Attention, Information, and Persuasion

John J. Conlon

January 12, 2025*

Abstract

I show experimentally that information persuades not only by shifting beliefs but also by redirecting attention. Participants repeatedly choose whether to "purchase" multi-attribute "goods." Information telling participants about the value of a desirable attribute—even when the information is already known, transparently redundant, and explicitly randomly assigned—greatly increases attention to the attribute it describes and distracts from other attributes. It also boosts average demand for the good, implying that inattention takes the form of neglect rather than shrinkage toward a prior. These forces combine to produce paradoxical responses to correcting beliefs: reducing overoptimism about an attribute nonetheless boosts demand for the good.

*Department of Social and Decision Sciences, Carnegie Mellon University: jconlon@andrew.cmu.edu. I thank Chiara Aina, Ned Augenblick, Katherine Coffman, Lucas Coffman, Nicola Gennaioli, Ben Enke, Muriel Niederle, Matthew Rabin, Gautam Rao, Chris Roth, Peter Schwardmann, Josh Schwartzstein, Cassidy Shubatt, Jason Somerville, Andrei Shleifer, and Jeffrey Yang for helpful comments. I am grateful for financial support from the National Science Foundation Graduate Research Fellowship under grant DGE1745303, and the Alfred P. Sloan Foundation Predoctoral Fellowship in Behavioral Macroeconomics, awarded through the NBER. The experiment described in this paper was pre-registered on the AEA registry (ID 0010434).

1 Introduction

How does information persuade? The traditional view in economics is that messages—from advertisers selling a product, mentors dispensing advice, or experimenters conducting RCTs—affect behavior solely by shifting recipients' beliefs. But many messages appear to convey information that their intended audience already knows: that a well-known soft drink tastes refreshing, that immigrants are different than you, or that a religion promises salvation. Evidently, these messages work: but how? One hypothesis is that they persuade by shifting what their recipients pay attention to. If so, what must be true about how attention is allocated for these messages to change people's behavior in the way their senders intend?

In this paper, I study these questions experimentally. I first show that choice data are generally insufficient for identifying attention unless agents' preferences and beliefs are known. This observation motivates my focus on a novel experimental task where such control is possible. Participants repeatedly face a binary choice between two abstract "goods" (see Figure 1 in Section 3 for example decisions). One good offers participants an amount of money with certainty plus a lottery that pays a larger sum with a small probability. The alternative good has six attributes (three types of coins and three boxes whose value depends on their color), each of which contributes to the participant's payment if she chooses that option. Crucially, preferences over attributes are known ex ante, since they all ultimately map onto cash, and participants' beliefs can be measured by asking them about this mapping (i.e., checking their comprehension and memory). Random variation in each good's attributes then allows me to estimate empirical responsiveness to their exogenously changing values, which I show is a sufficient statistic to estimate attention.[1]

I find first that, even in this simplified choice environment (compared to real-world economic decisions), the average participant is far from the full-attention "rational" benchmark: equivalent changes in the value of different attributes produce significantly different demand responses. Compared to the attribute participants most respond to, choices react between 25% and 89% less to equal-value changes in other attributes ($p < 0.01$ for all comparisons). These distortions appear despite the vast majority of participants being able, in unincentivized debriefing questions, to correctly

[1]Note that I use the term "attention" broadly to mean the weight that attributes receive (also sometimes called "decision weights"). I discuss in Section 2 the connection between this notion and other definitions of attention (e.g., gaze, "top of mind", etc.).

describe how the possible values of every attribute would contribute to their earnings. Very similar average distortions appear even among participants who are perfectly informed in this way, suggesting that deviations from the rational benchmark stem not from incorrect beliefs but from failures of full attention.

Next, a subset of participants receive information telling them about the value of one randomly selected attribute. Recall that most participants are already perfectly well-informed about these values. They are also told explicitly beforehand that information is randomly assigned independently of the importance of the attribute to the decision. These facts imply that such information cannot affect choices by changing participants' beliefs, but their inattentiveness at baseline raises the possibility that it nonetheless may alter their choices if it shifts which attributes they focus on. Indeed, I find that this information starkly increases responsiveness to the attribute it describes. These effects are large; responsiveness to an attribute increases by 69% ($p < 0.01$) on average in response to information describing it. I find similar effects (77% increase, $p < 0.01$) when restricting the data to the large majority of participants who already know this information, and thus these effects appear to operate primarily by redirecting participants' attention rather than by changing their beliefs.

These results show that information provision shifts attention, but three additional results shed light on the underlying drivers of attention and on how these effects operate. First, I find evidence of attention *spillovers*: information about one attribute boosts attention to it in part by decreasing responsiveness to other attributes (by 10% in both the full sample and among participants with correct beliefs, $p = 0.02$ for both comparisons). This spillover effect, though smaller in percentage terms than the direct effect on the attribute the information describes, applies to many more attributes. The total spillover effect, summing across all these attributes, therefore amounts to 75% of the direct effect. Because the direct and total spillover effects operate in opposite directions and are comparable in magnitude, the total effect of information on attention across all attributes is small and statistically indistinguishable from zero ($p = 0.43$). These findings point to an attentional capacity constraint. They imply that a communicator (a policy-maker, an advertiser) potentially faces tradeoffs when trying to manipulate attention through information provision: directing focus toward a particular feature distracts from others.

Second, information about an attribute boosts demand for its associated good by about 2 percentage points ($p < 0.01$), despite the fact (as described above) that information provision is known to be random and thus uncorrelated with the value of the attribute or of the good as a whole. I show theoretically that this effect depends on what I call the "default" value: how an attribute is treated when the agent fails to attend to it. The fact that the average effect of information is positive suggests that this default is *zero*. That is, when failing to attend to an attribute, agents fail to incorporate it into their decision at all, rather than simply relying on its expected or average value. Inattention therefore looks like *neglect*, rather than shrinkage toward a prior.

Third, though the attentional effect of information is large, it is also fragile. Information about one attribute continues to increase attention to it (by 86%, $p < 0.01$) even after it is no longer displayed, but only so long as no other information takes its place. If new information about a different attribute begins to appear, attention to the original attribute falls by 44% ($p < 0.01$), almost completely reverting to the level it would have maintained absent any information (an insignificant 5% difference, $p = 0.85$). These patterns corroborate my interpretation that attention drives these results and speak against alternative mechanisms such as biased priors, risk aversion, caution, or any other belief-based persuasion channel.

Finally, these effects of information on attention can combine to mask and even overturn their effects on beliefs. Recall that one of the alternatives participants can choose has a lottery associated with it. If they choose this option, they get to roll five six-sided dice, earning an additional bonus if their roll adds up to 12 or less. Participants substantially overestimate the odds of winning this lottery, with the average participant believing she has a 27% chance of winning, whereas the true odds are close to 10%. Nonetheless, information that alerts participants of the true odds of winning—and therefore provides bad news relative to their priors—actually *boosts* demand for the option that includes the lottery by 5.1 percentage points ($p < 0.01$). The explanation of this paradoxical result is that the information, in addition to telling participants about the unfavorable odds, also points their attention toward the lottery. Because it is a positive attribute of the good, which participants may otherwise neglect, this boosts demand despite the bad news it conveys. Adding evidence to this view, an almost identical piece of information that describes the lottery but does not provide the odds of winning boosts demand even more (by 9.1 compared to 5.1 percentage points, $p = 0.03$).[2]

[2] In addition, the difference in effects between the informative and non-informative messages is

My focus on an controlled experimental environment is intentional, as it allows me to cleanly identify the attention channel through which information might persuade. First, in field settings, it is typically impossible to rule out belief-based channels even if information changes choices in unexpected ways.[3] Second, how to estimate (or even define) attention is unclear unless agents have linear preferences, which my experiment induces. Third, my experiment allows me to measure preferences, beliefs, and attention regarding every dimension that is relevant for participants' choices. Finally, it also lets me provide information that is explicitly randomized and obviously not a signal that recipients should change their beliefs about anything else. All four of these features are crucial for untangling the beliefs and attention effects of information provision.

This paper contributes first to a literature studying non-traditional persuasion.[4] For example, previous work has studied how information senders can influence recipients' actions by shifting the analogies they use to solve problems (Mullainathan et al., 2008), making messages simpler or more visually appealing (Bertrand et al., 2010), invoking other-regarding preferences (Coffman & Niehaus, 2020), or shifting the causal model people adopt (Schwartzstein & Sunderam 2021, Aina 2023, Barron & Fries 2023, Charles & Kendall 2024). The distinction between belief-based and non-belief based persuasion is not always sharp, but my results provide a particularly clear example of the latter: providing information that recipients already know (and, given the experimental environment, cannot shift their beliefs about anything else) redirects attention and thereby choices.

These results add to the connection between the literature on persuasion and the literature more broadly studying decision-making in complex environments, where agents struggle to incorporate all relevant information before deciding (e.g., Enke & Zimmermann 2019, Martínez-Marquina et al. 2019, Oprea 2020, 2024, Enke & Graeber 2023, Ba et al. 2023, Bordalo et al. 2023; Bordalo, Gennaioli, et al. 2024,

larger for participants with particularly erroneous beliefs ($p < 0.01$).

[3]For example, a message about the current level of inflation might change recipients' beliefs about many things besides just the narrow fact it describes: e.g., about how important inflation is for economic outlook, about how important the sender *thinks* inflation is, about other facts like possible government responses or political repercussions, and so on. Similar concerns arise even if recipients already know the information the sender's message conveys or even if the message does not convey any hard information at all (e.g., the fact that a message is trying to make inflation salient to the recipient might reasonably change their beliefs).

[4]For a review of more standard Bayesian persuasion, see Kamenica (2019).

Bohren et al. 2024, Augenblick et al. 2024, Enke et al. 2024, Arrieta & Nielsen 2024, Shubatt & Yang 2024). My results highlight how manipulations seemingly operating through one "standard" channel—such as information targeting beliefs—may in fact alter behavior also (and possibly to a greater extent) by changing how agents simplify or resolve the complexities of a choice.

Second, my evidence contributes to a growing body of theoretical work investigating how agents allocate attention when decisions have multiple relevant features (See Loewenstein & Wojtowicz 2023 for a review of attention in economics).[5] My result that obviously uninformative messages have such large (but fragile) attentional effects points toward theories that give a central role to contextual factors such as visual prominence (e.g., Bordalo et al. 2022, 2023) rather than only the objective payoffs involved (as in, e.g., Koszegi & Szeidl 2013, Bordalo et al. 2013, Gabaix 2014, Matějka & Mckay 2015, and Bushong et al. 2021). Attentional capacity constraints appear in some existing models (e.g., Bordalo et al. 2022), though to my knowledge my experimental evidence on their importance is novel.[6] Next, my result that individuals downweight unattended-to attributes rather than relying on their priors is consistent with models of inattention as neglect; for example, though Koszegi & Szeidl (2013), Bordalo et al. (2022), and Bushong et al. (2021) make different assumptions about why attributes escape attention, all assume that when they do they are simply ignored or downweighted. This is in contrast to standard theories of rational inattention, which typically assume that absent attention agents fall back on their priors about attributes (e.g., Sims 2011, Gabaix 2019).[7] To my knowledge, my evidence testing between these possibilities is novel to this literature.

In addition to theoretical and lab-experimental studies, some applied work studies how people appear to neglect "shrouded attributes," and how reminders about such attributes can affect choices (e.g., Chetty et al. 2009, Brown et al. 2010, Allcott & Taubinsky 2015, Taubinsky & Rees-Jones 2018, Bradley & Feldman 2020). My

[5]For experiments testing some of these models, see among others Dertwinkel-Kalt et al. (2017), Frydman & Mormann (2017), Li & Camerer (2022), Somerville (2022), and Dean & Neligh (2023). Other work models attention as gathering additional (sometimes sensory) information before making a decision (e.g., Caplin & Dean 2015; Sepulveda et al. 2020, Yang & Krajbich 2022).

[6]My distraction result is reminiscent of Altmann et al. (2021), who find that incentivizing one task reduces engagement in another task. My findings show that these sorts of cognitive spillovers arise even within individual decisions and without modifying the objective incentives that participants face.

[7]The fact that participants react differently at baseline to equal-value changes across attributes is also inconsistent with rational inattention models based on Shannon entropy. See footnote 20.

results show that similar results apply even to aspects of decisions that are not, in any straightforward sense, hidden from decision makers. Further, unlike in my experiment, these studies necessarily estimate relative attention to only a small subset (usually a single pair) of a product's relevant attributes (e.g., price vs taxes), and thus they cannot speak to total attention allocation or distraction.

