Beyond Instrumental Value:

How Complexity Shapes Information Demand

Menglong Guan	Ryan Oprea	Sevgi Yuksel
Penn State	UC Berkeley	NYU

October 23, 2025

Abstract

We experimentally investigate how characteristics of information structures beyond their instrumental value influence information demand. Using novel methods that minimize Bayesian reasoning errors, we find that while participants strongly favor structures with higher instrumental value, they systematically deviate from optimal choices in predictable ways. We identify four key characteristics of information structures that drive these deviations and show that they are as predictive of behavior as individual characteristics, with stronger effects among participants who are otherwise more likely to act optimally. Our findings suggest that these distortions stem from the cognitive difficulty of evaluating information structures, with implications for models of information demand and the design of effective disclosure strategies.

We would like to thank numerous seminar participants for helpful comments. Guan: mkg6150@psu.edu; Oprea: roprea@gmail.com; Yuksel: sevgiy@gmail.com. This research was supported by the National Science Foundation under Grant SES-1949366 and was approved by UC Santa Barbara IRB. This paper replaces and subsumes our earlier working paper titled "Too Much Information".

1 Introduction

Few products are chosen more frequently by modern consumers than sources of information. From news sources to social media influencers, from product reviewers to scientific experts, decision makers are continuously tasked with deciding what competing sources of information to rely on. How do people make these decisions? Economic theory provides a sharp answer: demand for sources of information should be shaped by instrumental value—the expected utility improvement the information can be expected to produce in relevant decisions. However, as a growing literature in economics suggests, assessing the value of complicated objects like information structures requires intensive information processing that may make discerning instrumental value difficult. As a result, characteristics of information structures other than instrumental value may significantly influence the way people make these choices.

In this paper, we report a novel experiment designed to better understand the role instrumental value plays in people's demand for information and to identify what other formal characteristics of information structures also shape this demand. To do this, we ask subjects to preferentially *rank* hundreds of information structures that vary widely in a number of characteristics, including instrumental value. Higher-ranked structures are more likely to be later given to subjects to help them make an incentivized decision (a guess about the realization of a random variable), making this elicitation incentive compatible. We also collect information on subjects' ability to *use* signals from information structures exogenously assigned to them, allowing us to study the relationship between subjects' *demand* for information and their ability to optimally use it.

An important innovation in our design is a novel method for presenting information structures to subjects that we believe (and empirically verify) is particularly intuitive. States of the world are presented as a collection of events (black and white balls, shown on their screen), and information structures are presented as partitions of these events. Subjects receiving a signal from an information structure are simply told which element of the partition the true event lies in (see Figure 1 for an example). This allows us to present information structures graphically to subjects, facilitating relatively fast assessment and enabling us to ask individual subjects to rank a number of information structures during the experiment. As we verify in the data, this intuitive way of presenting information structures eliminates previously identified biases such as base-rate neglect and confirmation bias that are orthogonal to our experimental question.

As expected, we find that subjects' demand for (ranking of) information structures is strongly influenced by instrumental value. However, we also find that subjects frequently (in 27% of pairwise

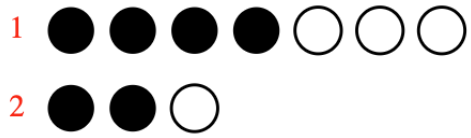


Figure 1: Example on the Representation of Information Structures in the Experiment *Notes: The subjects' task is to guess the color of a randomly selected ball. The information structure reveals whether the randomly selected ball is in group 1 or group 2.*

comparisons) rank less valuable structures above more valuable structures—an error that causes significant loss of earnings in the experiment. By contrast, we find that conditional on actually receiving a signal from an information structure, subjects make near-perfect use of this information. This strongly suggests that it is the task of ex-ante evaluation—that is, of assessing information structures and weighing relevant tradeoffs in pairwise comparisons—rather than the ex-post difficulty of optimally using the signals generated by these information structures, that drives the demand errors we observe.

Importantly, we find that these errors have, to a significant extent, predictable structure. Searching over a number of formal metrics on information structures, we find that four characteristics explain 77% of the explainable variation in deviations from optimality. These four factors appear to shape demand in notably diverse ways. Two of these factors are measures of the *size* of the information structure: the number of distinct posteriors it induces, and the number of signals that retain uncertainty, that is, signals that are not fully revealing of the state. We find that subjects tend to prefer information structures that can be interpreted as *simpler* along these two dimensions—those that produce fewer posteriors or fewer signals with residual uncertainty. Since the size of information structures likely influences the amount of information processing required to evaluate them, we interpret this as evidence that *complexity aversion* may have an important influence on information demand. We also find that subjects tend to prefer information structures that would have a higher instrumental value if (counterfactually) all signals from information structures were equally likely to occur. We interpret this as evidence that subjects use natural simplifying heuristics to approximate the value of structures: they ignore, or insufficiently account for, signal probabilities, reducing the intensity of the computations required to calculate instrumental value. Finally, we find that the total instrumental value of the two structures being evaluated tends to predict *attenuation*: subjects are less sensitive to differences in instrumental value between two structures when the sum of values is low. This suggests that information structures with low instrumental value are more difficult for participants to compare.

To benchmark these results on information structure characteristics, we compare them to par-

allel findings on how characteristics of *decision makers* influence information demand. We find that these structure-level predictors influence information demand to a similar degree as, for instance, decision makers' intelligence as measured by Raven's matrix scores. However, we also find a surprising relationship between measures of subject sophistication (e.g., Raven's scores) and the structure characteristics that predict demand errors: structure characteristics are *more* predictive of behavior for high-sophistication than low-sophistication subjects. This is true despite the fact that (unsurprisingly) more sophisticated subjects make fewer overall errors. What this result suggests is that errors made by unsophisticated subjects tend to be more random (and less predictable) than errors made by sophisticated subjects, perhaps because even assessing which structures are more complex requires some level of attention, effort, and sophistication. As a result, even the *errors* of sophisticated subjects are more predictable than those of unsophisticated subjects! Additionally, we find that the four structure characteristics described above are more predictive of behavior among ambiguity-averse participants, suggesting a connection between responses to complexity and attitudes towards ambiguity (see de Clippel, Moscariello, Ortoleva & Rozen (2024) for related evidence).

We interpret these results as evidence that evaluating the relative value of information sources is a complex task for many people—one that can require significant information processing. As a result, characteristics of information structures other than instrumental value can sometimes have a powerful secondary influence on the way people choose between information sources. Importantly, our findings suggest that these difficulties follow predictable patterns, meaning these information processing costs may be amenable to modeling, improving our ability to predict and characterize the way people seek out and consume information.

Related Literature

Our paper contributes to several literatures.

Information Demand. First, we make a methodological contribution to the growing experimental literature studying information demand. Our design allows us to study demand for information in a setting where making optimal use of information is very easy and does not require difficult Bayesian reasoning. Thus, our visual representation of information structures successfully excludes non-Bayesian reasoning or misinterpretation of information structures as confounds for studying the demand for information. Similar experimental designs can be used to study a wide range of questions about people's taste for information.

Second, our paper adds to a growing literature studying how factors other than instrumental value influence information demand. Prior studies have experimentally or theoretically examined

the role of confidence or confirmation seeking (Jones & Sugden 2001, Eliaz & Schotter 2007, 2010, Charness, Oprea & Yuksel 2021, Montanari & Nunnari 2022), preference for certainty (Ambuehl & Li 2018, Novk, Matveenko & Ravaioli 2024), timing of resolution of uncertainty (Grant, Kajii & Polak 1998, Nielsen 2020, Zimmermann 2015), skewness of information sources (Masatlioglu, Orhun & Raymond 2023), symmetry of information sources (Llorente-Saguer, Oliveros & Zultan 2025), anticipatory feelings (Caplin & Leahy 2001), motivated attention (Falk & Zimmermann 2024, Golman & Loewenstein 2018, Golman, Loewenstein, Molnar & Saccardo 2022), and behavioral motivations stimulated by changes in beliefs like disappointment aversion (Palacios-Huerta 1999, Dillenberger 2010, Andries & Haddad 2020), loss aversion (Koszegi & Rabin 2009), dissonance avoidance (Festinger 1957), suspense and surprise (Ely, Frankel & Kamenica 2015), etc., play in information demand.

Most closely related to our study in this literature is Liang (2023), a concurrent paper that studies suboptimal valuation of information structures and provides evidence that is broadly supportive of the mechanism underlying our main results. In particular, his results suggest that subjects have difficulty foreseeing and integrating payoffs from multiple information-contingent choices, but it is mostly difficulties with integration that get in the way of optimal valuation.

We contribute to this literature in several ways: (i) our experimental paradigm generates rich variation that enables us to examine how information demand responds to different structure characteristics; (ii) we compare the predictive power of these characteristics to individual-level characteristics such as cognitive ability; and (iii) we analyze interaction effects between structure characteristics and participant characteristics.

Complexity. Our study relates to a growing literature showing that cognitive limitations and complexity (i.e., costly information processing) plays an important role in a wide-range of economically important settings. For instance, recent work has provided evidence suggesting that complexity has an important influence on behavior in canonical settings like lottery choice (Enke & Graeber 2023, Oprea 2024*b*), intertemporal decision making (Enke, Graeber & Oprea 2025) and Bayesian updating (Ba, Bohren & Imas 2023). See Oprea (2024*a*) and Enke (2024) for recent reviews of this literature.

Perhaps most closely related to our paper in this literature is Enke & Shubatt (2024) which develops interpretable indices of complexity in lottery choice problems that are predictive of error rates in identifying the lottery with the highest expected value. Methodologically, we take a similar approach to identify characteristics of information structures that are predictive of mistakes in information demand.

Overall, we contribute to the literature on complexity by showing that information processing constraints extend to the no less canonical task of evaluating information. In this context, we find that complexity influences behavior through two distinct channels emphasized in prior work. First, we find evidence that decision makers are systematically *averse* to information structures that are complex in a specific, formal sense. Complexity aversion has been particularly emphasized in the context of lottery choice (Puri 2025), and recent work suggests it may be linked to ambiguity aversion arising from the complexity of the environment (de Clippel et al. 2024). Our findings are consistent with such a link. Second, we find that certain types of comparisons between information structures are complex for participants in the sense that they produce *attenuation*—an inelastic response to fundamentals that has been associated with cognitive limitations in recent work (Enke, Graeber, Oprea & Yang 2024). A distinctive feature of our study is that we document both effects—*aversion* and *attenuation*—and highlight the complementary ways in which they shape behavior within a unified experimental framework.

