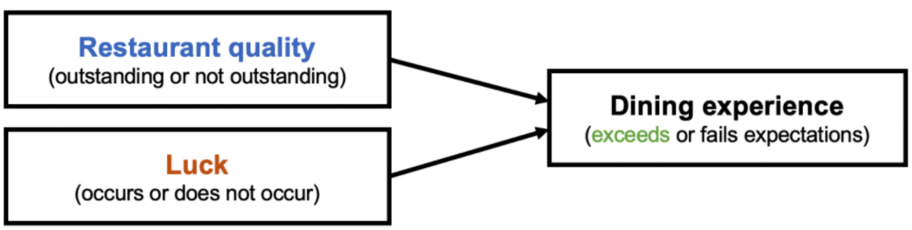
Outstanding | NOT outstanding

Figure 47: Screenshot of *Restaurant* vignette, treatment *Broad*, outcome *Belief Simple*.

Assume that you have dinner, for the first time, at a new restaurant called "The Dinnery". Assume further that your dining experience either exceeds or fails your expectations based on similar restaurants. Your dining experience is determined by 2 factors: (A) the restaurant's actual quality and (B) luck that is unrelated to the restaurant itself, e.g., you may have come on a sunny day, you were in an exceptionally good mood or dined with particularly enjoyable company.

(A) Whenever the actual restaurant quality is outstanding, it will cause your dining experience to exceed your expectations **with extremely high probability**, independent of whether good luck is also involved or not.
It is equally likely that actual restaurant quality is outstanding or not.

(B) Whenever good luck is involved, it will cause your dining experience to exceed your expectations **with high probability,** independent of whether actual restaurant quality is outstanding or not.
It is equally likely that good luck occurs or not.



Assume that your **actual dining experience exceeded your expectations**.

Given that your dining experience exceeded your expectations, **what do you think was...**

... the actual restaurant quality?

Outstanding	NOT outstanding

Figure 48: Screenshot of *Restaurant* vignette, treatment *Narrow*, outcome *Belief Simple*.