Next, my findings speak to a small literature looking at how information can influence attention (variously defined). For example, Golman et al. (2021) and Quispe-Torreblanca et al. (2022) study how people may avoid information or situations that would require thinking about unpleasant beliefs. Compared to these studies, I employ a distinct notion of attention: the extent to which already-held beliefs about a relevant feature of a decision are employed in a given choice.[8] There are also some examples of information provision experiments with apparent "backlash" effects (e.g., Barrera et al. 2020, Colonnelli et al. 2023, Alesina et al. 2023). In addition to being able to cleanly separate the beliefs and attention channels through which information may affect choices, my results also shed light on how attention must operate for it to underlie such effects. Perhaps relatedly, researchers designing information interventions sometimes worry that providing information about one topic might affect behavior over and above its effect on the targeted belief (see Haaland et al. 2023 for a review). Such effects could arise through belief-based channels (e.g., by shifting beliefs about some unobserved factor or through inferences about what the experimenter knows), but my results suggest that a simple attention story may be a first-order concern, even in experiments with active control groups.[9]

Finally, there is a large body of work in psychology on priming effects, whereby making certain concepts (e.g., religion, identity, or money-making) more salient can affect people's decisions and attitudes (see Cohn & Maréchal 2016 and Dai et al. 2023

[8]Relatedly, Bordalo et al. (2023) show that different but equivalent ways of describing hypotheses have large effects on what features people attend to while solving statistical problems, leading to well known biases in belief-updating.

[9]An active control design, where multiple treatment groups receive differing information, rules out behavior changes being driven only by attention shifts. But my findings that information boosts attention toward the attribute it describes and distracts from others imply that the sensitivity of behavior to the targeted belief is boosted relative to what it would be absent information. Thus, while an active control group prevents attention from affecting the *level* of the dependent variable, it may still affect the estimated *slope* with respect to beliefs. This is particularly relevant when attempting to interpret behavioral responses to information as reflecting underlying preferences (e.g., Wiswall & Zafar 2015) or how behavior would change absent biases in beliefs (e.g., Conlon & Patel 2023).

for meta-analyses). My results show that information aimed at changing behavior by correcting beliefs, because it inevitably involves making salient the dimensions the information describes, has effects that echo parts of this literature. Work on priming in psychology, however, tends not to connect such effects to information provision, beliefs, attentional capacity constraints, or the form that inattention takes (neglect vs priors).

This paper proceeds as follows. Section 2 describes a simple model of inattention-distorted choice to motivate the experimental design. Section 3 explains the experiment in detail, and then Section 4 presents the main results. Section 5 discusses how the results speak to recent theories of attention in economics. Section 6 describes and argues against other interpretations of the main results (e.g., experimenter demand, biased priors), and finally Section 7 concludes.

2 Theoretical Framework

I first describe a simple model to motivate the experimental design. Assume an agent decides whether to purchase a good characterized by a vector of attributes $\vec{a}$ (for example, a car has a price, certain safety features, fuel economy, etc.). Her utility from purchasing the good is linear in these attributes, according to equation 1:

$$v = \sum_{k=1}^{K} v_k a_k \tag{1}$$

I assume that the *perceived* value of the good depends on the agent's beliefs and on which attributes she pays attention to. Let $\theta_k \in [0, 1]$ indicate the extent to which the agent attends to dimension k. To the extent that she does, she incorporates her belief $\widehat{a_k}$ about the level of attribute k, which may differ from its true level a_k. To the extent that she neglects k, she uses a default value $\overline{a_k}$. Combining these assumptions, I define her attention-weighted perceived valuation of the good, which I denote by u, using equation 2:

$$u = \sum_{k=1}^{K} \theta_k v_k \widehat{a_k} + (1 - \theta_k) v_k \overline{a_k} \tag{2}$$

Finally, assume that there is some noise $\epsilon \sim F$ such that she purchases the good whenever $u + \epsilon > 0$. Then the probability that i purchases the good is $F(u)$. For simplicity, I assume that F is uniformly distributed in the relevant range: i.e., $F''(u) = 0$.

Note that this basic formulation is in principle compatible with many different theories of what draws attention. For example, attention could be driven (i.e., θ could be determined) by focusing, relative thinking, salience, memory of the same or similar products/choice scenarios, or noisy information-processing about attributes (Koszegi & Szeidl 2013, Bordalo et al. 2020, 2022; Bordalo, Burro, et al. 2024, Bushong et al. 2021, Gabaix 2019, Yang & Krajbich 2022). In Section 5, I summarize how my results speak to these theories of attention allocation.

My definition of the attention paid to an attribute is also purposely ambivalent about the cognitive process that allows the agent to respond to it. For example, by my definition, an agent could attend to an attribute even without looking at it (e.g., a voice from someone in the next room). They could even attend to an attribute without consciously thinking about it: for example, by my definition, a pedestrian pays attention to the unevenness of a sidewalk because she appropriately (if subconsciously) adjusts her gait to account for it. Although other notions are appropriate for different domains, my definition is appropriate for the research questions I am studying: when do we fail to account for an attribute in such a way that we make the wrong decision? How could a persuader use information a receiver already knows to alter her choices? How should policy-makers design and evaluate the effects of information interventions in light of such effects?

Is attention identifiable from choice data in such a setting? Consider equation 3, which compares how responsive demand for the good is to changes in the level of two attributes k and j.[10]

$$\frac{dF(u)/da_k}{dF(u)/da_j} = \underbrace{\frac{v_k}{v_j}}_{\text{Preferences}} \times \underbrace{\frac{d\widehat{a_k}/da_k}{d\widehat{a_j}/da_j}}_{\text{Beliefs}} \times \underbrace{\frac{\theta_k}{\theta_j}}_{\text{Attention}} \tag{3}$$

As is intuitive, equation 3 makes clear that demand responses depend on how much the agent cares about attributes (preferences), how much she knows about changes in those attributes (beliefs), and how attentive she is to each of them. It also makes clear that if agents' preferences and beliefs are known ex ante, then relative demand

[10]Equation 3 makes a few implicit, but substantive, assumptions. Namely, it assumes that changes in the level of one attribute do not change beliefs about the levels of *other* attributes, the amount of attention paid to any attribute, or the default values $\overline{a_k}$.

shifts are a sufficient statistic for (relative) attention. This is the setting that the experiment described below constructs.

We can now ask what effect information provision will have on agents' choices. To model this, suppose we can employ an information treatment t, which I take to be continuous for illustration purposes (e.g., perhaps t denotes the forcefulness or credibility of the information). Assume for simplicity that t only changes beliefs about one attribute a_k. Equation 4 then follows:

$$\frac{d}{dt}\left[\frac{dF(u)}{da_k}\right] = F'(u)v_k\left(\underbrace{\theta_k\frac{d}{dt}\left[\frac{d\widehat{a}_k}{da_k}\right]}_{\text{Effect on Beliefs}} + \underbrace{\frac{d\widehat{a}_k}{da_k}\frac{d\theta_k}{dt}}_{\text{Effect on Attention}}\right) \tag{4}$$

The first term of equation 4 captures a traditional belief-based persuasion effect: information might increase how responsive demand is to an attribute because it affects how responsive *beliefs* are to that attribute. For example, an agent might switch from being ignorant of the level of k ($\frac{d\widehat{a}_k}{da_k} = 0$) to being fully informed ($\frac{d\widehat{a}_k}{da_k} = 1$). The second term captures an attention effect: information might increase attentiveness to the attribute it describes, over and above any effect on beliefs. Note that in the extreme case where beliefs are already correct and thus do not react to information ($\frac{d}{dt}\left[\frac{d\widehat{a}_k}{da_k}\right] = 0$), choices only react through this attentional channel. In addition, though by assumption information describing k has no effect on beliefs about other attributes, it could affect the attention paid to other attributes. If so, the first term of equation 4 for such attributes would be zero, but the second might not be. My experiment tests for both direct effects of information on the attribute it describes and these indirect effects on other attributes.

In addition to asking how information affects the responsinveness of demand to changes in attributes, we can also ask how it might affect the total level of demand, as given by equation 5:

$$\frac{dF(u)}{dt} = F'(u)\left[\underbrace{\theta_k v_k\frac{d\widehat{a}_k}{dt}}_{\text{Effect on Beliefs}} + \underbrace{\sum_{f=1}^{K} v_f\frac{d\theta_f}{dt}(\widehat{a}_f - \overline{a}_f)}_{\text{Effect on Attention}}\right] \tag{5}$$

Here again we see two terms corresponding respectively to a beliefs channel and an attention channel. First, of course, if the information increases the agent's belief about the level of attribute k (and she has a positive preference v_k for it), this will increase demand. Note that if the agent's beliefs are unbiased and information is randomly assigned, such that on average it does not change her beliefs about the value of k, then information will have no average beliefs-based effect. As we will see, this is the case for some information in my experiment.

The second term of equation 5 says that, even absent an effect on beliefs, information will boost demand if it tends to increase attention to attributes that are higher than their default value $\overline{a}_f$. Thus any such effects depend crucially on how attributes are treated when they are not attended to. Suppose again that information is randomly assigned, such that on average it arrives when attributes are equal to their expected value. If the default is this expected value (as in some models of optimized limited attention, see Gabaix 2019), then such information should have no average effect. If instead inattention takes the form of neglect or downweighting, as in other models (e.g., Koszegi & Szeidl 2013, Bushong et al. 2021, Bordalo et al. 2022), then information can have attentional effects on choices even when it does not focus the agent on especially positive attributes. Simply increasing attention to attributes that are positive *at all* (not necessarily unusually so) can boost demand. The experiment described below tests between these two possibilities.

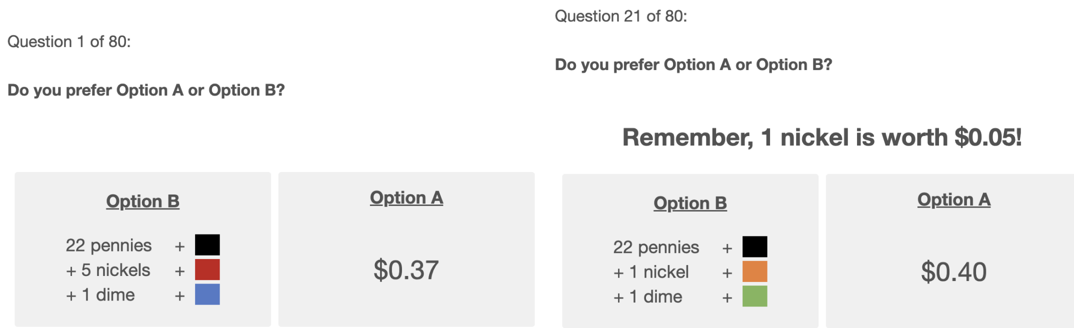
3 Experimental Design

3.1 Option A vs Option B

Participants were recruited through Prolific to participate in an online survey (see Appendix B for details on recruitment, sample, compensation, comprehension checks, and pre-registration). The main part of the experiment asked participants to repeatedly choose between two options, labeled Option A and Option B, for how their bonus payment would be calculated. They made 80 such choices, and one of these was randomly chosen at the end of the experiment to actually determine their bonus payment. The order in which the two options were presented (Option A on the left and Option B on the right, or vice versa) was randomized across participants but held fixed throughout the experiment. Figure 1 shows screenshots of two such choices.

Option A had two "attributes." First, it listed an amount of money that would,

Figure 1: Main Experimental Task: Choosing between Options A and B



Notes: This figure gives two examples of the decision screen participants saw when choosing between Options A and B. The left panel is an example of a decision without any additional information being shown. The right panel shows a decision screen for a participant receiving information about one attribute of Option B (in this case, the number of nickels). The left-right placement of Option A vs B was randomized across participants but was constant throughout the experiment.

with certainty, be added to participants' bonus if they chose this option. The exact amount (though participants were not told this) was chosen independently across choices from a normal distribution with a mean of $0.40 and a standard deviation of $0.20 (with a minimum of $0.00). Second, if they chose Option A in the decision that was randomly selected to be implemented, they also got to roll five virtual six-sided dice. If the sum of these rolls added up to 12 or less, an extra $1.00 or $2.00 (randomized across participants) was added to their bonus. Right after the instructions page that described this lottery to them, participants were asked their belief about the percent chance of winning such a lottery. The average [median] answer was 27% [20%], significantly higher than the true chance ($p < 0.01$), which is approximately 10%.

Option B had six attributes. I will later look at the impact of providing information about one of these attributes on how responsive participants are to them (more details below). They were therefore designed to be as similar to each other as possible while remaining distinct enough to be considered separately. Three of these attributes were listed numbers of coins (pennies, nickels, and dimes), the value of all of which would be added to their bonus if they chose Option B. There were always 2, 12, or 22 pennies; 1, 3, or 5 nickels; and 0, 1, or 2 dimes. Each of these three values was equally likely to be chosen. Notice that the three coins therefore took on almost the same range of monetary values and differed from choice to choice by similar amounts (always 0, 10, or 20 cents different). Further, the participant pool was restricted to people living in the US for whom the value of these coins is familiar (which I confirm below).