The remainder of the paper is organized as follows. Section 2 presents our theoretical framework. Section 3 describes the experimental design. Section 4 presents results. Section 5 concludes by discussing the implications of our results.

2 Theoretical Framework

2.1 Instrumental Value of Information

Consider a finite state space Ω, with a typical state denoted by ω. The prior distribution on Ω is denoted by p. An information structure, indexed by i, is a stochastic mapping from the state space Ω to a finite set of signals S. It is useful to think of i as inducing a distribution over posteriors.[1] That is, given p, an information structure i induces (i) a distribution q^i over S and (ii) conditional on each signal s, a posterior distribution p^i_s over the state space.

In standard economic theory, the value of an information structure to a decision-maker (DM) depends on the decision problem the information structure is meant to inform. Suppose the DM faces a decision problem in which she observes signal s from information structure i and takes action $a \in A$ to maximize $\mathbb{E}[u(a,\omega)|s] := \sum p^i_s(\omega)u(a,\omega)$, where $u(a,\omega)$ describes the decision-maker's state-dependent utility function. The *instrumental value* (or simply *value*) of i, given the

[1] For each ω, let $\sigma^i_\omega(s) \in \Delta(S)$ denote the probability that signal s is produced by structure i. The probability of observing signal s is $q^i(s) := \sum_\omega p(\omega)\sigma^i_\omega(s)$. For each signal, the posterior distribution on Ω can be computed using Bayes' rule. Conditional on each signal s, $p^i_s(\omega) = \frac{p(\omega)\sigma^i_\omega(s)}{q^i(s)}$.

set of actions A available and utility function u, is the expected increase in utility made possible by the DM being able to condition her action on the realized signals:

$$V^i = \underbrace{\sum_{s \in S} q^i(s) \underbrace{\max_{a \in A} \mathbb{E}[u(a,\omega)|s]}_{\substack{\text{Expected utility} \\ \text{conditional on } s}}}_{\text{Expectation over } s} - \underbrace{\max_{a \in A} \mathbb{E}[u(a,\omega)]}_{\substack{\text{Expected utility} \\ \text{without } s}}.$$

Instrumental value is the key characteristic shaping information demand (i.e., preferences for information) in standard information economics.

2.2 A Guessing Task

We focus on a simple guessing task in our experiment to study the demand for information. The state of the world is binary, i.e. $\Omega = \{b, w\}$, where b (as will be clear in the next section) can be interpreted as referring to the color *black* and w referring to the color *white*. The decision-maker's prior is fixed at $p := p(b) = 0.6$. The decision-maker takes a binary action, $A = \{b, w\}$, with the goal of matching the state such that $u(a, \omega) = B$ (where B is a bonus) if $a = \omega$ and zero otherwise.

With this simple specification, the instrumental value of an information structure reduces to the following:

$$V^i = \Big(\underbrace{\sum_{s \in S} q^i(s) \underbrace{\max\{p^i_s, 1 - p^i_s\}}_{\substack{\text{Guessing accuracy} \\ \text{conditional on } s}}}_{\text{Expectation over } s} - \underbrace{p}_{\substack{\text{Guessing accuracy} \\ \text{without information}}} \Big) B, \tag{1}$$

where $q^i(s)$ is the probability that signal s is generated by structure i and p^i_s is the probability that $\omega = b$ conditional on signal s. Note that the expected utility of the decision-maker in this problem is equal to their guessing accuracy times the bonus associated with guessing correctly. Thus, the value of an information structure for a decision-maker, facing such a guessing task, is directly linked to the expected improvement in their guessing accuracy. In the following, we will simply refer to the expected improvement in guessing accuracy that a certain information structure induces as its instrumental value.

Although this setting is simple, as we will show in the next section, it allows us to construct a rich set of information structures that differ not only in their instrumental value but also in a wide variety of characteristics that plausibly influence the information processing required to evaluate information structures.

3 Experimental Design

3.1 Guessing Task and Representation of Information

Our experimental design is built around the simple guessing task introduced in the last section. A ball is randomly drawn from a set of 10 balls, six of which are always black and four of which are always white. The subjects' task is to correctly guess the color of that randomly selected ball. They earn a bonus of \$10 if they guess correctly and zero otherwise.

Before making their guesses, subjects receive information about the randomly selected ball from one information structure. Each information structure is represented as a partition of the 10 balls, i.e., a grouping of the 10 balls into non-empty subsets. An information structure provides information about the randomly selected ball by revealing to the subject *which subset of the partition* the randomly selected ball belongs to. Thus, each subset of the partition can be interpreted as a distinct signal that the structure might generate. The size of any subset (the number of balls in this subset) visually represents the probability with which that signal will be realized, and the composition of the balls within each subset (relative share of black to white balls) visually represents the posterior conditional on the signal being realized. Presenting information in this way therefore makes the characteristics of an information structure particularly transparent to subjects.

Figure 1 provides an example of an information structure that partitions the 10 balls into two subsets (two possible signals) displayed in two numbered rows. The first subset (labeled "1" and referred to as "group 1" in the experiment) consists of four black balls and three white balls; the second subset ("group 2") consists of the remaining two black balls and one white ball. This information structure provides information about the randomly selected ball by revealing which subset (group 1 vs. group 2) the randomly selected ball belongs to. Formally, the presented information structure generates a binary signal. The first (second) signal, corresponding to group 1 (group 2), is realized with a 70 (30) percent chance. Conditional on the first (second) signal, the posterior probability that the color of the randomly selected ball is black is $\frac{4}{7}$ ($\frac{2}{3}$), and intuitively, the optimal guess is b (b). In addition, as expressed in equation 1, these properties (the likelihood of each signal and the posterior conditional on this signal) are sufficient to determine the instrumental value of the information structure.

Using the partition representation, we construct *all* information structures with up to five signals. These correspond to all possible partitions of the 10 balls to one, two, three, four and five subsets, yielding 237 distinct structures, in total. Our analysis focuses on pairwise comparisons between these structures. Theoretically, there are 27,966 possible pairwise comparisons. In our

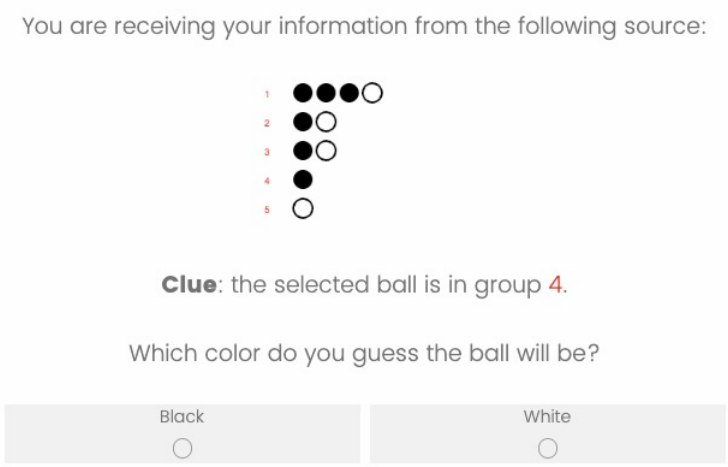


Figure 2: Guessing Task

data, 23,859 of these are observed at least once, with some comparisons appearing up to 14 times. This induces substantial variation in instrumental value and also variation in a number of other structure characteristics, allowing us to explore how subjects respond to changes in instrumental value and other factors that characterize information structures.[2]

3.2 Eliciting Guesses

Part 1 of the experiment consists of fifteen *guessing* tasks. In each task, subjects guess the color of a randomly selected ball based on a signal generated from a given information structure. The 15 distinct information structures used across tasks are randomly selected at the subject level from the set of 237 structures, without replacement. Figure 2 shows a screenshot of a guessing task. Subjects observe a partition of the black and white balls, with different subsets (groups) of the partition displayed in numbered rows. Subjects learn which group the selected ball is from and use this information to guess its color.

The purpose of the guessing tasks in the experiment is to study whether subjects can make optimal use of signals when they are presented using the partition design. This is important for interpreting their demand for information structures in Part 2.

3.3 Eliciting Preferences for Information Structures

Part 2 of the experiment elicits subjects' preferences for information structures using *ranking* tasks. In each task, subjects are asked to rank seven distinct information structures from most preferred to

[2]The complete list of structure characteristics considered is presented in Table 2 of Appendix A.

least preferred. In determining their rankings, participants were incentivized to consider all possible pairwise comparisons between the seven structures presented to them. Specifically, participants were told that two of the seven structures would be randomly selected, and they would receive information from the structure they had ranked higher. Our analysis accordingly focuses mainly on these pairwise comparisons.

Each subject is presented with eight ranking tasks. For each subject, the eight sets of seven structures (56 in total) used across these tasks are randomly selected without replacement from the 237 possible structures, subject to the constraint that each set includes structures spanning the full range of instrumental values.[3] This allows us to evaluate, in each ranking task, how demand for information responds to instrumental value and how this responsiveness is influenced by other characteristics of information structures that are not directly relevant to instrumental value (our primary question).

Figure 3 provides a screenshot of a ranking task. In the task, the seven selected information structures are presented in a randomized order and subjects are asked to reorder them using a drag-and-drop interface.[4]

3.4 Additional Tasks

The experiment also includes tasks designed to provide deeper insight into how individuals value information structures, as well as additional tasks that measure individual characteristics which we included to help us to interpret subjects' preferences for and use of information structures, allowing us to better assess heterogeneity across subjects.

At the end of Part 2, we ask subjects to indicate how confident they are that they ranked information structures in a way that maximizes their chance of winning the \$10 bonus.[5] Part 3

[3]Given the prior $p = 0.6$, the baseline guessing accuracy without any additional information is 0.6. Thus, the instrumental value of an information structure, defined as the expected increase in accuracy it induces, can take on values of 0, 0.1, 0.2, 0.3, or 0.4 in our experiment. Each of the eight ranking tasks includes at least one information structure corresponding to each of these distinct values.