The other three attributes of Option B were colored boxes (arranged vertically such that there was a top, middle, and bottom box), each of which could take on one of three colors (a different three colors for each box). At the beginning of the experiment, before any other instructions, participants were asked to rank each set of three colors according to "how much you like them." Whichever color they ranked highest then added \$0.20 to their bonus, the one they ranked second added \$0.10, and the one they ranked last added \$0.00. Participants were told this was how the survey chose these values. They were also truthfully told that each color was equally likely to appear. Notice that all the colored boxes therefore take on a similar range of monetary values as each other and as the coins, and they differ from choice to choice by similar amounts (again, always 0, 10, or 20 cents different). The purpose of assigning the values of the colors according to participants' preferences was to make them easy for participants to remember (which I confirm below).

The values of five of the six attributes of Option B were chosen randomly and independently across each choice, with each of the three possible values being equally likely to be chosen. The sixth attribute, randomly selected, was "frozen" at one particular value for the entire experiment. This value was also equally likely to be any of the three possible values for the attribute, but simply did not vary from choice to choice.

3.2 Information about Attributes

Unknown to participants, the 80 choices were divided into four periods, each of which lasted for 20 choices. Periods differed in whether and what type of information participants in different treatment groups were provided while they made their choices. Table 1 summarizes each period for the four different treatment groups.

During Period 1, participants simply chose between Options A and B, as described above, without receiving any additional information. This period was intended to give participants' experience with the decision environment and expose them to the distribution of each attribute's values. During Period 2 (choices 21 to 40), 80% of respondents (Treatments 2, 3, and 4) began to see information at the top of the

Table 1: Information across Treatment Groups

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
Pre-Decisions	⟵ Identical Instructions ⟶			
Period 1: Decisions 1-20	None	None	None	None
Period 2: Decisions 21-40	None	Target	Target	Target
Period 3: Decisions 41-60	None	None	Target	Alternative
Period 4: Decisions 61-80	⟵ None, Lottery, or Lottery + Odds ⟶			
N	134	239	114	103

Notes: This table describes the distribution of participants across treatment groups and the information presented to each group throughout the experiment. "Target" denotes information about the randomly chosen target attribute of Option B. "Alternative" denotes information about the alternative attribute, which was chosen randomly from the five non-target Option-B attributes. "Lottery" indicates information mentioning the lottery associated with Option A but not its odds. "Lottery + Odds" indicates information that also mentioned the numerical odds of winning Option A's lottery. The first row indicates that instructions were identical to all participants, regardless of treatment group. The row for Period 4 indicates that information about lotteries (or not) were randomly assigned independently of treatment group. The experiment sorted participants into treatment groups at the beginning of the survey, with 20% probability of being assigned to Treatments 1, 3, and 4, and a 40% probability of being assigned to Treatment 2. Variation in sample sizes from this distribution is due to chance.

screen about a randomly selected attribute of Option B (chosen with equal likelihood from among the five non-frozen attributes), which I call the "target attribute." This information told participants how much the target attribute in the current choice was worth. For example, it might read, "Remember, a gold top box adds $0.20!" or "Remember, 12 pennies add $0.12!" This message would change as the value of the target attribute changed across choices.

Participants were explicitly told, directly before beginning to choose between Options A and B, that information like this might appear, that any information they were provided was chosen randomly, and therefore that what the information described "is no more likely to be important to your decision" than aspects it did not describe. They then had to correctly answer a comprehension question confirming their understanding of this fact before continuing with the survey. This explicit (truthful) description of the random assignment of information provision shuts down any potential explanation for treatment effects whereby participants infer an implicit recommendation from certain aspects of the experimental manipulation.

During Period 3 (choices 41 to 60), respondents who received no information in Period 2 (Treatmtent 1) continued to see no information. Among participants who received information in Period 2 (Treatments 2-4), Treatment 2 reverted to seeing no information in Period 3. Treatment 3 continued to see information about the target attribute just as they did in Period 2. Treatment 4 was instead shown information about a new attribute (picked at random from the remaining four non-frozen attributes), which I call the "alternative" attribute.

Period 4 did not provide any information about the attributes of Option B. Rather, and within treatment groups, participants were randomized to either receive no information or to receive one of two messages about the lottery associated with Option A. The first message, which I call "Lottery," simply described the lottery (which, though described in the instructions and comprehension checks, was not mentioned on the later decision screens). The second message, "Lottery + Odds," was almost identical but included the numerical odds of winning. In particular, the "Lottery + Odds" message read "Remember, Option A also comes with a 10% chance to win an additional $1 [or $2] prize!" with smaller text at the bottom of the screen telling them the details of how the lottery worked. The "Lottery" message was identical except "10%" did not appear. Because the lottery did not vary from choice to choice, the message remained unchanged at the top of the screen for the entirety of Period 4.

3.3 Beliefs about Attributes

After making all 80 decisions, participants were asked whether they knew how each possible value of Option B's attributes contributed to their bonus payment. In particular, they were first asked how much money a penny, a nickel, and a dime were worth. Reassuringly, 95% of participants get all of these questions correct. Next, for each of the colored boxes associated with Option B (recall there was a top, middle, and bottom box), they were asked to match each possible color to its monetary value ($0.00, $0.10, $0.20). For each box, between 87 and 89% of participants get all three values correct. Seventy-four percent of participants get all 12 questions (three coins, and three values for each of the three boxes) right. Note that this high accuracy appears despite these questions being unincentivized. We see similar levels of accuracy (indeed slightly higher, 75%) among participants who received no relevant information about the attributes during the 80 choices (Treatment 1). This high level of accuracy likely in part reflects the fact that the mapping between colors and money depended on participants' preferences over colors (and thus participants could reconstruct values by thinking about these preferences).

4 Results

4.1 Empirical Strategy

For most of the results below, I estimate variants of equation 6 by OLS:

$$ChoseOptionB_{i,t} = \beta_0 T_{i,t} + \sum_k \beta_k a_{i,k,t} T_{i,t} + \mu_i + \epsilon_{i,t} \tag{6}$$

In the above equation, $ChoseOptionB_{i,t}$ indicates whether participant i chose Option B in decision t. $T_{i,t}$ indicates some treatment status (e.g., whether/what type of information was visible for i during t). $a_{i,k,t}$ is the monetary value of attribute k in that choice, where there are eight possible attributes: the three coins, three colored boxes, the certain payment in Option A, and the value of the lottery in Option A. I define $a_{i,k,t}$ for each attribute such that β_k should be positive for all of them (i.e., I multiply Option A's attributes by negative one). I also scale variables such that coefficients can be interpreted as the effect of increasing the value of an attribute by \$0.10. In practice, I often add multiple attributes together (e.g., the value of all the colored boxes, or all non-frozen Option B attributes) to increase power and interpretability. I also typically recenter all attribute values such that they have mean zero. Finally, I also usually include individual-fixed effects μ_i, which could represent person-specific biases toward one or the other option, or heterogeneity in default values $\overline{a_k}$.

Estimating equation 6 lets us answer several questions. First, it tells us how responsive demand for Option B is to the values of various attributes within any given treatment (i.e., any information environment). Recall from equation 3 that responsiveness to attributes reveals a combination of 1) preferences over attributes, 2) beliefs about changes in the attributes, and 3) relative attention to attributes. By design, the experiment fixes preferences over attributes, as each attribute of Option B is straightforwardly worth a certain amount of money.[11] In addition, as described in

[11] In principle, the presence of the lottery for Option A raises the possibility that risk aversion could affect how participants are willing to trade-off between Options A and B. In practice, as we will see, participants are *more* responsive to the certain payment in Option A than to Option

Section 3.3, I measure participants' beliefs about attributes. I can thus ask whether differential responsiveness to attributes is due to differential beliefs or instead to inattention. Similarly, I can ask whether changes in this responsiveness across treatment groups are due either to treatment effects on beliefs or on attention, as in equation 4.

Next, the first term in equation 6 tells us how demand responds on average to information about an attribute, pooling across the particular values that attribute takes on. Recall from equation 5 that this effect depends both on whether information shifts attention and on what the "default" value is (i.e., how an attribute is treated when it is neglected).

Note that equation 6 has a clear "rational" benchmark: if agents' pay equal attention to all attributes, and if they have correct beliefs about those attributes, then demand responses should all be equal (i.e., $\beta_k = \beta_j$ for all k and j). Further, information should have no effect ($\beta_0 = 0$ and β_k should not depend on treatment). For example, estimating this regression using the expected-payoff maximizing choice as the dependent variable (as opposed to participants' actual choices) yields a coefficient of approximately 0.12 for all attributes. That is, increasing an attribute's value by \$0.10 increases the chance that its option is the payoff-maximizing choice by 12 percentage points on average.

4.2 Baseline Attention

Is attention systematically distorted at baseline, i.e., in Period 1 before participants received additional information about any attribute? The left panel of Figure 2 (and column 1 of Table A.I) show estimates of equation 6 (without individual-fixed effects or treatment dummies) for four attributes: Option A's certain monetary value, the subjective value of Option A's lottery, Option B's coins, and Option B's colored boxes. To calculate each participants' subjective value of Option A's lottery, I multiply its monetary prize by participants' stated priors about their odds of winning.[12]

We see large differences across attributes in this average measure of attentiveness. Compared to Option B's colored boxes, participants are 113% more attentive to the value of Option B's coins and 180% more attentive to Option A's certain value.

B's attributes, the opposite of what we would expect if risk aversion were a substantial factor in participants' decisions. Since this issue would not substantially affect interpretation any of the main results, I assume participants are risk-neutral.

[12]In the main text I use a simple linear probability model, since it makes the fewest assumptions. But Table A.II shows qualitatively identical patterns when I instead estimate a logit regression.

They appear least attentive to Option A's lottery (which, recall, was not visually represented on each decision screen), as the coefficient on it is only 31% of that on Option B's colored boxes.[13] For each pair of attributes, we can reject that the responsiveness of demand is equal ($p < 0.01$ for all pairwise comparisons).

By construction, these differences cannot be due to differences in how participants actually value these attributes. Some differences could, however, be due to misperceptions about how each attribute would contribute to their bonus payment (e.g., not remembering what each colored box is worth). To explore this possibility, the light blue bars in Figure 2 (and column 4 of Table A.I) restricts the analysis to the "correct-beliefs sample," the 74% of participants who correctly reported, in the unincentivized questions at the end of the experiment, the value of each type of coin and the values of each possible colored box.[14] We see similar estimates to those from the full sample: these particpants are 147% more responsive to Option A's certain value, 92% more responsive to Option B's coins, and 69% less responsive to Option A's lottery than they are to Option B's colored boxes. Again, we can reject equality of responsiveness for each pair of attributes ($p < 0.01$). Thus, these differences appear to be driven by selective attention, rather than by mistaken beliefs.

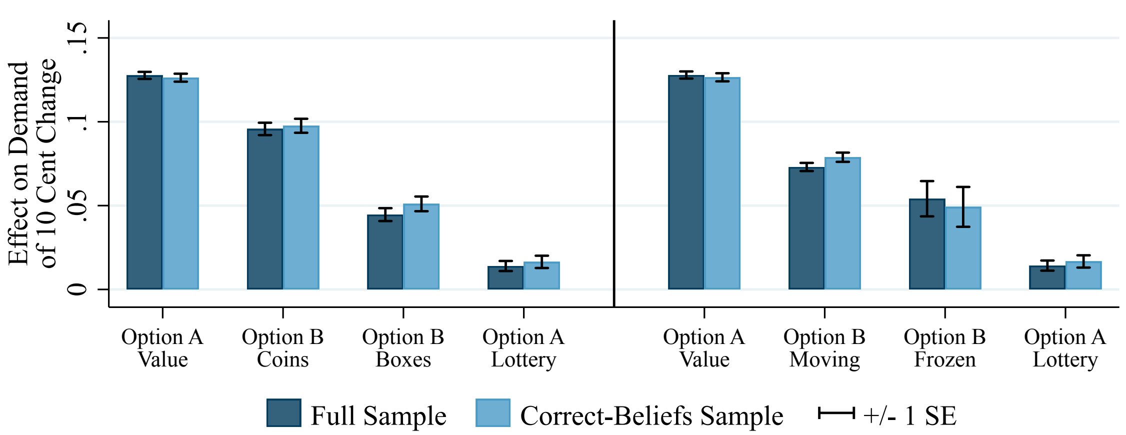
A natural question is whether these attentional differences arise due to differences in their associated payoffs within a given decision. Option A's certain value tended to be larger than that of other attributes; if people pay more attention to attributes during decisions in which they take on more extreme values, that might explain why on average Option A's certain value draws more attention. Columns 2 and 5 of Table A.I explore this possibility by adding quadratic terms to the regression specification.[15]

[13] This coefficient could suffer from attenuation bias to the extent that participants' reported beliefs about the lottery are noisy. However, if I instead simply use the objective value of the lottery, whose prize was randomized to be \$1.00 or \$2.00, I find that this has no significant effect on whether participants choose Option B and can reject responsiveness equal to even that of the colored boxes ($p = 0.01$, results available upon request). This estimate does not suffer from attenuation bias (since it does not incorporate participants' potentially noisily measured beliefs). Also notice that because participants on average greatly overestimate the probability of winning the lottery, with full attention they "should" react more to increases in the objective value of the lottery than to increasing the value of the colored boxes by an equivalent amount. Thus, my finding that participants react least to the lottery does not seem to be merely a product of attenuation bias.

[14] Some of these participants, between Period 1 and belief elicitation, saw information telling them about one (Treatments 2 and 3) or two (Treatment 4) attributes of Option B. Table A.III shows similar results even for participants in Treatment 1, who were never provided information about Option B's attributes.

[15] For Option B's coins and colored boxes, I square each component attribute (e.g., the value of the dimes or the middle box) and then add up these squared values. This allows me to test whether

Figure 2: Attention at Baseline



Notes: This figure depicts a subset of the OLS estimates from Table A.I, which estimates equation 6 using the Period 1 decisions of all treatment groups. The dependent variable is whether the participant chose Option B. The independent variables in the left-hand panel are the certain value of Option A, the subjective value of Option A's lottery, the sum of Option B's coins, and the sum of Option B's colored boxes. The subjective value of the lottery is calculated by multiplying the prize for winning the lottery with each participants' prior belief about their odds of winning (winsorized at the 90th percentile). The independent variables in the right-hand panel are identical except the Option-B attributes instead include the sum of the five changing attributes and the attribute that was frozen throughout the experiment at its initial value. Dark blue bars show estimates including all participants, while the light blue bars show estimates including only participants who correctly respond to unincentivized questions at the end of the experiment about the monetary value of each coin and every possible colored box. Whiskers show robust standard errors, clustered at the individual level. Table A.I shows the full regression results for these specifications.

Though the estimates are not always statistically significant, the coefficients on the squared terms show if anything the opposite pattern. When Option B's attributes are more valuable, their marginal contribution to choosing Option B is smaller (negative coefficients). Conversely, when Option A's attributes are more valuable, their marginal contribution to choosing option A is smaller (positive coefficient).[16] This result suggests that if anything there is diminishing sensitivity to the value of attributes, so range effects cannot explain the average differences between them.[17] Further, note that Option B's coins and colored boxes have almost identical ranges of values, so the difference between them cannot not be due to such effects.

Next, recall that a randomly chosen attribute of Option B is frozen at its initial value throughout the whole experiment. The right panel of Figure 2 (and Column 2 of Table A.I) show that participants are 26% less responsive to this attribute than to the attributes that are changing for them from decision to decision ($p = 0.08$). We see similar relative inattention among the correct-beliefs sample (38%, $p = 0.02$). I return to these facts in Section 5 when I discuss implications of my findings for models of attention.

4.3 Responsiveness to Information

I now turn to the main question of how providing information about Option B's attributes affects participants' choices. Given that the large majority of respondents are able to recall how each attribute's possible values contribute to the total value of Option B, and that they know such information is randomly assigned, it might be natural to think that it should have little effect. However, recall from equations 4 and 5 the possibility that information might affect choices even without any effect on beliefs to the extent that it changes the relative attention that agents pay to different attributes.

extreme individual attributes draw more attention. This is equivalent to regressing choices on the value of each individual attribute and its squared value but with the restriction that there be equal coefficients across coins and across colored boxes.

[16]The fact that positive coefficients on Option-A attributes implies less sensitivity to more extreme values can be understood as follows. As the square of Option A's certain value gets larger, this boosts demand for Option B. Thus, though the main effect of making Option A more valuable of course reduces demand for Option B, this marginal effect favors B more as the level gets larger.

[17]In addition, unlike in the analyses to follow, Table A.I does not recenter each attribute to have mean zero. Thus, the coefficients on the main effects can be interpreted as the marginal effect of increasing each attribute from a value of zero. These main effects look qualitatively similar those in column 1 where quadratic terms were excluded.

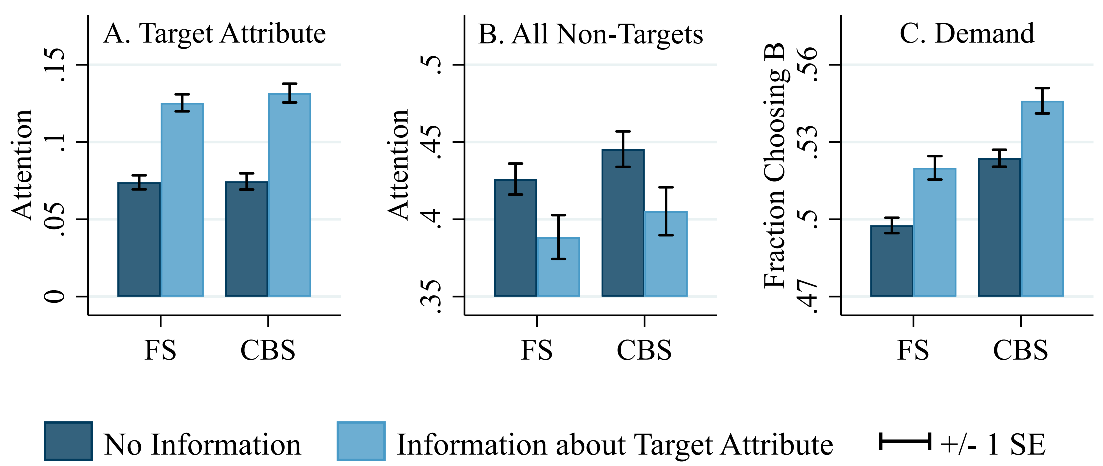
To investigate this possibility, we can compare participants in treatment groups 2-4, who received information in Period 2 about a randomly selected attribute of Option B, to those in Treatment 1, who continued to see no additional information. In practice, the experiment implemented these treatments by choosing, for each person regardless of treatment group, a random attribute of Option B to be the "target" attribute. In Treatments 2-4, participants then began to receive information about this attribute in Period 2. I can therefore compare responsiveness to this target attribute depending on whether participants were (Treatments 2-4) or were not (Treatment 1) receiving information about it.

Figure 3 summarizes OLS estimates of equation 6 (the full estimates are reported in Table A.IV), pooling data from all participants for Periods 1 and 2. In these regressions, the included attributes are the certain value of Option A, the target attribute of Option B, and the non-target attributes of Option B (all summed together). I interact these attributes with a dummy for receiving information about the target attribute (i.e., being in Treatments 2-4 during period 2). We see in the leftmost pair of bars in Panel A (and in column 1 of Table A.IV) that the information has a large effect on responsiveness to the target attribute, increasing by 69% the attention participants pay to it (from 0.074 to 0.125, $p < 0.01$). We see similarly sized effects if we restrict the sample to the "Correct-Beliefs Sample" (right pair of bars in Panel A, and column 4 of Table A.IV), the respondents who correctly identify how each attribute of Option B contributes to their bonus in the unincentivized questions at the end of the experiment. This result is consistent with the information primarily operating by changing how much attention participants pay to different attributes, rather than through its effect on their beliefs.

Columns 2-3 and 5-6 of Table A.IV split the sample by whether the target attribute was a coin or colored box. We see that effects are larger for information about the colored boxes ($p < 0.05$ for both the full and correct-beliefs samples), suggesting that attention effects are larger when baseline attention is lower. However, even information about coins significantly affects the difference in responsiveness to the target attribute ($p < 0.05$ for both samples). Because it is obvious (and the later unincentivized questions confirm) that participants are already perfectly aware that, say, two dimes are worth $0.20, the most natural interpretation of these results is that the information persuades primarily by shifting attention.

These results show that information boosts the attention participants pay to the

Figure 3: Effects of Information about Target Attribute



Notes: This figure summarizes OLS estimates of equation 6 using the Periods 1 and 2 decisions of all treatment groups. The dependent variable is whether the participant chose Option B. The independent variables are the attributes of Options A and B interacted with a dummy variable for receiving information about the target attribute (i.e., being in Treatments 2-4 during Period 2). Panel A shows the coefficient on the target attribute. Panel B shows the sum of attention to all non-target attributes. Panel C shows the fraction choosing Option B depending on treatment, controlling for the other variables in the regressions. The left-hand and right-hand pairs of bars within each panel shows estimates from the full sample (FS) and correct-beliefs samples (CBS), respectively. Whiskers show robust standard errors, clustered at the individual level. See Table A.IV for the full regression results and more details on the specification.

attribute it describes. Does it also have effects on other attributes? Panel B of Figure 3 shows the sum of the coefficients on all other attributes for participants who are and are not receiving information about the target attribute. We see that this measure of the total attention paid toward the other attributes in the decision significantly declines when information directs attention toward the target attribute ($p < 0.05$ for both the full and correct-beliefs samples). In percent terms, this decline is smaller (about 10%) than the direct effect on the target attribute, but it applies to many more attributes. Thus in absolute terms the total spillover effect is comparable to the effect on the target attribute (-0.038 compared to 0.051). The effect on "total" attention, summing the effects on the target and non-target attributes, is therefore statistically indistinguishable from zero ($p = 0.43$).

The effects described so far concern how information affects the responsiveness of demand to various attributes. Panel C of Figure 3 shows that information about the target attribute also had a significantly positive average effect on demand for Option B (by 2.2pp in both the full and correct-beliefs samples, $p < 0.01$ for both comparisons).[18] These effects appear despite the information being uncorrelated with the value of the attribute it described. That is, on average the treatment provided neutral information (which, in any case, most participants already knew) about the value of the target attribute. This result is consistent with the default value—how an attribute is treated when it is not attended to—being *zero* rather than agents' priors about its expected value.[19] Consistent with this effect operating through attention, these effects are larger for information about colored boxes than about coins ($p = 0.05$ and $p = 0.10$ for the full and correct-beliefs samples), corresponding to the larger boost in attention for these attributes compared to coins.

4.4 Dynamics of Attentional Effects

How stable are these attentional effects? Figure 4 and Table A.V show estimates of equation 6 using data from only Period 3, where the four attributes are the certain value of Option A, the target attribute of Option B, the "alternative" attribute of Option B (which Treatment 4 sees information about in Period 3), and the remaining

[18]The values of each attribute in Table A.IV are recentered to have mean zero, so the main effect of information can be interpreted as the effect at the mean of these values.

[19]Note that, by this point in the experiment, all participants had made at least 20 previous choices, and so had experience with the distribution of values that each attribute took on.

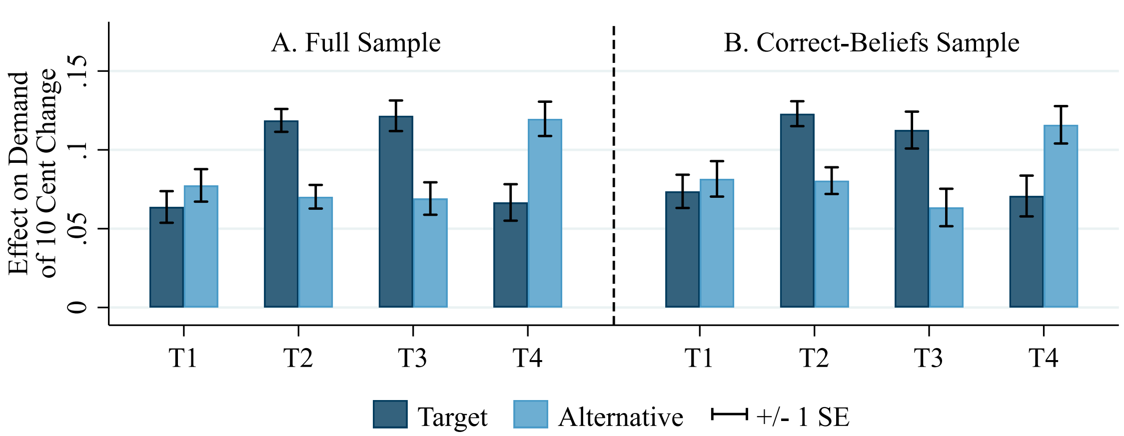
attributes of Option B (all summed together). I estimate this regression separately for each treatment group. We see that having received information about the target attribute still has a significant effect on attention paid to it, even when this information is no longer visible (Treatments 2 vs 1, 0.119 vs 0.064, $p < 0.01$). This increased attention is almost identical to the attention paid to the target attribute when the information is still visible (Treatments 2 vs 3, 0.119 vs 0.122, $p = 0.81$). In contrast, when information about a new attribute begins appearing (Treatment 4), the effect of the previous information disappears entirely: attention to the target attribute reverts to a very similar level as if participants had never received the information (0.064 vs 0.067 in Treatments 1 vs 4, $p = 0.55$) and much less than if the new information had not appeared (0.119 vs 0.067 in Treatments 2 vs 4, $p < 0.01$). We also see a large and significant effect on attention paid to the alternative attribute in Treatment 4 (0.120 vs 0.077 in Treatments 4 vs 1, $p < 0.01$), as expected.