[4]We include seven information structures in each ranking task to elicit preferences over a substantial share of the 237 possible structures without overtaxing participants. Without assuming transitivity, it would have required eliciting 21 pairwise choices to yield equivalent data to a ranking task. As discussed in Section 4, the high optimality rate of rankings and their similarity to those in a laboratory experiment with 16 structures per screen suggest participants found this elicitation manageable.

[5]Specifically, we ask subjects: for any two information structures A and B where they ranked A higher than B, how confident they are that their chance of winning the \$10 bonus (i.e., their guesses being correct) is as high or higher when receiving information from A versus B.

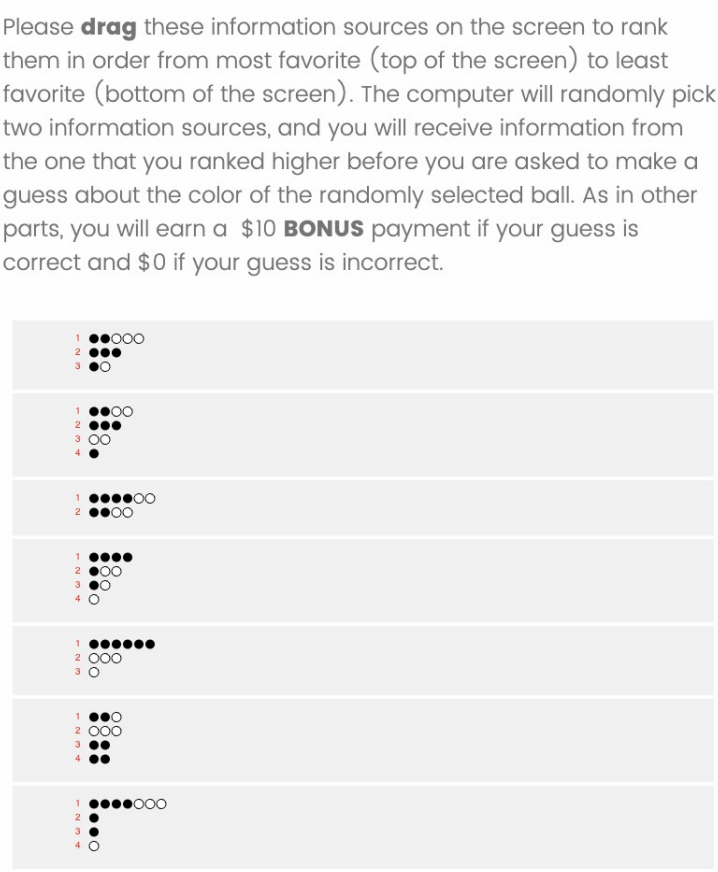


Figure 3: Ranking Task *Notes: As subjects drag and reorder structures, an automatically updating order number appears to the left of each structure. These numbers reflect the current ranking after each move. Check the experiment interface screenshots in Appendix E for an example.*

of the experiment consists of the following tasks: (i) *Information structure computation*: subjects are asked to assess the chance that a computer would guess the color of a randomly selected ball correctly when it makes the optimal guess based on information from a given information structure, and report how confident they are that their assessments are correct;[6] (ii) *Risk and ambiguity attitude elicitation* (3 tasks): we use multiple price lists to elicit subjects' certainty equivalents for three lotteries with equal expected value, including a simple 50/50 lottery, a compound lottery, and an Ellsberg lottery;[7] (iii) *Compound lottery computation*: subjects are asked to compute the

[6]This task provides a measure of subjects' ability to identify the instrumental value of an information structure. The information structure used in this task is one of two pre-selected information structures (randomized at the subject level) that have the same instrumental value but differ in other characteristics. The two structures can be found in the experiment interface screenshots in Appendix E.

[7]The order of the three certainty equivalent elicitation tasks is randomized at the subject level. The tasks are taken from de Clippel et al. (2024).

exact chance that a certain outcome in the compound lottery is realized and indicate how confident they are about their computations.[8] Part 4 contains ten cognitive questions: one Wason selection task, three cognitive reflection test questions, three Raven's test questions, and three belief bias syllogism questions.[9] Details of all these additional tasks can be found in Appendix E.

3.5 Incentives and Implementation Details

The experiment was conducted on Prolific in November 2024 with 400 subjects, using software programmed by the authors in Qualtrics.

All subjects received a base payment of \$6 and an extra \$0.5 if they answered all seven comprehension questions correctly on their first attempt. Comprehension questions are extensive and directly test participants' understanding of the partition design. In particular, participants are tested on the prior, the likelihood of different signal realizations from a structure and the posterior conditional on these signals (see Appendix E). We keep track of the number of mistakes each participant makes in the comprehension questions. We include all participants in the main analysis and discuss the robustness of our results to different subsamples in Section 4.5. 25% of subjects were randomly selected to receive an additional bonus payment based on their decisions in one randomly selected part of the experiment. If part 1 was selected for payment, one of the 15 guessing tasks was randomly chosen and the subject received \$10 if they guessed the color of the selected ball correctly in that task. If part 2 was selected, the software randomly chose one ranking task and drew two information structures from that task. The subject was assigned to receive information from the higher-ranked structure in a final guessing task and received \$10 if guessing correctly in that task. If part 3 was selected, one randomly chosen task in the part determined the subject's bonus payment. If part 4 was selected, the subject received \$0.5 for each correctly answered cognitive question. The median time participants spent on the experiment was 30 minutes and the average (median) bonus payment is about \$7.1 (\$5.5).

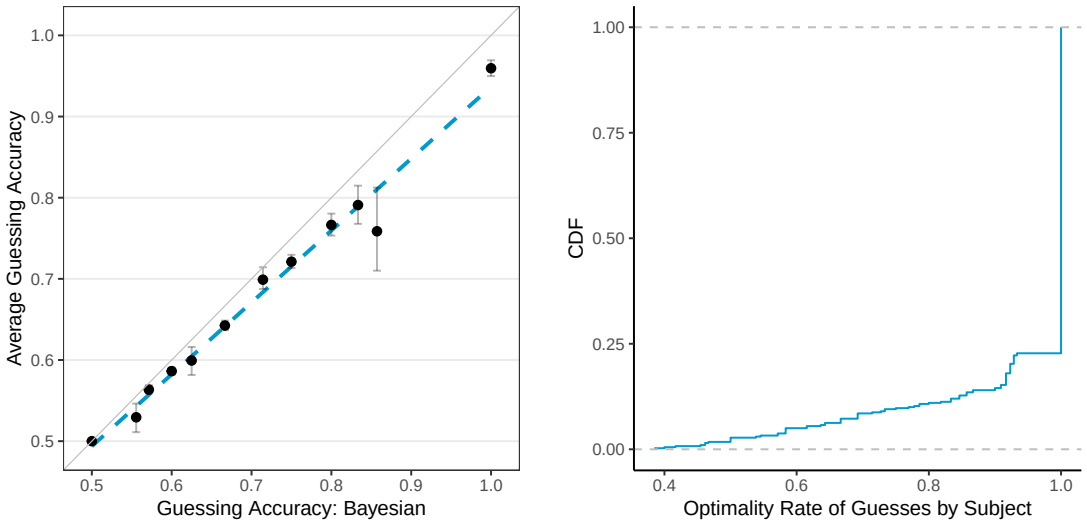
4 Results

4.1 Optimality of Guesses

We begin by confirming that our experimental design is successful in mostly removing inferential errors like failures of Bayesian reasoning when subjects use signals from our information structures

[8] Both the information computation and compound lottery computation tasks are incentivized by the Binarized Scoring Rule (Hossain & Okui 2013). Inspired by Danz, Vesterlund & Wilson (2022), we do not explicitly present the detailed payoff calculation rules to subjects but allow them to access this information if they want.

[9] These questions are the same as those implemented in Oprea & Yuksel (2022).



(a) Guessing Accuracy by Value

(b) Distribution of Optimality of Guesses

Figure 4: Optimality of the Use of Signals *Notes: In panel (a), the best achievable (Bayesian) accuracy level depicts* $\max\{p_b^s, 1-p_b^s\}$ *where* p_b^s *is the posterior on the state being black given the signal* s*. Gray lines denote 95 percent confidence intervals by bootstrapping. The light blue dashed line is the best linear fit, and the gray solid line is the 45 degree line. In panel (b), the optimality of guesses is computed on the individual level and denotes the share of signals for which guess was optimal.*

to inform their guesses. To do this, we study how subjects make use of the information provided to them. We focus on two measures. *Accuracy* is defined as the likelihood of a guess matching the state (the true color of the randomly selected ball). We contrast accuracy of a guess with the best achievable accuracy level given the signal realization. The *optimality* of a guess denotes whether the guess is in the direction (black or white) that maximizes accuracy given the signal realization.[10] To allow for clean interpretation of this measure, we restrict attention to signals (with Bayesian posterior different from 0.5) for which there is a unique optimal guess.

Figure 4a presents data on guessing accuracy, and shows how this varies with the best achievable (Bayesian) accuracy level given the signal. Note that the best achievable depends on the Bayesian posterior associated with each signal realization. Specifically, the best achievable accuracy level is $\max\{p_s^i, 1-p_s^i\}$ (where p_s^i is posterior on the state being black given signal s from structure i). The figure shows that the accuracy of participants' guesses on average closely tracks this best achievable

[10] Let p_s^i be the posterior on the state being black given signal s, the expected accuracy of a guess is p_s^i $(1-p_s^i)$ if the guess is black (white). The optimality of the guess is coded as 0 if $p_s^i > 0.5$ and the guess is white or $p_s^i < 0.5$ and the guess is black, and 1 otherwise.

level.[11,12] This suggests that subjects make near-perfect use of signals to inform their guesses. In particular, participants make mistakes in guesses rarely and even when they do, they do so in cases where these mistakes are not very costly.

The high rate of guessing optimality is particularly clear in Figure 4b, which plots the optimality rate of guesses and depicts the distribution of this measure, the share of signals for which the subject's guess is optimal, computed at the individual level. The vast majority of subjects (77.3 percent) *always* make optimal choices. Calculating across subjects, 94.6 percent of the guesses are optimal. This near-perfect optimality strongly suggests that our methods for providing information typically avoid Bayesian errors (and other errors like confirmation bias), removing a major barrier to measuring information preferences.