Taken together, these results suggest that, while information can have large effects on attention and that these effects can outlast the information itself, they are also quite fragile. I interpret this result as another manifestation of distraction: just as information boosts attention to the attribute it describes by decreasing focus on other attributes, so the attentional impact of new information comes at the expense of the effect of previous information.

4.5 Implications for Information Correcting Beliefs

Thus far, I have focused on the effect of information in cases where it communicates only what decision makers already know. Clearly, this is a special case, as information is often provided at least in part with the aim of changing recipients' beliefs. The results from the previous sections, however, suggest that such information can have paradoxical effects. For example, suppose an agent is neglecting (i.e., her default is zero) an attribute she has a positive preference for. She then receives *bad news* about that attribute: its value is less than her prior. She might nonetheless act as though this information is good news, increasing her demand for its associated good. The explanation is that the information may increase her attention toward the attribute, boosting the value she treats it as having from the default of zero to a positive (if lower than expected) level. Such paradoxical effects would have implications both for how to interpret responses to information interventions and for what sorts of information

Figure 4: Dynamics of Information Effects



Notes: This figure depicts a subset of the OLS estimates from Table A.V, which estimates equation 6 using the Period 3 decisions of each treatment group (separately). The dependent variable is whether the participant chose Option B. The independent variables are the certain value of Option A, the target attribute of Option B, the alternative attribute of Option B, and the sum of the four other attributes of Option B. Treatment 1 (T1) never received information about the target attribute. Treatments 2-4 (T2-T4) received information about the target attribute in Period 2, but differed in the information presented during Period 3. During this period, Treatment 2 received no information, Treatment 3 continued to receive information about the target attribute, and Treatment 4 received information about the alternative attribute (randomly chosen from the remaining four non-frozen attributes of Option B). Panel A shows estimates including all participants, while Panel B shows estimates including only participants who correctly respond to unincentivized questions at the end of the experiment about the monetary value of each coin and every possible colored box. Whiskers show robust standard errors, clustered at the individual level. This figure shows only the coefficients on the target attribute (dark blue bars) and alternative attribute (light blue bars). Table A.V shows the full regression results for these specifications.

persuaders might be willing to divulge.

In this section, I explore this possibility by looking at Period 4 of the experiment. Recall that during Period 4, participants were randomly (and independently of their treatment group) sorted into three groups. A third of participants saw no additional information in Period 4. Another third were shown the "Lottery" message, which informed them that Option A also came with a lottery that added \$1 or \$2 to their bonus if their roll of five 6-sided dice add up to 12 or less. The final third of participants received the almost identical "Lottery + Odds" message, which additionally included the fact that such a lottery pays off about 10% of the time, much lower than participants' priors (mean 27%, median 20%, both significantly different from 10% at $p < 0.01$).

What effect should we expect the "Lottery + Odds" message to have on demand for Option A? As shown in equation 5, there may be two competing effects. First, in one sense this information clearly conveys bad news about Option A: it should reduce participants' beliefs about the value of the lottery, and hence of Option A. But second, if participants would otherwise fail to pay attention to the lottery, the information's attentional effect depends on the lottery's default value $\overline{a}$. The results in 4.3 suggested that this default is zero. If so, then boosting attention toward the lottery could nonetheless boost demand for Option A by pointing attention toward one of its positive attributes (even if it is not as positive as participants' priors suggested).

Table 2 shows OLS estimates where the dependent variable is whether participants chose Option A (which included the lottery), pooling all decisions across all periods of the experiment. I regress this variable on the certain value of Option A, the total value of all six of Option B's attributes, individual-fixed effects, an indicator variable for whether participants saw the information about the lottery that did not include odds, and a similar indicator for the information that also included the odds of winning. In column 1, we see that the "Lottery + Odds" message significantly *boosted* demand for Option A by 5.1pp, ($p < 0.01$), despite (in a sense) delivering bad news about it for the average participant.

We can disentangle the beliefs and attention effects of the "Info + Odds" treatment by comparing it to the "Info" treatment, which did not change beliefs about the lottery's odds. Column 1 of Table 2 shows that this message boosted demand for Option A by even more (9.1pp, $p < 0.01$). This result suggests that the "pure" attention effect is quite large, enough to countervail the significant beliefs effect of

Table 2: Effect of Information that Changes Both Beliefs and Attention

	(1)	(2)
Lottery Info without Odds	0.091***	0.077***
	(0.013)	(0.017)
Lottery Info including Odds	0.051***	0.085***
	(0.013)	(0.017)
Lottery Info without Odds X Error in Prior		0.099
		(0.076)
Lottery Info including Odds X Error in Prior		-0.217***
		(0.074)
Option A Value	0.128***	0.128***
	(0.002)	(0.002)
Option B Total Value	-0.076***	-0.076***
	(0.002)	(0.002)
Observations	47,200	47,200
Individuals	590	590
R^2	0.50	0.50
p-value: Main Effects of Information Equal	0.03	0.72
p-value: Interactions Equal		0.00

Notes: This table shows OLS regression estimates, pooling data from all Treatments and all Periods of the experiment. The dependent variable is a dummy indicating whether the participant chose Option A (which included the lottery). I regress this variable on the certain value of Option A ("Option A Value"), the sum of all six of Option B's attributes ("Option B Total Value"), individual-fixed effects, and dummy variables indicating whether the participant was shown information mentioning the lottery, where this information either came without numerical information about the odds of winning ("Lottery Info without Odds") or with such information ("Lottery Info with Odds"). In Column 2, I additionally interact these dummy variables with the error (belief minus truth) in participants' previously reported beliefs about the odds of winning the lottery (winsorized at the 90th percentile). Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

4.0pp (9.1pp minus 5.1pp, $p = 0.03$).

In column 2, I additionally interact these indicators with the error (beliefs minus truth) in participants' priors about the odds of winning the lottery. For participants who do not receive information about the odds of winning, we see directionally larger effects for respondents who overestimated the odds of winning by more ($p = 0.20$). In contrast, for participants who also learned the true odds of winning, the interaction term is negative and statistically significant ($p < 0.01$), as we would expect from the information correcting misperceptions. The main effects of both interventions—the effect for participants whose beliefs are correct, which we can interpret as the pure attentional effect—are large and positive (around 8 percentage points, $p < 0.01$, for both), and statistically indistinguishable from each other ($p = 0.72$). Note also that the fact that the "Lottery + Odds" message boosted demand for Option A even for participants whose beliefs were already correct (the interacted main effect in Table 2) is particularly stark evidence that participants' default is less than their priors, consistent again with inattention taking the form of neglect. Similarly, the fact that the simple "Lottery" message boosted demand for Option A without altering beliefs about the odds of winning is again inconsistent with participants' default being their priors.

4.6 Analyzing Open-Ended Text Responses

At the very end of the experiment, participants were asked to describe in their own words how they tended to choose between Options A and B. I then have a large language model categorize their responses to shed some (suggestive) light on participants' choice processes (see Appendix B for details on the question and categorization procedure). The average (median) respondent wrote 23 (18) words in response to this prompt, so these answers provide short summaries of what strategies participants felt they were employing. However, as with many (unincentivized) open-ended questions, the level of engagement from respondents shows considerable variation: 22% of participants write less than ten words.

I find, reassuringly, that the large majority of participants (66.9%) mention attempting to add up the monetary value of each option before choosing. At the same time, many participants are aware that they are doing so only imperfectly, with 42.5% mentioning that they were merely trying to identify or guessing (rather than knowing) what the right option to choose was. Perhaps surprisingly given the large effect that they had on participants' choices, only 1.2% (7 participants) mention the information treatments at all in these open-ended questions. Finally, no participant makes any mention of reacting to what they perceived the experimenter to want them to do, or of interpreting the information provided as an implicit suggestion of the right course of action, facts I return to in Section 6.

5 Discussion: Implications for Models of Attention

Taken together, what do these results imply for our understanding of how attention is allocated? Most existing theories in economics model attention as a function solely of objective payoffs. For example, several recent theories posit a role for the range of payoffs across attributes within a given choice. We saw in Section 4.2 some evidence for such a channel: participants pay somewhat less attention to attributes when they take on larger values, consistent with relative thinking (Bushong et al. 2021) or diminishing sensitivity (Bordalo et al. 2013) though not with focusing (Koszegi & Szeidl 2013). But these forces appeared modest compared to the other effects in the experiment.

In addition, few of the results described in Section 4 are predicted by standard models of rational inattention. First, the fact that inattention takes the form of neglect challenges the contention that, conditional on what agents attend to, they processes information rationally. Second, the workhorse model of rational inattention (see Mackowiak et al. 2022 for a review) posits that Shannon entropy governs the costs of attention. An implication of this assumption is that the relative probability of different choices should respond only to differences in their objective payoffs, which is inconsistent with large differences I find in responsiveness at baseline to equal-value changes across attributes.[20] Gabaix (2014) assumes that agents allocate their limited

[20] More precisely, the following equation should hold:

$$\ln P(B|\omega_1) - \ln P(B|\omega_2) = \frac{1}{k}\Big[u(B|\omega_1) - u(B|\omega_2)\Big]$$

where ω_1 and ω_2 are two states (i.e., realizations of the attributes of Options A and B), $P(B|\omega)$ is the probability that Option B is chosen given ω, $u(B|\omega)$ is the payoff from choosing B given ω, and k is the cost of attention. By letting ω_1 and ω_2 be states that differ only along one attribute, we can see that equal-value changes in any attribute should yield identical differences in choice probabilities. See Dean & Neligh (2023).

attention toward attributes that are more important for their choice. Consistent with this idea, participants reacted most to Option A's certain value, which had the largest variance and therefore was most often pivotal. On the other hand, we saw significant differences in attention to Option B's coins compared to its colored boxes, despite (by design) extremely similar distributions of payoffs across these two types of attributes and therefore no difference in their importance. Next, the sensitivity of attention to obviously redundant information is not predicted by any of these models, unless they are added as separate assumptions about the cognitive costs of accessing information (which I discuss more below).

Instead, many of my experimental results point toward the need to incorporate contextual factors as important determinants of attention. An example of such a model is salience theory (Bordalo et al. 2022), in which limited attention is drawn bottom-up to features of the environment that are prominent or surprising. Two aspects of my results resonate with this framework. First, and most obviously, the fact that clearly irrelevant information ("12 pennies are worth $0.12!") has such large effects on attention highlights the importance of the immediate choice environment and which things are made visually prominent within it. Baseline attention toward the attributes in my experiment also corroborate this channel: the feature drawing the most attention (Option A's certain value) was also the one displayed separately and in larger font, while the attribute that most escaped attention (Option A's lottery) was the only one not explicitly depicted on every decision screen.[21]

Second, past contexts and choices have significant lasting effects on attention. For example, having directed attention toward an attribute in the past by providing information continues to dramatically boost attention toward that attribute in the future (so long as no new information about a different attribute takes its place). More speculatively, the fact that coins (which participants have much experience attending to) draw more attention than colored boxes (which they do not) is consistent with a similar channel. Additionally, attributes draw attention in part when they are surprising, in that they take on different values than what the agent has experienced

[21]The importance of visual prominence also brings to mind theories like the attentional drift-diffusion model (e.g., Yang & Krajbich 2022), where gaze modulates cognitive accumulation of evidence from different aspects of the decision. Some of my results, like those on distraction and neglect, are consistent with this framework. However, these theories tend not to model the agent's decision of where to look, and so to that extent they cannot explain why attention shifts in the way it does in my experiment.

in the past, as we saw by comparing attention toward attributes that change vs are frozen throughout the experiment.[22]

One natural way to reconcile some of my results with models in the spirit of the rational inattention literature would be to postulate an attention-cost function that depends on more than Shannon entropy or objective stakes (e.g., see Hébert & Woodford 2021). For example, perhaps certain attributes are simply cognitively cheaper to assess (e.g., depending on how they are visually displayed, whether information is describing them, or whether similar types of information have been considered in the past). Many participants, after all, were aware that they were only imperfectly attempting to choose the highest-value option (see Section 4.6), an attitude very much in the spirit of costly information processing models and one not typically emphasized in so-called "bottom-up" theories of attention. I view incorporating (perhaps salience-driven) contextual factors into such models as providing a way of combining attractive elements of both bottom-up and top-down theories of attention. Further, even if some of the results of my experiment are influenced by non-payoff-based attention costs, this would not undermine any of the implications these results have for persuasion or for the design and interpretation of information interventions.