Result 1. *Subjects mostly make optimal use of information. Guesses conditional on signal are optimal 94.6% of the time. 77.3% of subjects make optimal guesses 100% of the time.*

4.2 Impact of Instrumental Value on Information Demand

Next, we analyze the demand for information by examining ranking data. Since the specific set of seven structures varied randomly across participants and ranking screens, and participants were incentivized to focus on pairwise comparisons (Section 3.3), we use pairwise comparisons as the unit of analysis.[13,14] This approach allows us to systematically examine how differences in instrumental value, and other structure characteristics, affect relative rankings.[15]

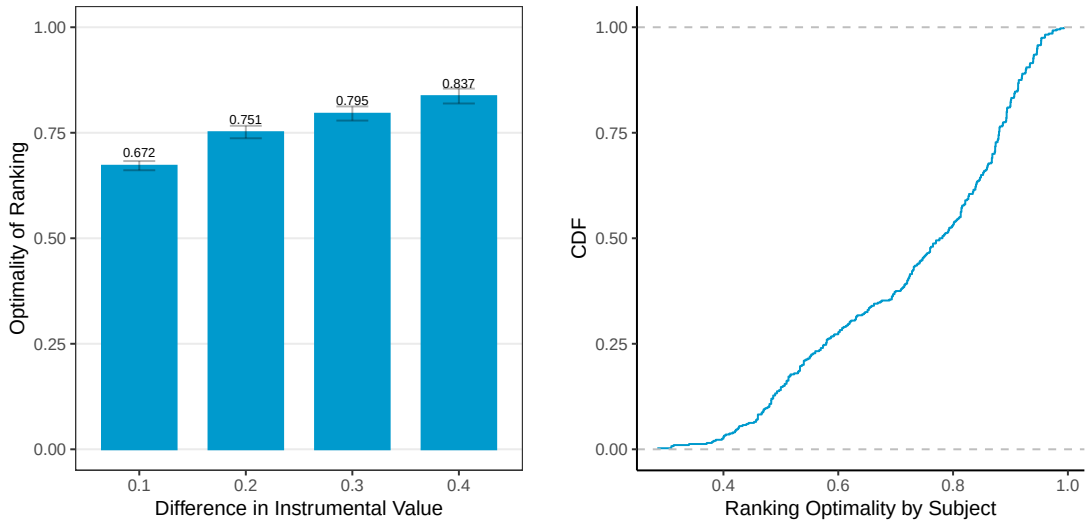
[11] One outlier (as observed in Figure 4a) is a signal that consists of seven balls: six black and one white. Out of 58 participants who randomly observed this signal, eight suboptimally guessed the color of the ball to be white. As discussed in Section 4.5, our main results are robust to removing such participants from the analysis.

[12] Figure 9a in Appendix B plots the distribution of losses due to suboptimal guesses.

[13] While our primary analysis focuses on pairwise comparisons, we can also exploit random variation across ranking screens to examine how the full set of seven structures presented on each screen affects rankings. Table 7 in Appendix B provides such an analysis. Specifically, we use the number of structures that fully reveal the state (out of seven) as a measure of the simplicity of the ranking task to study whether this measure influences the main results reported in the paper. The analysis indicates that our main results remain robust when accounting for how this full-set-level measure moderates the impacts of instrumental value and other identified structure characteristics (discussed in the next section) on demand for information.

[14] We also consider the full ranking of seven structures (rather than pairwise comparisons) as the unit of analysis. Table 8 in Appendix B presents the results of rank-ordered logit regressions that examine how the ranking of a structure within the set of seven is determined by its instrumental value and other identified structure characteristics. These results are consistent with our analysis of pairwise comparisons.

[15] Rankings within each screen impose transitivity over pairwise comparisons. To verify that this does not impact our main results, we replicate our analysis using a single random pairwise comparison from each ranking screen. While this approach substantially reduces the amount of data, it leaves our results unchanged.



(a) Optimality of Ranking by Value Difference (b) Distribution of Ranking Optimality

Figure 5: Optimality of Ranking Information Structures *Notes: We focus on pairwise comparisons. In panel (a) bars depict the rate at which one structure was preferred to another by value difference between the two structures. Gray lines denote 95 percent confidence intervals by bootstrapping. In panel (b) ranking optimality is computed at the individual level and denotes the share of pairwise comparisons (with strict value difference between two structures) for which ranking was optimal.*

Figure 5a provides a first look at how demand for information is shaped by instrumental value. The figure examines all pairwise comparisons between information structures in which one structure is strictly more instrumentally valuable than the other. The bars show the likelihood with which the more instrumentally valuable structure was ranked higher than (as more preferred to) the other. We plot a separate bar for each possible value difference. When the value difference is relatively low, the optimal structure is preferred 67.2 percent of the time. This increases to 83.7 percent when the value difference is very high.

Figure 5b plots the cumulative distribution of optimality of pairwise comparisons at the individual level, focusing on cases where there is a strict value difference between the two information structures. In contrast to our findings, above, on guessing optimality (as seen in Figure 4b), there is substantial heterogeneity across participants in ranking optimality. None of the 400 participants in our data set achieves full optimality of 100 percent. While only 17.8 percent of participants achieve an optimality rate above 90 percent, 46.8 percent of participants achieve an optimality rate above 80 percent. It is also noteworthy that the optimality rate is below 60 percent for 27.3 percent of our participants.[16]

[16]Figure 9b in Appendix B plots the distribution of losses due to suboptimal rankings.

Furthermore, information structures that fully reveal the state, thereby guaranteeing 100 percent guessing accuracy, are ranked as more preferred than strictly lower-value structures in 82.6 percent of relevant pairwise comparisons. Interestingly, this percentage increases to 87.6 percent when we focus on the *simplest* information structure that fully reveals the state. This is the information structure with only two signals; namely, the structure that partitions the 10 balls into two groups, one consisting of six black balls and the other consisting of four white balls. This is a first piece of evidence suggesting that characteristics of information structures beyond their instrumental value impact the way they are valued.

Although these results show that (as expected) instrumental value is a major driver of information demand, they also reveal significant deviation from payoff maximizing behavior. Overall, as Figure 5a shows, in 26.8 percent of relevant cases, subjects rank a strictly less instrumentally valuable information structure higher (i.e., as more preferred) than a more valuable information structure.[17] These mistakes come at significant costs to subjects. By ranking less valuable structures as more preferred, conditional on the pairwise comparison being relevant for payment, on average subjects reduce their guessing accuracy by 17.2 percentage points, leaving approximately \$1.7 on the table.[18] These relatively severe errors in *comparing* information structures contrast sharply with the very low rate of errors in *using* information structures (to guess states) reported in the previous subsection.

Result 2. *Ranking of information structures is strongly influenced by instrumental value, but there are serious deviations from payoff maximizing behavior.*

4.3 Structure Characteristics Beyond Instrumental Value That Impact Demand For Information

In this section, we ask whether there are structure characteristics—attributes of information structures other than instrumental value—that are systematically predictive of demand for information.

To motivate this question and set up our approach to answering it, consider a participant in our experiment tasked with comparing two information structures to determine which one to rank higher. When contrasting one structure with another, participants might be behaviorally drawn

[17] There is no evidence of learning by introspection in the absence of feedback: the error rate remains stable across the eight ranking tasks, ranging from 26.0 to 27.8 percent. However, the amount of time participants spend on each ranking screen declines with experience. Importantly, this is not due to reduced attentiveness: the share of pairwise rankings that participants change relative to the randomly generated original rank fluctuates only between 39.4 and 44.1 percent.

[18] To compute this, we assume optimal use of information by participants.

to structures that are *simpler* relative to some criteria. For example, they might favor structures that are easier to evaluate, perhaps structures that contain fewer signals or generate fewer distinct posteriors. The key idea here is that differences between structures, beyond their instrumental value, may influence demand.

Alternatively, demand for information may be driven not by (or not solely by) *differences* between structures but by aggregate properties of the choice set. For instance, participants might deviate more from optimal behavior when the overall *complexity* of the choice set increases, perhaps when the total number of signals generated by the two structures is high, making it harder to determine which of the two structures is more valuable. The key idea here is that aggregate-level characteristics of the choice set could serve as predictors of participants' likelihood of making ranking errors.

We will examine the influence of both types of characteristics on information demand: those capturing differences between structures and those reflecting properties of the choice set.

Selection of structure characteristics

We do not take an ex ante stance on which characteristics are predictive of information demand. Instead, as outlined in Section 3, we construct a dataset designed to introduce variation across a broad range of characteristics. The complete set of characteristics considered in our analysis, and the variation observed in our data set with respect to each characteristic, can be found in Tables 2 and 3 in Appendix A. In addition to primitive characteristics such as the number of signals and the probability of fully revealing signals, we include information-theoretic measures such as entropy and skewness, as well as relational notions like Blackwell dominance and refinement (on partitions). We also incorporate measures inspired by recent literature that aim to capture the difficulty of comparison or reflect potential simplification strategies.

From a set of 26 characteristics, we identify a small subset that is sufficient to capture the majority of the explainable variation in rankings, using methods similar to the ones used by Enke & Shubatt (2024) to identify the drivers of complexity in lottery choice.[19] We pick these characteristics

[19] The specific characteristics applicable in our setting overlap but also differ from the ones considered in Enke & Shubatt (2024). For instance, since all information structures induce a lottery that can simply be characterized by the likelihood of guessing the state correctly, the most important characteristic identified in their paper, *excess dissimilarity*, is always equal to zero in our setting. For the same reason, the value-dissimilarity measure proposed by Shubatt & Yang (2024) is not directly applicable. Nonetheless, we construct two choice-set-level measures of comparison complexity inspired by these papers. Rather than focusing on the reduced-form lottery over correct guesses, our measures are based on the compound lottery induced by each information structure over guessing

guided by the following procedure. Our goal is to identify interpretable characteristics with large predictive power.

- We fit LASSO regressions for a grid of regularization parameter (λ) values. Higher λ values produce simpler models by driving more characteristic coefficients to zero. We identify characteristics that consistently maintain non-zero coefficients across a wide range of λ values, indicating their robust predictive importance.

- We train classification decision trees across a grid of tuning parameters (which control the complexity of a tree model) to explore which characteristics are most important for predicting information demand. We identify characteristics that appear in the top levels of trees (as early splits), as they represent the strongest predictors.

Both approaches are valuable for guiding the selection of characteristics, particularly given the high correlation among certain characteristics in the set, which introduces multicollinearity. The results of these analyses are presented in Appendix C. These approaches produce strikingly consistent results. Based on these results, we pick four characteristics. These characteristics are easy to interpret and, as will be shown in a later section, account for 77% of the explainable variation (based on structure characteristics) in the ranking data.[20]

The first three of these predictive characteristics have to do with *differences* between the two structures being compared. The first characteristic concerns differences in the number of distinct posteriors the two structures generate:

Difference in number distinct posteriors (Δn_{dp}) Each structure induces a distribution of posteriors. Let n^i_{dp} denote the number of distinct posteriors generated by structure i.[21] If the

accuracy conditional on signal realization. Another key distinction is that both of the above papers focus primarily on characteristics defined at the level of the choice set, whereas our study also examines how differences between the two information structures affect choices.

[20]We focus on only four characteristics in the main analysis because adding a fifth yields only marginal gains in explanatory power, increasing model completeness (discussed in a later section) by less than two percentage points. While focusing on just three variables—Δn_{dp}, $\Delta \tilde{V}$, and $\sum V$ (see definitions below)—also achieves high completeness, this omits the variable with the highest individual explanatory power, Δn_{ru}. For this reason, we report results using all four. We note that, given the intercorrelation among structure attributes, the specific characteristics we identify may not be uniquely important. Our goal is to shed light on how individuals attempt to simplify the cognitively demanding task of comparing information structures.

[21]Consider a structure that partitions the 10 balls into four groups, each of the first three consisting one black ball and one white ball and the last consisting of three black balls and one white ball. Posterior on the randomly selected ball being black is 0.5 for the first three cases and 0.75 in the last case. The number of distinct posteriors generated

agent is comparing structure i to j, $\Delta n_{dp} := n^i_{dp} - n^j_{dp}$.

Similarly, the second characteristic quantifies differences in the number of signals with residual uncertainty:

Difference in number signals with residual uncertainty (Δn_{ru}) A signal with residual uncertainty is a signal that induces a posterior belief with support on multiple states. Let n^i_{ru} denote the number of signals with residual uncertainty produced by structure i.[22] If the agent is comparing structure i to j, $\Delta n_{ru} := n^i_{ru} - n^j_{ru}$.

Since a crucial step in assessing an information structure's value involves determining the distribution of actions it induces and their corresponding payoff consequences in terms of guessing accuracy, these two characteristics (the number of distinct posteriors and the number of signals with residual uncertainty) may intuitively influence the complexity (difficulty) of this evaluation.

The third characteristic selected by our procedure captures the difference in the accuracy induced by the two structures that would be inferred *if the subject were to heuristically ignore the likelihood of signals*. Specifically, it quantifies the value each structure would have if decision makers correctly computed the posterior for each signal but assumed that all signals were equally likely, rather than weighting them by their true likelihood. This is a natural way to heuristically simplify the problem and can be interpreted as an attempt to ease the aggregation challenge described above. The fact that this heuristic difference between structures is predictive of rankings suggests that subjects may be using heuristics like this when evaluating information structures:[23]

Difference in unweighted guessing accuracy ($\Delta \tilde{V}$). We define unweighted guessing accuracy induced by structure i as $\tilde{V}^i := \frac{1}{n_s} \sum_s \max\{p^i_s, 1 - p^i_s\}$, where n_s is the total number of signals generated.[24] If the agent is comparing structure i to j, $\Delta \tilde{V} := \tilde{V}^i - \tilde{V}^j$.

Finally, the fourth characteristic selected by our procedure does not relate to *differences* between the two structures being compared, but rather to the properties of the choice set as a whole. Specifically, controlling for the difference in instrumental value, mistakes in pairwise comparisons

by this structure is two.

[22]Consider the example in footnote 21. For all four signals, there is a positive probability that the ball could be black or white. The number of signals with residual uncertainty is four.

[23]Note that this occurs despite the fact that comprehension questions confirm participants can identify differences in signal likelihoods when asked to.

[24]Consider the example in footnote 21. Optimal guesses would induce a guessing accuracy of 0.5 for the first three signals and 0.75 for the last signal. Unweighted guessing accuracy is $\frac{0.5+0.5+0.5+0.75}{4}$. Note that this differs from the true (weighted) induced guessing accuracy of 0.6.

are more likely to occur when the total instrumental value of the two structures is low.[25]

In other words, participants' responsiveness to differences in instrumental value increases with the overall value of the choice set, suggesting that comparisons between high-value structures are cognitively easier. This is intuitive: high-value structures are typically associated with signals that fully reveal the state, while low-value structures tend to generate ambiguous signals, making comparisons between information structures more difficult.[26]

Total instrumental value ($\sum V$). If the agent is comparing structure i to j, let $\sum V = V^i + V^j$, where V^i and V^j are the instrumental values of structures i and j, respectively.

Impact of structure characteristics on rankings

The four graphs in Figure 6 illustrate the impact of these four selected characteristics on the ranking of information structures. These graphs plot the rate at which one structure is preferred to another as a function of the value difference between the two structures. Under optimal behavior, the resulting curve would resemble a step function, rising from zero to one at zero.

Focusing on the first three characteristics (differences in number distinct posteriors, number signals with residual uncertainty and unweighted guessing accuracy), panels (a)-(c) of Figure 6 show that, when the difference in instrumental value is held constant, participants are more likely to prefer structures that induce relatively (i) fewer distinct posteriors, (ii) fewer signals with residual uncertainty, and (iii) higher unweighted guessing accuracy. These effects are substantial overall and particularly pronounced when the instrumental value difference between the two structures is small.

In all pairwise comparisons where one structure has strictly higher instrumental value than the other, participants are more likely to rank it as preferred when it induces weakly fewer distinct posteriors. Specifically, the likelihood of selecting the higher-value structure increases from 65% to 76% compared to cases where it induces strictly more distinct posteriors. An even larger increase is observed when the higher-value structure induces weakly fewer signals with residual uncertainty,

[25] Among choice-set-level variables, total entropy informativeness of the two structures is selected before total instrumental value. We use the latter because (i) it is highly correlated with the former (correlation coefficient of 0.96) and thus offers similar predictive power (model completeness changes by less than one percentage point); and (ii) it is easier to interpret in our setting. As demonstrated in Figure 11 of Appendix B, reproducing panel (d) of Figure 6 using total entropy informativeness rather than total instrumental value produces nearly identical results.

[26] While our analysis does not pick up either of the two measures of comparison complexity included (see footnote 19) as a characteristic most predictive of rankings, as shown in Figure 10 of Appendix B, these measures produce a similar attenuation effect. This reinforces the idea that some pairs of information structures are easier to compare than others.

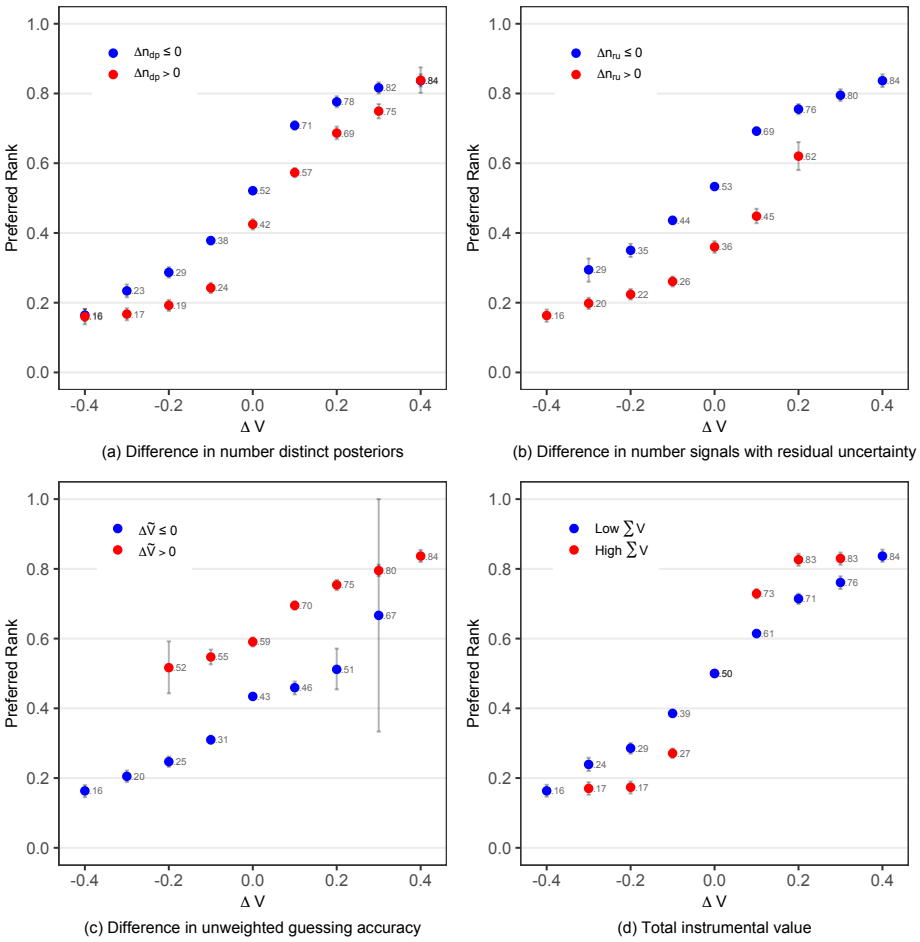


Figure 6: Pairwise Ranking of Information Structures by Value Difference *Notes: We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

with selection likelihood rising from 48% (in cases with strictly such signals) to 74%. Finally, the likelihood of choosing the higher-value structure increases from 46% to 75% when it is associated with strictly higher unweighted guessing accuracy compared to cases where it has a weakly lower accuracy.