6 Ruling out Potential Confounds

To summarize the results in Section 4, we saw evidence for five main results. First, at baseline (i.e., absent any additional information), attention across attributes deviates from a rational full-attention benchmark. Participants are less attentive to the colors and coins of Option B than to the certain value of Option A, and they are least attentive to Option A's lottery. Second, providing information (even information that participants already know) boosts attention paid to the attribute it describes. Third, this boosted attention to the described attribute comes at the cost of distracting attention away from other attributes. Fourth, though these attention effects are large, they are also fragile: switching to providing information about a new attribute completely undoes the effect on previously described attributes. Finally, randomly assigned information about one of Option B's attributes boosts demand for that

[22] In addition to giving a key role to prominence and the role of past experiences, salience theory also typically assumes that attention toward one feature is decreasing in the salience of other features (distraction) and that the default value absent attention is zero (neglect).

option, implying that the "default" value for attributes is below participants' priors about them, consistent with inattention as "neglect". This latter finding suggests that even bad news about an attribute (relative to participants' priors) can nonetheless boost demand for its associated option, an implication I test and find support for in Section 4.5.

In this section, I consider potential confounds to my interpretation of these results and describe how the design and findings either rule them out or render them unlikely.

Pessimistic Priors

One might worry that incorrect priors about the distribution of attribute values might drive some of these results. For example, suppose a participant believes that one of Option B's attributes has a lower average value than it in fact does. Suppose further that she only incorporates the value of that attribute into her choice when information about it is being provided, and that otherwise she relies on her biased prior. If so, she will tend to react to accurate information about it by increasing her demand for Option B. She might do so even if the default value she uses absent attention is equal to her prior. Thus, a positive average effect of information on demand for Option B would not be a result of inattention taking the form of neglect.

Several aspects of the experimental design and results rule this explanation out. First, for some attributes, participants are explicitly told the correct priors about the distribution of values. In particular, the instructions told participants that each color (for each of the colored boxes) was equally likely to be chosen. Despite this, information about these attributes boosted demand for Option B more than information about other attributes (Table A.IV, columns 2-3 and 5-6), the opposite of what we would expect if pessimistic priors were driving these responses. Second, participants had extensive experience in the first period of the experiment (20 decisions with randomly generated attribute values) during which they encountered the true distribution and could update any incorrect priors. It is only after this initial period that the experiment begins to provide information, and thus pessimistic priors would have to persist despite repeated feedback.

Finally, the results in Section 4.5 provide particularly clear evidence against the view that participants' default value is their prior. In particular, we saw that information describing Option A's lottery boosted demand for Option A *even among those who were not told the true odds of winning* (Table 2, column 1). Such participants must have relied on their prior about the lottery's odds, and so if the default were their prior this treatment should have had no effect. In addition, we also saw positive effects on demand for Option A of information describing the odds of winning *even among participants who already had correct priors* (Table 2, column 2). If participants were already relying on their prior, information confirming this belief would have had no effect.

Information as an Implicit Suggestion

Next, a potential worry is that participants interpreted the information, despite it containing no wording suggesting they ought to, as an implicit recommendation of the action they should take or of the attributes that happen to matter more in the decision. This possibility is ruled out by the design of the experiment, as participants are explicitly told that any information is about randomly selected attributes and therefore not a signal of what they should react to. A comprehension question immediately prior to the beginning of the choice task required participants to confirm their understanding of this fact, including the implication that information should not alter their beliefs about which attributes are most important. The fact that almost no participants (1.2%) mentioned the information, when asked at the end of the experiment how they decided between Options and B (see Section 4.6), is consistent with this understanding that the information did not serve as an implicit endorsement of any action.

Experimenter Demand

Next, one might worry that the results could be driven by experimenter demand: i.e., that participants might attempt to anticipate the researcher's hypotheses and then modify their behavior to confirm these hypotheses out of sense of altruism. Several observations speak against this concern. First, there are a fairly large number of results consistent with some (but not all) behavioral-economic theories that participants would need to anticipate and then conform their behavior to. Some of these are subtle enough that participants' anticipating them seems quite implausible. For example, participants' boost their responsiveness to attributes the information describes *and* reduce their responsiveness to attributes it does *not* describe (Figure 3); participants' attention to described attributes reverts to its no-information baseline level *only* when it is replaced by information about a new attribute (Figure 4); and participants' increase their demand for options whose attributes the information describes *unless* their priors about that attribute are sufficiently pessimistic (Table 2).

Second, and more fundamentally, there is no empirical evidence that experimenter demand is a salient concern in economics lab experiments. De Quidt et al. (2018) find that explicitly asking participants to change their behavior as a favor to the experimenters has effects around 0.13 standard deviations across 13 different standard experimental tasks. Similarly, Winichakul et al. (2024) study four decision domains among three different subject pools and find no evidence that even the strongest experimenter-demand treatments change behavior enough to spuriously produce large effects. Crucially, these studies estimate *upper bounds* on experimenter demand effects, where the change in behavior required to benefit the experimenter is explicitly described and requested. Effects in studies such as mine, where participants would also need to formulate beliefs about the hypotheses that the experimenters are interested in—and independently decide that conforming to these hypotheses would constitute a favor to the experimenter—are likely to be yet smaller. The fact that no participant mentions anything resembling experimenter demand effects in their description of how they made choices (see Section 4.6) further argues against this concern.

That said, I of course cannot fully rule out that participants in my study might have managed to anticipate the hypotheses being tested, (spuriously) adjusted their behavior to an unprecedentedly large extent, and avoided mentioning this at the end of the experiment.

Risk Aversion and Caution

Next, a reasonable concern is that, if participants are uncertain about the value of various attributes, some results might be driven by small-stakes risk-aversion or caution à la Cerreia-Vioglio et al. (2015). For example, if participants cannot remember exactly how much money a red box or a nickel is worth, risk aversion or caution might lead them both to react less to its true value and to shy away from options that include more subjective uncertainty. Several results point against this interpretation. First, the value of some attributes is straightforward (pennies, nickels, dimes). Second, all my results hold even among the correct-beliefs sample who accurately reports the value of all attributes.[23] Third, and most tellingly, notice that if uncertainty (plus either risk aversion or caution) were driving the main effects, we would expect that information about the target attribute should continue to boost responsiveness to it even when information about a new attribute is provided in Period 3 (Treatment 4): presumably, uncertainty about the target attribute would be eliminated after 20 choices where its value is prominently displayed. Instead we see responsiveness to the target attribute revert to its baseline level, consistent with an attention interpretation.

7 Conclusion

In this paper, I show experimentally that information affects choices by shifting attention, and that these effects can be large enough to overturn the traditional beliefs-based channel of persuasion. Additional aspects of my results may have further implications for persuasion, both for well-meaning policy makers and profit-minded firms. First, these attentional shifts appear despite the information in my experiment often being transparently unhelpful, suggesting that in other contexts salient messages may redirect attention even when they are (recognized as) redundant or manipulative.

Second, I find that attention is capacity constrained: shifting focus to one attribute distracts from others. This dynamic may have implications for policymakers hoping to improve decision-making by providing information or reminders: the net benefit of such interventions will depend on what attributes or considerations such messages crowd out. And firms may find attentional manipulations particularly valuable if they also can use them to "signal jam," simulataneously boosting attention toward their products' positive features and distracting away from their competitors' advantages.

Next, I find that attention is fragile. Information starkly shifts what agents attend to, but new information can quickly undo such effects. This suggests that to be effective attention-boosting policies need to operate close to the moment where a relevant decision is being made. An open question is what factors contribute to the longevity of attention effects, and how these forces interact with similar fragility in

[23] Of course, strictly speaking participants could be uncertain about attribute values but nonetheless manage to correctly guess their values in my beliefs questions. Even in this case, we should expect results to be *weaker* for the correct-beliefs sample, under the intuitive assumption that people who report correct beliefs are less uncertain than those who report incorrect beliefs. This is not what I find.

belief updating (e.g., see Graeber et al. 2022).

Finally, my results suggest that inattention takes the form of *neglect*: features that escape attention appear to drop out of agents' decision procedures entirely, rather than simply being treated as having some expected or average value. Clearly, such effects have limits, and exploring these boundaries is an important question for future work. But my results suggest that at least sometimes failures to pay attention will lead us away from a sensible default, with implications for when and how reminders and information will improve decisions.

References

Aina, C. (2023). Tailored stories.

Alesina, A., Miano, A., & Stantcheva, S. (2023, 1). Immigration and redistribution. *The Review of Economic Studies*, *90*, 1-39. doi: 10.1093/restud/rdac011

Allcott, H., & Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, *105*(8), 2501–2538. doi: 10.1257/aer.20131564

Altmann, S., Grunewald, A., & Radbruch, J. (2021). Interventions and cognitive spillovers. *The Review of Economic Studies*. doi: 10.1093/restud/rdab087

Arrieta, G., & Nielsen, K. (2024). *Procedural decision-making in the face of complexity* (Tech. Rep.). Working Paper.

Augenblick, N., Backus, M., Little, A., & Moore, D. (2024). *Model uncertainty, disagreement, and over-precision: Theory and evidence.* (Working paper)

Ba, C., Bohren, J. A., & Imas, A. (2023). *Over-and underreaction to information.* Retrieved from https://ssrn.com/abstract=4274617

Barrera, O., Guriev, S., Henry, E., & Zhuravskaya, E. (2020, 2). Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics*, *182*. doi: 10.1016/j.jpubeco.2019.104123

Barron, K., & Fries, T. (2023). Narrative persuasion. *CESifo Working Paper no. 10206*.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., & Zinman, J. (2010). What's advertising content worth? evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics*, *125*, 263-306.

Bohren, J. A., Hascher, J., Imas, A., Ungeheuer, M., & Weber, M. (2024). *A cognitive foundation for perceiving uncertainty.* Retrieved from https://ssrn.com/abstract=4706147

Bordalo, P., Burro, G., Coffman, K., Gennaioli, N., & Shleifer, A. (2024). Imagining the future: Memory, simulation, and beliefs about covid. *Working paper.*.

Bordalo, P., Conlon, J. J., Gennaioli, N., Kwon, S. Y., & Shleifer, A. (2023). How people use statistics. *NBER Working Paper No. 31631*.

Bordalo, P., Gennaioli, N., Lanzani, G., & Shleifer, A. (2024). A cognitive theory of reasoning and choice. *Working paper*.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and Consumer Choice. *Journal of Political Economy*, *121*(5), 803–843. doi: 10.1086/673885

Bordalo, P., Gennaioli, N., & Shleifer, A. (2020). Memory, attention, and choice. *Quarterly Journal of Economics*.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, 1-42.

Bradley, S., & Feldman, N. E. (2020). Hidden baggage: Behavioral responses to changes in airline ticket tax disclosure. *American Economic Journal: Economic Policy*, *12*, 58-87. doi: 10.1257/pol.20190200

Brown, J., Hossain, T., & Morgan, J. (2010). Shrouded Attributes and Information Suppression: Evidence from the Field. *The Quarterly Journal of Economics*, *125*(2), 859–876.

Bushong, B., Rabin, M., & Schwartzstein, J. (2021). A model of relative thinking. *Review of Economic Studies*.

Caplin, A., & Dean, M. (2015, 7). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, *105*, 2183-2203. doi: 10.1257/aer.20140117

Cerreia-Vioglio, S., Dillenberger, D., & Ortoleva, P. (2015). Cautious expected utility and the certainty effect. *Econometrica*, *83*, 693-728. doi: 10.3982/ecta11733

Charles, C., & Kendall, C. (2024). Causal narratives. *Working paper*.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*, 1145-1177. doi: 10.1257/aer.99.4.1145

Coffman, L., & Niehaus, P. (2020, 11). Pathways of persuasion. *Games and Economic Behavior*, *124*, 239-253. doi: 10.1016/j.geb.2020.08.008

Cohn, A., & Maréchal, M. A. (2016, 12). Priming in economics. *Current Opinion in Psychology*, *12*, 17-21. doi: 10.1016/j.copsyc.2016.04.019

Colonnelli, E., Gormsen, N. J., & Mcquade, T. (2023). Selfish corporations.