Figure 6d depicts the impact of the last selected characteristic—the total instrumental value of the two structures—on the optimality of pairwise rankings. Since this characteristic operates at the aggregate level, reflecting properties of the choice set, its influence produces a distinct effect, visible in the figure. Specifically, deviations from optimal behavior are more pronounced when the total instrumental value, $\sum V$, is low, that is, when $\sum V$ falls weakly below the median value in

our sample. In particular, the participants' responses to differences in instrumental value more closely resemble a step function, in accordance with theoretical predictions, when $\sum V$ is high (strictly above the median value in our sample). By contrast, for lower values of $\sum V$, the response pattern more closely follows an attenuated logistic S-shape. Overall, when we focus on all pairwise comparisons where one structure has strictly higher value than the other, the likelihood of optimal ranking increases from 70% to 78% when we compare low $\sum V$ to high $\sum V$ cases.[27]

Thus, we find that different characteristics have different qualitative effects on rankings. Characteristics related to differences between structures produce what looks like systematic aversion to those characteristics (perhaps driven by complexity aversion). They also point to specific ways in which individuals may rely on simplifying heuristics when evaluating information structures. Our one selected characteristic related to the choice set as a whole produces instead an attenuated response to instrumental value. These are distinct effects of complexity documented in the prior literature, and our results suggest that each plays an important role in how individuals evaluate and compare information structures.

Completeness

To further quantify the impact of these characteristics on the ranking of information structures, we evaluate the *completeness* of alternative models with and without these characteristics. Formally, following the definition in Fudenberg, Kleinberg, Liang & Mullainathan (2022), the completeness of a model is given by $\frac{\varepsilon_{base}-\varepsilon_{model}}{\varepsilon_{base}-\varepsilon_{irreduciable}}$, where ε_{base} is the prediction error of a baseline model, ε_{model} is the prediction error of the model under consideration, and $\varepsilon_{irreduciable}$ represents the irreducible error. In other words, completeness measures the model's reduction in prediction error relative to the baseline, scaled by the maximum achievable reduction in prediction error. We focus on all pairwise rankings where there is a strict value difference between the two structures. For the baseline, we fit a logit model where we use the only characteristic that is theoretically relevant for rankings: which structure is of higher value. To compute $\varepsilon_{irreduciable}$ we use predictions from a fully-

[27] We look closely at two potential drivers of this result. (1) If participants were more careful when ranking their top choices than their bottom choices, this could generate the observed pattern. To test this, we separate the top and bottom three structures in each participant's final ranking and estimate, for each set, the share of pairwise reversals relative to the original ranking presented. Consistent with this concern, we find more changes on average among the top three. However, a substantial share of participants (32.8 percent) show the opposite pattern—making (weakly) more changes among the bottom three—and our main results replicate within this subset. (2) As documented in the previous section, participants make fewer mistakes when comparing maximally valuable information structures that induce guessing accuracy of 100 percent to others. Removing such comparisons from the analysis does not eliminate the predictive power of total instrumental value.

Table 1: Impact of Selected Structure Characteristics on Completeness

Structure Characteristics	Model Completeness
Difference in number distinct posteriors, Δn_{dp}	0.27
Difference in number signals with residual uncertainty, Δn_{ru}	0.56
Difference in unweighted guessing accuracy, $\Delta \tilde{V}$	0.32
Total instrumental value, $\sum V$	0.27
$\Delta n_{dp} + \Delta n_{ru} + \Delta \tilde{V} + \sum V$	0.77

Notes: Completeness quantifies the model's reduction in prediction error relative to a baseline, scaled by the maximum achievable reduction in prediction error. The baseline is a logit model that includes an indicator variable for the structure with higher instrumental value. The models reported above incorporate this indicator variable along with the specified additional variables. The best achievable reduction is calculated using predictions from a convolutional neural network (CNN). Each model is trained on 80 percent of the data and prediction error is computed on the remaining 20 percent. The reported values represent the average results from five independent repetitions of this procedure.

flexible convolutional neural network (CNN). The CNN is trained to predict ranking of information structures based on raw characteristics of the structures, namely the specific partitioning of the 10 balls corresponding to each structure, plus the grand set of structure characteristics that were inputs to LASSO and decision tree.

We compare these benchmark models to alternative logit models that augment the baseline model with the four characteristics discussed above. We use five-fold cross validation to test performance of the models. Iteratively, we divide our data set into two parts: a training set and a test set. All models are estimated on the former set (corresponding to 80 percent of the data) and their performance is evaluated on the test set. Our measure of error is log loss.[28]

Table 1 presents completeness measures for alternative logit models. The table suggests that each characteristic individually has a sizable impact on completeness. For instance, adding the difference in the number of signals with residual uncertainty, Δn_{ru}, alone increases completeness to 56 percent. The corresponding values for the other characteristics are 27 percent (difference in the number of distinct posteriors, Δn_{dp}), 32 percent (difference in unweighted guessing accuracy, $\Delta \tilde{V}$), and 27 percent (total instrumental value, $\sum V$).[29,30] When all four characteristics are included,

[28] Assume the prediction of the model is that structure i will be chosen over j with probability p_{ij}, the log loss is $-\ln p_{ij}$ for a particular observation i is ranked as preferred to j; the loss $-\ln(1-p_{ij})$ if j is ranked as preferred to i.

[29] By comparison, including ΔV in the model—representing the difference in instrumental value, and thus the cost of making a mistake—yields a completeness of 29 percent.

[30] Table 5 in Appendix B presents model completeness when including all 26 structure characteristics together and when including each individually.

completeness improves to 77 percent. These results suggest that identifying the structure with higher instrumental value, along with the four selected characteristics, is sufficient to capture a substantial portion of the explainable variation in pairwise rankings.[31]

Explanatory power of selected characteristics on the individual level

The evidence so far suggests that a small set of characteristics predicts rankings at the aggregate level. We next show that these characteristics are also highly predictive for a majority of participants at the individual level. To explore this, we focus on the first three characteristics, which capture differences between two structures in a pairwise comparison. We examine whether alternative decision rules based solely on these characteristics are more predictive at the individual level than the theoretical benchmark, which assumes that the structure with the higher instrumental value should always be chosen.[32] We focus on pairwise comparisons where one structure is of strictly higher value than the other and consider the following natural decision rules:

Optimal: Choose the structure with the higher value.

Alternative 1: Choose the structure that induces a lower number of distinct posteriors.[33,34]

Alternative 2: Choose the structure that induces a lower number of signals with residual uncertainty.

Alternative 3: Choose the structure that induces a higher unweighted guessing accuracy.

A striking 67 percent of our participants ranked information structures in a manner that aligns more closely with one of these (suboptimal) alternative rules rather than the optimal rule! Among these participants, a majority (58 percent) exhibited behavior most consistent with the third alternative rule, which conditions on unweighted guessing accuracy, suggesting that many subjects may simply be heuristically ignoring the likelihood of signals when evaluating information structures. The remaining participants were nearly evenly split between the first two rules, with 22 percent and 23 percent aligning with each, respectively.[35]

[31] Table 6 in Appendix B presents regression analysis on the impacts of the four structure characteristics on ranking preference.

[32] It is not possible to formulate a simple decision rule based on the last characteristic, $\sum V$, which is defined at the choice set level.

[33] In all rules, cases of indifference are assumed to be resolved randomly. But, because we focus on subset of data where there is always a strict value difference between the two structures, there are no cases of indifference with respect to the optimal rule.

[34] Note that alternative versions of this rule (and other rules) that condition on both the selected characteristic and instrumental value will fit our data better. Here, we focus on the simplest possible version.

[35] The shares do not sum to exactly 100 percent because the ranking behavior of a small number of participants was equally and maximally consistent with multiple rules.

We summarize the results from this section below:

Result 3. *We find that several characteristics of information structures (beyond their instrumental value) are predictive of mistakes in rankings. Four key characteristics are sufficient for capturing the majority of the explainable variation in rankings.*

4.4 Impact of Structure Characteristics Relative to Individual Characteristics

To help contextualize earlier results and understand other drivers of mistakes in the evaluation of information structures, we next consider characteristics of decision makers that influence ranking behavior. To do this, we follow the same procedure described in the previous section for structure characteristics to identify participant characteristics that are most predictive of ranking behavior (see Appendix C for a detailed analysis and Table 4 in Appendix A for the full list of characteristics considered). Below, we list the decision-maker characteristics we select.[36]

Raven score This can be interpreted as a measure of intelligence (or alternatively as a measure of engagement with the experiment).

Ambiguity aversion[37] This is intended to capture the participant's attitudes towards ambiguity.

Absolute mistakes in information structure valuation This provides us with a measure of a participant's ability to infer (expected) guessing accuracy induced by an information structure.

Finally, although we do not consider this as a participant characteristic that can be interpreted independently of the ranking task, we also include our elicited measure of confidence.

Confidence in the optimality of rankings This gives us a measure of how certain participants are in the optimality of their rankings.

To assess the predictive power of the participant characteristics above and compare them to the selected structure characteristics, we conduct the following analysis, with results presented in Figure 7. Focusing on all pairwise comparisons where one structure is of strictly higher value than the other, we compute the impact of a one standard deviation increase in each variable on ranking optimality. For ease of comparison, Figure 7 reports the absolute values of these effects, with line color indicating whether the effect is positive (blue) or negative (red).

[36] Among these characteristics, the only one not selected as highly predictive of behavior is ambiguity aversion. We nonetheless include it because it is qualitatively distinct from the other characteristics and allows us to make connections to recent work exploring the relationship between ambiguity and complexity.

[37] We measure ambiguity aversion as the difference between a participant's certainty equivalent for a lottery with a known 50% payout probability and a lottery with an unknown payout probability of x. See Section 3 for implementation details.

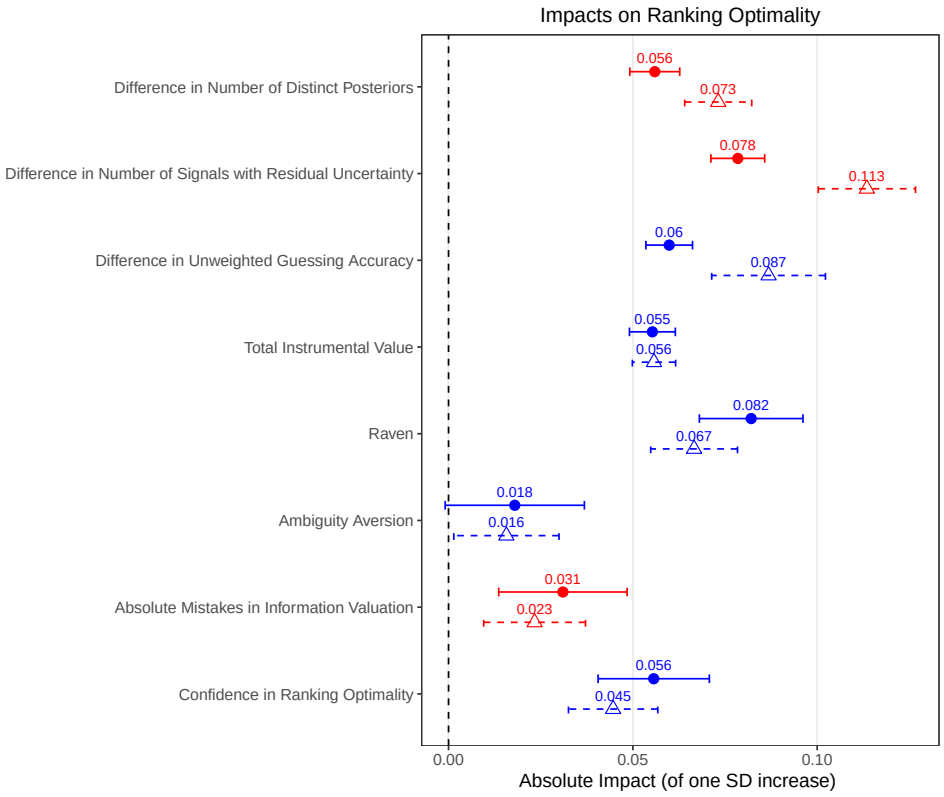


Figure 7: Impact of Structure and Participant Characteristics on Ranking Optimality *Notes: The figure plots the absolute impact of one standard deviation increase in a structure characteristic or participant characteristic on ranking optimality, with the color denoting whether the effect is positive (blue) or negative (red). The solid dots show the results when we focus on all pairwise comparisons where one structure is of strictly higher value than the other; the non-solid triangles show the results when we focus on comparisons where one structure is exactly more instrumentally valuable than the other by a 0.1 increment in guessing accuracy. Short lines denote 95 percent confidence intervals.*

Several variables stand out in terms of their predictive power. The number of signals with residual uncertainty and the Raven score are the strongest predictors of deviations from optimality, with a one-standard-deviation increase in these variables altering ranking optimality by approximately 7.8 and 8.2 percentage points. These are followed by the remaining structure characteristics and participant confidence in their rankings, each of which has an effect of approximately 6 percentage points. The remaining participant characteristics exhibit smaller impacts on ranking behavior.

These findings highlight both the role of individual differences—some participants are more cognitively equipped and more willing to identify higher-value structures—and the persistent influence of structure characteristics despite these individual differences. Notably, the effect of increasing the number of signals with residual uncertainty on ranking optimality is comparable in magnitude to that of participants' Raven scores, underscoring the significance of structure-driven complexity in

the evaluation of information structures.

Interactions between structure characteristics and participant characteristics

Since these results highlight the importance of heterogeneity in our participant population, a natural question is whether structure characteristics are more predictive of rankings for some types of participants than others? Answering this question is challenging, as participant characteristics and structure characteristics can interact in many different ways to jointly influence ranking optimality. Our goal in this section is to show that, to the degree that such interactions are present, they might not arise in ways that are anticipated.

To examine whether the behavior of certain participant types is more strongly predicted by structure characteristics, we conduct the following analysis. We focus on pairwise comparisons in which deviations from optimal behavior are most frequent, specifically, cases where the difference in instrumental value between the two structures is small.[38] In these instances, selecting the suboptimal information structure results in a 10 percentage points reduction in guessing accuracy. To measure the impact of each structure characteristic f, we compare the rate at which rankings are optimal when f is strictly above the median value in our sample to when it is weakly below.[39] This yields an intuitive measure of the extent to which characteristic f influences ranking optimality. We compute this measure separately for several subsamples of our participant population. For example, we divide participants into two roughly equal groups based on their Raven scores. This allows us to examine whether the impact of characteristic f varies by participant type, where type is defined in terms of intelligence as measured by the Raven score.

The results of this analysis are presented in Figure 8, which illustrates the impact of each selected structure characteristic on ranking optimality. The top panel separates participants based on their Raven scores, as described earlier. Two key observations emerge: (1) All four characteristics significantly influence optimality for both subgroups (high and low Raven scores). (2) The impact is notably larger for participants with high Raven scores, with some of these characteristics affecting ranking optimality nearly twice as much in this group compared to those with low Raven scores. This may seem surprising, as the group that is, on average, more optimal is also the one whose behavior is more influenced by structure characteristics that are irrelevant to optimality.[40]

[38] The same analysis considering all pairwise comparisons is shown in Figure 16 of Appendix B.

[39] This corresponds to the vertical gap between the red and blue dots in the four graphs of Figure 6 when $\Delta V = 0.1$.

[40] Figure 17 in Appendix B studies more closely the estimated impact of these selected structure characteristics as a function of the overall optimality rate in rankings. The non-monotonic relationship depicted in the figure is reminiscent of similar findings in Gonçalves (2024).

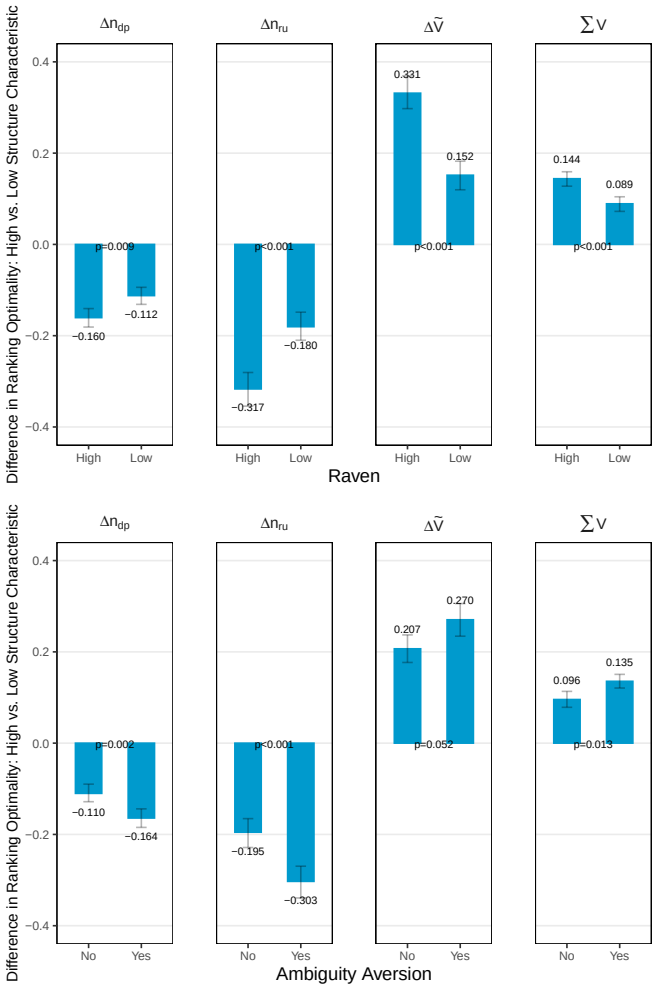


Figure 8: Interaction Between Participant and Structure Characteristics *Notes: Analysis is on pairwise rankings where* $\Delta V = 0.1$. *Bars represent the difference in ranking optimality when characteristic* f *is strictly above the median value in our sample versus weakly below, separated by participant type. Gray lines denote 95 percent confidence intervals by bootstrapping. P values are computed using two-sample t-tests.*

A similar pattern emerges when we separate participants based on other individual characteristics. These results are presented in Figure 15 of Appendix B. Here, we summarize the main takeaways. Dividing participants into two groups based on the magnitude of their mistakes in the information evaluation question yields similar results. Participants who make no mistakes or only small mistakes (formally, those whose mistake size is weakly below the median in our sample) tend to rank information structures more optimally. However, the impact of selected structure characteristics on their ranking optimality is at least as strong (or stronger) as for those with larger mistakes.[41] Remarkably, we observe the same pattern with respect to participants' confidence in

[41] The only exception is the difference in unweighted guessing accuracy, $\Delta\tilde{V}$, which has a greater effect on ranking optimality for participants with higher mistakes.

the optimality of their rankings. Participants who are more confident in their rankings—who are also, on average, more optimal—are the ones whose behavior is more responsive to these structure characteristics.

Overall, these results demonstrate that while participant types matter for ranking behavior, the extent to which structure characteristics influence optimality does not diminish among those who perform better on average. Instead, participants who are generally more aligned with optimal rankings—due to intelligence, attentiveness, or confidence—are also the ones whose behavior is most shaped by structure characteristics that should, in principle, be irrelevant to optimality. This suggests that deviations from optimal rankings are not simply driven by lack of ability or effort but may instead reflect systematic tendencies in how even high-performing individuals process the value of different information structures. We discuss the interpretation of this finding further in Section 5.

Separating participants based on their ambiguity attitudes reveals an intriguing pattern, as shown in the bottom panel of Figure 8. Ranking behavior is more influenced by the selected characteristics for ambiguity-averse participants compared to those who are ambiguity-neutral or ambiguity-loving. This may point to a role for ambiguity attitudes in driving complexity aversion in the evaluation of information structures (see de Clippel et al. (2024) for recent evidence on the link between ambiguity attitudes and complexity aversion).

Result 4. *Structure characteristics are similarly (or more) predictive of rankings than individual-level characteristics that could influence demand for information. Moreover, their impact on rankings is greater for participant subgroups that are more likely to exhibit optimal behavior and for those participants who are ambiguity averse.*

4.5 Robustness of Findings and Treatments on Mechanisms

Robustness across subsamples. The findings in the previous section indicate that structure characteristics exhibit stronger predictive power over rankings for certain subpopulations. To ensure that our main results are not driven by inattentiveness, lack of engagement, or general confusion with the experiment, we replicate Figure 6 for different subsamples in Appendix B. Figure 12 restricts the analysis to participants who answered all comprehension questions correctly on their first attempt; Figure 13 focuses on participants who made perfectly optimal guesses (across all 15 questions) in the guessing portion of the experiment. The patterns in these subsamples closely align with those observed in the full sample, reinforcing the robustness of our main findings.