Conlon, J. J., & Patel, D. (2023). *What jobs come to mind? stereotypes about fields of study.*

Dai, W., Yang, T., White, B. X., Palmer, R., sanders, E. K., McDonald, J. A., ... Albarracín, D. (2023). Priming behavior: A meta-analysis of the effects of behavioral and nonbehavioral primes on overt behavioral outcomes. *Psychological Bulletin*. doi: 10.1037/bul0000374.supp

Dean, M., & Neligh, N. (2023). Experimental tests of rational inattention. *Journal of Political Economy*.

De Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, *108*(11), 3266–3302.

Dertwinkel-Kalt, M., Köhler, K., Lange, M. R., & Wenzel, T. (2017). Demand shifts due to salience effects: Experimental evidence. *Journal of the European Economic Association*, *15*, 626-653. doi: 10.1093/jeea/jvw012

Enke, B., & Graeber, T. (2023). Cognitive uncertainty. *The Quarterly Journal of Economics*.

Enke, B., Graeber, T., Oprea, R., & Yang, J. (2024). *Behavioral attenuation* (Tech. Rep.). National Bureau of Economic Research.

Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *The review of economic studies*, *86*(1), 313–332.

Frydman, C., & Mormann, M. (2017). The role of salience and attention in choice under risk: An experimental investigation. *Working paper*.

Gabaix, X. (2014). A Sparsity-Based Model of Bounded Rationality. *The Quarterly Journal of Economics*, 1661–1710. doi: 10.1093/qje/qju024.Advance

Gabaix, X. (2019). *Behavioral inattention* (Vol. 2). Elsevier B.V. doi: 10.1016/bs.hesbe.2018.11.001

Golman, R., Loewenstein, G., Molnar, A., & Saccardo, S. (2021). The demand for, and avoidance of, information. *Management Science*.

Graeber, T., Roth, C., & Zimmermann, F. (2022). *Stories, statistics, and memory.*

Haaland, I., Roth, C., & Wohlfart, J. (2023, 3). Designing information provision experiments. *Journal of Economic Literature*, *61*, 3-40. doi: 10.1257/jel.20211658

Hébert, B., & Woodford, M. (2021, 10). Neighborhood-based information costs. *American Economic Review*, *111*, 3225-3255. doi: 10.1257/AER.20200154

Kamenica, E. (2019). Bayesian persuasion and information design. *Annual Review of Economics*. doi: 10.1146/annurev-economics

Koszegi, B., & Szeidl, A. (2013). A Model of Focusing in Economic Choice. *Quarterly Journal of Economics*, 53–104. doi: 10.1093/qje/qjs049.Advance

Li, X., & Camerer, C. F. (2022). Predictable Effects of Bottom-up Visual Salience in Experimental Decisions and Games. *Quarterly Journal of Economics*.

Loewenstein, G., & Wojtowicz, Z. (2023). *The economics of attention.* Retrieved from https://ssrn.com/abstract=4368304

Mackowiak, B., Matejka, F., & Wiederholt, M. (2022). Rational inattention: A review. *Annual Review of Economics*. doi: 10.2866/417246

Martínez-Marquina, A., Niederle, M., & Vespa, E. (2019). Failures in contingent reasoning: The role of uncertainty. *American Economic Review*, *109*(10), 3437–3474.

Matějka, F., & Mckay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, *105*, 272-298.

Mullainathan, S., Schwartzstein, J., & Shleifer, A. (2008). Coarse thinking and persuasion. *The Quarterly Journal Of Economics*.

Oprea, R. (2020). What makes a rule complex? *American economic review*, *110*(12), 3913–3951.

Oprea, R. (2024). Decisions under risk are decisions under complexity. *American Economic Review*.

Quispe-Torreblanca, E., Gathergood, J., Loewenstein, G., & Stewart, N. (2022). Attention utility: Evidence from individual investors.

Schwartzstein, J., & Sunderam, A. (2021). Using models to persuade. *American Economic Review*, *111*, 276-323. doi: 10.1257/aer.20191074

Sepulveda, P., Usher, M., Davies, N., Benson, A. A., Ortoleva, P., & Martino, B. D. (2020, 10). Visual attention modulates the integration of goal-relevant evidence and not value. *eLife*, *9*, 1-58. doi: 10.7554/eLife.60705

Shubatt, C., & Yang, J. (2024). Similarity and comparison complexity. *arXiv preprint arXiv:2401.17578*.

Sims, C. A. (2011). Rational inattention and monetary economics. *Handbook of Monetary Economics*, *3A*. doi: 10.1016/S0169-7218(11)03004-8

Somerville, J. (2022). Range-dependent attribute weighting in consumer choice: An experimental test. *Econometrica*, *90*, 799-830. doi: 10.3982/ecta18412

Taubinsky, D., & Rees-Jones, A. (2018). Attention variation and welfare: Theory and evidence from a tax salience experiment. *Review of Economic Studies*, *85*(4), 2462–2496. doi: 10.1093/restud/rdx069

Winichakul, K. P., Lezema, G., Mustafi, P., Lepper, M., Wilson, A., Danz, D., & Vesterlund, L. (2024). The effect of experimenter demand on inference.

Wiswall, M., & Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, *82*, 791-824.

Yang, X., & Krajbich, I. (2022). A dynamic computational model of gaze and choice in multi-attribute decisions. *Psychological Review*.

A Additional Tables

Table A.I: Attention at Baseline

	Full Sample			Correct-Beliefs Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Option A Value	0.128***	0.143***	0.128***	0.126***	0.131***	0.127***
	(0.002)	(0.007)	(0.002)	(0.002)	(0.008)	(0.002)
Option A Lottery	0.014***	0.022**	0.014***	0.016***	0.018	0.017***
	(0.003)	(0.011)	(0.003)	(0.004)	(0.012)	(0.004)
Option B Coins	0.096***	0.126***		0.098***	0.122***	
	(0.004)	(0.013)		(0.004)	(0.015)	
Option B Boxes	0.045***	0.052***		0.051***	0.073***	
	(0.004)	(0.013)		(0.004)	(0.014)	
Option A Value^2		0.002**			0.001	
		(0.001)			(0.001)	
Option A Lottery^2		0.001			0.000	
		(0.001)			(0.001)	
Option B Coins^2		-0.012**			-0.010	
		(0.005)			(0.006)	
Option B Boxes^2		-0.004			-0.011	
		(0.006)			(0.007)	
Option B Changing			0.073***			0.079***
			(0.002)			(0.003)
Option B Frozen			0.054***			0.049***
			(0.011)			(0.012)
Constant	0.621***	0.621***	0.643***	0.619***	0.592***	0.641***
	(0.023)	(0.030)	(0.023)	(0.027)	(0.035)	(0.027)
Observations	11,800	11,800	11,800	8,740	8,740	8,740
Individuals	590	590	590	437	437	437
R^2	0.35	0.35	0.34	0.35	0.35	0.34
p-value: Option A Value = Boxes	0.00	0.00		0.00	0.00	
p-value: Option A Lottery = Boxes	0.00	0.06		0.00	0.00	
p-value: Coins = Boxes	0.00	0.00		0.00	0.02	
p-value: Squared terms all zero		0.01			0.22	
p-value: Moving = Frozen			0.08			0.02

Notes: Each column shows OLS estimates of modifications of equation 6 using the Period 1 decisions of all treatment groups. The dependent variable is whether the participant chose Option B. The independent variables in columns 1 and 4 include the certain value of Option A, the subjective value of Option A's lottery, the sum of Option B's coins, and the sum of Option B's colored boxes. The subjective value of the lottery is calculated by multiplying the prize for winning the lottery with each participants' prior belief about their odds of winning (winsorized at the 90th percentile). The certain and lottery values of Option A are multiplied by negative one so that the expected sign of all uninteracted attribute coefficients is positive. Columns 2 and 5 additionally include the squared values of these attributes (squaring the individual components and then taking the sum). The independent variables in columns 3 and 6 are identical to columns 1 and 4 for Option A, but for Option B include, first, the sum of the five changing attributes and, second, the attribute that was frozen throughout the experiment at its initial value. Columns 1-3 include all participants, while columns 4-6 include only participants who correctly respond to unincentivized questions at the end of the survey about the monetary value of each coin and every possible colored box. Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.II: Attention at Baseline: Logit Specification

	Full Sample			Correct-Beliefs Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Option A Value	0.792***	1.099***	0.779***	0.806***	1.122***	0.796***
	(0.025)	(0.071)	(0.024)	(0.026)	(0.077)	(0.025)
Option A Lottery	0.089***	0.137*	0.089***	0.108***	0.115	0.108***
	(0.021)	(0.074)	(0.021)	(0.027)	(0.080)	(0.023)
Option B Coins	0.605***	0.796***		0.633***	0.785***	
	(0.030)	(0.098)		(0.032)	(0.080)	
Option B Boxes	0.286***	0.341***		0.336***	0.472***	
	(0.025)	(0.080)		(0.030)	(0.104)	
Option A Value^2		0.036***			0.037***	
		(0.007)			(0.008)	
Option A Lottery^2		0.005			0.001	
		(0.007)			(0.008)	
Option B Coins^2		-0.077**			-0.060*	
		(0.036)			(0.033)	
Option B Boxes^2		-0.027			-0.067	
		(0.038)			(0.051)	
Option B Changing			0.456***			0.509***
			(0.020)			(0.024)
Option B Frozen			0.340***			0.319***
			(0.060)			(0.076)
Constant	0.685***	1.094***	0.797***	0.682***	1.061***	0.797***
	(0.132)	(0.212)	(0.155)	(0.169)	(0.217)	(0.172)
Observations	11,800	11,800	11,800	8,740	8,740	8,740
Individuals	590	590	590	437	437	437
p-value: Option A Value = Boxes	0.00	0.00		0.00	0.00	
p-value: Option A Lottery = Boxes	0.00	0.07		0.00	0.00	
p-value: Coins = Boxes	0.00	0.00		0.00	0.03	
p-value: Squared terms all zero		0.00			0.00	
p-value: Moving = Frozen			0.05			0.01

Notes: Each column shows logit estimates of modifications of equation 6 using the Period 1 decisions of all treatment groups. The dependent variable is whether the participant chose Option B. The independent variables in columns 1 and 4 include the certain value of Option A, the subjective value of Option A's lottery, the sum of Option B's coins, and the sum of Option B's colored boxes. The subjective value of the lottery is calculated by multiplying the prize for winning the lottery with each participants' prior belief about their odds of winning (winsorized at the 90th percentile). The certain and lottery values of Option A are multiplied by negative one so that the expected sign of all uninteracted attribute coefficients is positive. Columns 2 and 5 additionally include the squared values of these attributes (squaring the individual components and then taking the sum). The independent variables in columns 3 and 6 are identical to columns 1 and 4 for Option A, but for Option B include, first, the sum of the five changing attributes and, second, the attribute that was frozen throughout the experiment at its initial value. Columns 1-3 include all participants, while columns 4-6 include only participants who correctly respond to unincentivized questions at the end of the survey about the monetary value of each coin and every possible colored box. Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.III: Attention at Baseline for Treatment 1 Correct-Beliefs Sample

	OLS			Logit		
	(1)	(2)	(3)	(4)	(5)	(6)
Option A Value	0.124***	0.108***	0.124***	0.764***	0.931***	0.760***
	(0.006)	(0.018)	(0.006)	(0.065)	(0.216)	(0.056)
Option A Lottery	0.024***	-0.025	0.025***	0.151***	-0.173	0.156**
	(0.009)	(0.026)	(0.009)	(0.056)	(0.193)	(0.062)
Option B Coins	0.092***	0.116***		0.580***	0.734***	
	(0.008)	(0.034)		(0.071)	(0.210)	
Option B Boxes	0.055***	0.079***		0.359***	0.498***	
	(0.008)	(0.025)		(0.066)	(0.147)	
Option A Value^2		-0.002			0.018	
		(0.002)			(0.021)	
Option A Lottery^2		-0.005*			-0.035*	
		(0.003)			(0.021)	
Option B Coins^2		-0.010			-0.060	
		(0.014)			(0.080)	
Option B Boxes^2		-0.012			-0.066	
		(0.011)			(0.065)	
Option B Changing			0.075***			0.478***
			(0.006)			(0.050)
Option B Frozen			0.063**			0.397**
			(0.026)			(0.173)
Constant	0.666***	0.541***	0.685***	0.902***	0.567	1.002***
	(0.052)	(0.067)	(0.053)	(0.328)	(0.584)	(0.329)
Observations	2,020	2,020	2,020	2,020	2,020	2,020
Individuals	101	101	101	101	101	101
R^2	0.33	0.34	0.33			
p-value: Option A Value = Boxes	0.00	0.37		0.00	0.10	
p-value: Option A Lottery = Boxes	0.02	0.00		0.01	0.00	
p-value: Coins = Boxes	0.00	0.39		0.00	0.34	
p-value: Squared terms all zero		0.19			0.08	
p-value: Moving = Frozen			0.67			0.64

Notes: Each column shows OLS (columns 1-3) or logit (columns 4-6) estimates of modifications of equation 6 using participants' Period 1 decisions with the following modifications. The dependent variable is whether the participant chose Option B. The independent variables in columns 1-2 and 4-5 include the certain value of Option A, the subjective value of Option A's lottery, the sum of Option B's coins, and the sum of Option B's colored boxes. The subjective value of the lottery is calculated by multiplying the prize for winning the lottery with each participants' prior belief about their odds of winning (winsorized at the 90th percentile). Columns 2 and 5 include the squared values of these attributes (squaring the individual components and then taking the sum). The certain and lottery values of Option A are multiplied by negative one so that the expected sign of all uninteracted attribute coefficients is positive. The attributes in columns 3 and 6 are identical to columns 1 and 4 for Option A, but for Option B include, first, the sum of the five changing attributes and, second, the attribute that was frozen throughout the experiment at its initial value. Data are restricted to Treatment-1 participants (who do not receive any information about Option-B attributes throughout the experiment) who correctly respond to unincentivized questions at the end of the survey about the monetary value of each coin and every possible colored box. Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.IV: Effects of Information about the Target Attribute

	Full Sample			Correct-Beliefs Sample		
	Pooled (1)	Coin (2)	Box (3)	Pooled (4)	Coin (5)	Box (6)
Info	0.022***	0.013	0.041***	0.022***	0.014	0.039***
	(0.007)	(0.010)	(0.010)	(0.008)	(0.011)	(0.011)
Option A Value X No Info	0.127***	0.132***	0.123***	0.125***	0.131***	0.121***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Option A Value X Info	0.129***	0.133***	0.125***	0.128***	0.134***	0.124***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Option A Lottery X Info	0.000	-0.000	0.001	-0.001	0.000	-0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Target Attribute X No Info	0.074***	0.101***	0.047***	0.074***	0.097***	0.053***
	(0.005)	(0.006)	(0.006)	(0.005)	(0.007)	(0.007)
Target Attribute X Info	0.125***	0.122***	0.132***	0.132***	0.122***	0.143***
	(0.006)	(0.008)	(0.008)	(0.006)	(0.009)	(0.008)
Other Changing Option B Attributes X No Info	0.075***	0.070***	0.080***	0.080***	0.073***	0.087***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Other Changing Option B Attributes X Info	0.069***	0.063***	0.074***	0.073***	0.062***	0.082***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)
Frozen Option B Attribute X Info	-0.017*	-0.030**	-0.003	-0.014	-0.023*	-0.004
	(0.009)	(0.012)	(0.012)	(0.009)	(0.013)	(0.012)
Observations	23,600	11,400	12,200	17,480	8,200	9,280
Individuals	590	285	305	437	205	232
R^2	0.513	0.517	0.512	0.523	0.526	0.525
Effect on Target	0.051	0.022	0.085	0.057	0.025	0.089
p-value: Effect on Target is Zero	0.000	0.020	0.000	0.000	0.021	0.000
Total Effect on Non-Targets	-0.038	-0.056	-0.025	-0.040	-0.065	-0.021
p-value: Total Effect on Non-Targets is Zero	0.019	0.011	0.272	0.020	0.009	0.376
Total Effect on All Attributes	0.014	-0.034	0.059	0.017	-0.040	0.069
p-value: Total Effect on All Attributes is Zero	0.430	0.141	0.023	0.375	0.127	0.013

Notes: This table shows OLS estimates of modifications of equation 6 using the Periods 1 and 2 decisions of all treatment groups. The dependent variable is whether the participant chose Option B. The independent variables are certain value of Option A, the subjective value of Option A's lottery, the target attribute of Option B, the sum of the four other changing attributes of Option B, and the frozen Option B attribute, and individual-fixed effects. The certain and lottery values of Option A are multiplied by negative one so that the expected effect of increasing all attributes is positive. "Info" is a dummy variable for receiving information about the target attribute (i.e., being in Treatments 2-4 during Period 2). Columns 1-3 include all participants, while columns 4-6 include only participants who correctly respond to unincentivized questions at the end of the survey about the monetary value of each coin and every possible colored box. The main effect of the Option A Lottery and Frozen Option B Attribute are excluded because there are co-linear with the individual-fixed effects (and interactions of these attributes with information represent the difference in responsiveness when receiving information compared to not). The row showing the "Effect on Target" (and associated p-value) refers to the difference between the coefficient on "Target Attribute" with and without information. The row showing the "Effect on Non-Targets" (and associated p-value) first adds the coefficients for all attributes other than the target attribute (multiplying the coefficient on the other changing Option B attributes by four since there were four such attributes) and takes the difference between this value with and without information. "Total Effect on All Attributes" is analogous but also includes the target attribute. Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.V: Dynamics of Information Effects

	Full Sample				Correct-Beliefs Sample			
	T1	T2	T3	T4	T1	T2	T3	T4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Target Option B Attribute	0.064***	0.119***	0.122***	0.067***	0.074***	0.123***	0.113***	0.071***
	(0.010)	(0.007)	(0.010)	(0.012)	(0.011)	(0.008)	(0.012)	(0.013)
Alternative Option B Attribute	0.077***	0.070***	0.069***	0.120***	0.082***	0.080***	0.063***	0.116***
	(0.010)	(0.007)	(0.010)	(0.011)	(0.011)	(0.008)	(0.012)	(0.012)
Other Option B Attributes	0.068***	0.075***	0.070***	0.059***	0.073***	0.080***	0.077***	0.073***
	(0.006)	(0.004)	(0.007)	(0.007)	(0.007)	(0.005)	(0.007)	(0.007)
Option A Value	0.132***	0.131***	0.127***	0.134***	0.130***	0.132***	0.124***	0.128***
	(0.004)	(0.003)	(0.004)	(0.005)	(0.005)	(0.003)	(0.005)	(0.006)
Observations	2,680	4,780	2,280	2,060	2,020	3,580	1,600	1,540
Individuals	134	239	114	103	101	179	80	77
R^2	0.56	0.55	0.57	0.55	0.57	0.56	0.57	0.55
p-value: Target Same as T1		0.00	0.00	0.85		0.00	0.01	0.86
p-value: Target Same as T2			0.81	0.00			0.46	0.00
p-value: Target Same as T3				0.00				0.02
p-value: Alternative Same as T1		0.57	0.57	0.00		0.94	0.27	0.04

Notes: This table shows OLS estimates of modifications of equation 6 using the Period 3 decisions of each treatment group (separately by column, treatments indicated in the column headings). The dependent variable is whether the participant chose Option B. The independent variables are the certain value of Option A, the target attribute of Option B, the alternative attribute of Option B, the sum of the four other attributes of Option B, and individual-fixed effects. The certain value of Option A is multiplied by negative one, such that the expected sign of all coefficients is positive. Treatment 1 (T1) never received information about the target attribute. Treatments 2-4 (T2-T4) received information about the target attribute in Period 2, but differed in the information presented during Period 3. During this period, Treatment 2 received no information, Treatment 3 continued to receive information about the target attribute, and Treatment 4 received information about the alternative attribute. Columns 1-4 include the full sample, while columns 5-8 include only participants who correctly respond to unincentivized questions at the end of the survey about the monetary value of each coin and every possible colored box. Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

B For Online Publication: Data Appendix

Here I provide additional details on the experiment. Participants were recruited through Prolific to participate in a "Quick Survey on Decision Making." Potential survey-takers were not told anything of the content of the survey except that it was part of a research study and that it would be more difficult to complete if they struggled to tell colors apart. A total of 590 participants completed the survey during December 2022. All participants are US residents, the average participant is 41 years old, and 51% are women. The median respondent took 21 minutes to complete the experiment. The experiment paid a $3.00 completion fee plus any bonus that participants earned from their choices. The full experimental survey can be viewed by following this link: `https://harvard.az1.qualtrics.com/jfe/form/SV_5jbhR2VLKqq2n6S`.

Participants had to correctly answer a series of comprehension checks in order to continue with the survey. If they initially answered a question incorrectly, they were forced to revise their answer before proceeding. I do not exclude anyone from the data for poor performance, but 95% of comprehension questions were answered correctly on the first try, suggesting a high level of engagement and understanding. Table B.I gives more details on these comprehension questions and provides a link to the experimental instructions.

In Section 4.6, I describe open-ended responses that participants gave describing how they made their decisions. These come from a question at the very end of the experiment (i.e., after all choices and belief elicitation) asking "In your own words, how did you choose between option A and option B? Did you use any strategies to try to simplify the problem?". An open text box appeared below the question where participants could (though they were not required to) write a response. I then feed these responses into a large language model (GPT-4o) to derive binary yes-no responses to the following questions about participants' responses (rates of "Yes" responses in parentheses):

1. "Does this respondent mention attempting to add up the value of the options, or choosing the option that seemed to add up to the highest amount?" (66.8%)
2. "Does this respondent mention trying to identify or guessing or approximating what the best option was (rather than knowing what it was)?" (42.5%)
3. "Does this respondent mention anything reacting to reminders or information (other than the baseline rules of the experiment)?" (1.2%)
4. "Does this respondent mention anything about what the researcher's hypothesis might have been (or altering their decisions at all based on what the researcher seemed to want)?" (0.0%)
5. "Does this respondent mention anything about doing what the experiment *wanted*

or *suggested to* her to do (or any variation thereof), rather than what she otherwise thought was the right choice?" (0.0%)

Two differences between the preregistration and the analysis that appears in the main text bear mentioning. First, the preregistration mentions that some participants see information in period 2 about the alternative attribute and no information in period 3. However, because participants are not told which are the target and alternative attributes, this treatment is equivalent to being told about the target attribute (i.e., simply changing labels for which are the "target" and "aternative" attributes). I combine these into Treatment group 2 after this relabeling.

Second, the preregistration mentions a second experiment ($N = 211$) with the following differences from the experiment in the main text. First, there was no lottery associated with Option A. Second, the level of attributes was chosen such that the target attribute of Option B was always pivotal in deciding which option yielded the higher bonus. That is, whenever the pivotal attribute took on either its intermediate or high value (recall that all Option B attributes had three possible values), Option B was the payoff-maximizing choice. The other "non-pivotal" attributes were by construction uncorrelated with the payoff-maximizing decision.

Just like in the main experiment, no one saw information in the first period. Unlike in the main experiment, in each of the remaining three periods, participants either saw no information, information about the (pivotal) target attribute, or information about a randomly selected non-target (and therefore non-pivotal) attribute. This randomization occurred across periods and within participant.

This experiment was intended to test how the welfare effect of information depends on whether it directs attention toward pivotal or non-pivotal attributes. Table B.II shows an OLS regression where the dependent variable is an indicator for whether the participant chose the lower-value option. It regresses this variable on individual-fixed effects and indicators for whether the participant was receiving information about the pivotal target attribute or a non-pivotal non-target attribute. We see that, while both types of information reduce the rate at which participants mistakenly choose the lower-value option, these effects are larger when the information is about a pivotal attribute ($p < 0.01$). One explanation for why even information about non-pivotal attributes may have improved decision-making is that in Period 1 (i.e., absent any information) participants only choose Option B 41% of the time despite it being the payoff-maximizing choice 70% of the time. Thus, because paying more attention to the non-pivotal attribute boosts demand for Option B, it therefore also reduces mistakes. These results are again consistent with the default value being less than participants' priors.

Table B.I: Comprehension Questions

	Topic	% Correct on First Attempt
Question #1	One choice is randomly chosen to be implemented	97.1%
Question #2	How bonus is calculated if Option A is chosen	87.4%
Question #3	Value of coins in Option B	98.3%
Question #4	Value of each color for top box	93.9%
Question #5	Value of each color for middle box	95.6%
Question #6	Value of each color for bottom box	96.5%
Question #7	Information is provided randomly	Not recorded*

Notes: The instructions that participants saw, along with the text of each comprehension question can be found at **this link**. *The fraction of participants who correctly answered question #7 on the first attempt was not recorded due to a coding error, so this statistic is unavailable.

Table B.II: Effects of Information about Pivotal and Non-Pivotal Information

	Mistake (1)
Information about Pivotal Attribute	-0.083***
	(0.011)
Information about Non-Pivotal Attribute	-0.046***
	(0.012)
Observations	16,880
Individuals	211
R^2	0.23
p-value: Information Effects Equal	0.00

Notes: This table shows an OLS regression where the dependent variable is an indicator for whether the participant chose the lower-value option. It regresses this variable on individual-fixed effects and indicators for whether the participant was receiving information about the pivotal target attribute and non-pivotal non-target attribute. Robust standard errors, clustered at the individual level, are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.