Connections to guessing behavior. While results in Section 4.1 indicate that participants

rarely make mistakes in guessing questions, such mistakes do occasionally occur. We examine the extent to which structure characteristics predict these errors, as well as their impact on response times, that is, the time it takes participants to determine their guess.[42] Results from this analysis are reported in Figure 14 of Appendix B. Overall, we find guesses are more likely to be optimal for information structures with high instrumental value, fewer signals with residual uncertainty, and higher unweighted guessing accuracy. The impact of these variables on response times follows the opposite pattern: participants respond more quickly when structures have high instrumental value, fewer signals with residual uncertainty, and high unweighted guessing accuracy.[43,44] Namely, we find that structure characteristics influencing ranking behavior appear to be linked to those that make it more challenging for participants to determine optimal guesses given a signal.

Laboratory Experiment. An earlier version of this paper reported results from a laboratory experiment with a similar design.[45] A key difference is that all participants in the lab study saw the same 16 information structures, chosen to vary certain structure characteristics (e.g., number of signals, entropy informativeness) while holding instrumental value constant. This limited the variation in structure characteristics. The current study addresses this by expanding the set of information structures from 16 to 237.

Behavior in the laboratory study closely mirrors that of the main study.[46] Most importantly, as shown in Figure 19 of Appendix D, the four key structure characteristics identified in the main study similarly predict ranking behavior in the lab data. These findings reinforce the robustness of the main study's results, confirming that they are unlikely to be driven by specific details of the

[42]Since we analyze guesses conditional on signal realization, we can also directly assess whether characteristics of the signal itself, rather than just structure-level characteristics, influence guessing accuracy and response times.

[43]We do not find the number of distinct posteriors associated with an information structure to significantly affect guessing accuracy or response time.

[44]Similar patterns emerge with participant-level characteristics. Guessing optimality is higher among participants with higher Raven scores, greater ambiguity aversion, and fewer errors in information evaluation. Response times follow the opposite trend, increasing with the same changes in these variables. A similar relationship is observed with confidence in rankings, where participants who are more confident in the optimality of their rankings make more optimal guesses and take a shorter time in determining these guesses. In fact, as shown in Figure 14 in Appendix B participant-level characteristics have a much greater impact on guessing behavior compared to structure-level or signal-level variables.

[45]Appendix D provides further details.

[46]For instance, participants guess optimally 98.8% of the time and rank less valuable structures above more valuable ones in 22.8% of relevant pairwise comparisons. These rates closely align with those from the main study (94.6% and 26.8%) and with those from a dedicated replication study using the same interface and subject pool as the main study but limited to the 16 structures used in the lab study (93.2% and 25.9%).

experimental interface or the characteristics of the online subject pool.[47]

In addition to gathering *ranking* data (as in our main study), we also asked subjects to explicitly *value* information structures via elicited willingness to pay (WTP) measures. We found that explicit WTP measures differ from ranking results in two interesting ways. The first is that WTP responses were considerably noisier than rankings and violated instrumental value in 39.7% of relevant pairwise comparisons, substantially exceeding the 22.8% violation rate observed in ranking data. This finding mirrors other evidence suggesting that choice tasks (a'la our ranking task) are often less noisy in certain respects than explicit valuation tasks (see, e.g., Bouchouicha et al. (2025)). The second is that WTP and rankings do not always agree: WTP measures rank structures in ordinal agreement with ranking data only 74.9% of the time. What's more, the impacts of the four structure characteristics on WTP do not always align with those observed in the ranking data (see Figures 24 and 25 in Appendix D). This finding mirrors a long line of literature (especially in risky choice) suggesting that the cognitive processes involved in choice tasks (a'la our ranking tasks) can be quite different from explicit valuation tasks and often result in different implied preferences (see, e.g., Lichtenstein & Slovic (1971), Grether & Plott (1979)).[48]

We also ran several treatment variations in the laboratory study that are useful in understanding the mechanism behind our main results. First, we examined to what extent difficulties in evaluating information structures are due to the difficulty of forecasting how those information structures will be used to improve behavior (see Niederle & Vespa (2023) for review of the literature on failures in contingent reasoning like these). To test this, in the laboratory study we included a Reverse treatment in which participants made guesses for all possible signal realizations before ranking the structures, and were later reminded of those guesses during the ranking task. Under the hypothesis

[47] In an earlier version of this paper that was based on the laboratory data, we found that differences in entropy informativeness were an important predictor of pairwise rankings. However, this finding turns out to be special to the restricted set of 16 information structures selected for that earlier study (indeed, we find the same result in a Replication study (described in Appendix D) that uses the 16 structures from the laboratory study but uses the Prolific subject pool and the experiment interface of the main study). When we dramatically expand the set of information structures to 237, we find instead that Δn_{ru}, Δn_{dp}, and $\Delta \tilde{V}$, rather than differences in entropy informativeness, are better and more robust predictors of behavior.

[48] We elected not to include WTP elicitations in our main study for several reasons. First, ranking data is less noisy (and closer to optimal behavior) and we decided it would therefore be a better fit for an online sample. Second, we think ranking data is more similar to the types of choices over information structures people typically make in the field (people are rarely asked to attach explicit monetary values to information). Finally, in comparison to the lab study, we wanted to study preferences over far more information structures, which is relatively easy with ranking tasks but would be difficult using WTP elicitations.

that our main findings are driven by subjects' failure to contingently reason about how different signals induce different guesses, we would expect this treatment to eliminate or at least reduce the severity of our results. However, the Reverse treatment had no effect on reducing ranking errors: the same structure characteristics continued to strongly influence rankings (see Figures 21 and 22 in Appendix D).

Second, we examined whether non-standard preferences over risk, loss, or the timing of uncertainty resolution could explain the systematic deviations from optimal behavior observed in the ranking of information structures. To test this, the laboratory study included a *No Uncertainty* treatment, which eliminated uncertainty using methods similar to those used in Martinez-Marquina, Niederle & Vespa (2019) and Oprea (2024*b*) while preserving the cognitive demands of the task. Despite the absence of risk and the fact that each structure's value was objectively equal to its instrumental value, participants' ranking patterns remained similar (see Figure 23 in Appendix D that replicates Figure 6). This suggests that preferences over timing or risk are unlikely to be the primary drivers of deviations from optimal ranking behavior in our data.

5 Discussion

According to standard economic theory, the demand for information is driven by its instrumental value—the extent to which it improves decision-making in relevant tasks. We test this hypothesis and study whether *other* characteristics of information structures also systematically influence demand. To do this, we use a novel experiment that measures participants' relative preference rankings for hundreds of distinct information structures that differ not only in their instrumental value but also in a number of other formal characteristics. These non-instrumental characteristics plausibly influence the type of information processing required to compare information structures and therefore may shape the difficulty of identifying optimal sources of information.

As predicted by standard theory, we find that instrumental value plays a major role in shaping demand for information structures. However, in 27% of cases, participants prefer less instrumentally valuable structures over more valuable ones, leading to substantial earnings losses. These errors in *assessing* information structures contrast sharply with subjects' near-perfect ability to *use* signals from the same information structures. Our results therefore suggest that even when information is perfectly easy to use (ex post), assessing the relative value of information structures (ex ante) can nonetheless be difficult and prone to error.

Importantly, we find that these errors in evaluating information structures have a significant, predictable structure. In particular, we identify four key characteristics of information structures

that predict demand errors. Participants seem to be systematically *averse* to information structures that (i) produce more distinct posteriors and that (ii) generate more signals that retain uncertainty, that is, signals that fall short of fully revealing the state. Since evaluating information structures with these characteristics seem intuitively to require more intensive information processing, we speculate that this may be evidence that a form of complexity aversion shapes the demand for information structures. We also find (iii) that subjects prefer information structures that would (counterfactually) have higher instrumental value if all signals had been equally likely. We hypothesize that this pattern arises because many participants neglect differences in signal likelihood when computing the relative value of structures, causing them to systematically overestimate the value of structures that would have been more useful if indeed all signals had been uniformly distributed. This appears to reflect a natural approximation strategy—a heuristic to simplify the complex task of fully assessing the value of information structures. Finally, we find that the total instrumental value of the two structures being compared predicts, not systematic aversion (as in the first three predictors), but instead *insensitivity* to the difference in instrumental value between structures.

Putting these drivers together, we interpret our findings as suggesting that information processing constraints are a significant secondary driver of demand for information in addition to instrumental value, often causing decision makers to prefer inferior sources of information. These distortions appear to arise because evaluating the value of one information structure relative to another is a cognitively demanding task. It requires (i) considering all possible signal realizations, (ii) identifying the optimal action conditional on each signal, (iii) computing the associated payoffs, and (iv) aggregating over these outcomes to arrive at an overall value. The complexity of this process increases systematically with certain structure characteristics, introducing predictable patterns of deviations from optimality. While the particular characteristics we identify may not be uniquely important, given the intercorrelation among structure attributes, they nonetheless reveal how individuals attempt to simplify a cognitively demanding task. Because deviations from optimal behavior exhibit systematic structure, our findings offer useful insights into how cognitive constraints shape information demand, suggest a path toward more predictive models of how individuals select and value information structures, and have implications for efficient information design and provision.

Interestingly, we find that there is a counterintuitive relationship between the predictive effects of these structure characteristics and the cognitive sophistication of our participants. We find that our four structural predictors of errors are in fact *more* predictive for more sophisticated subjects (e.g., subjects who score highly in Raven tasks), despite the fact that more sophisticated subjects

also make fewer errors in ranking structures. While counterintuitive, this result does not constitute a paradox.[49] Less sophisticated subjects make errors at a higher rate, but they also make errors that are less predictable and plausibly more random, while more sophisticated subjects make errors that are better organized by systematic characteristics of information structures. We think this may point to a potentially important insight about the role complexity aversion plays in behavior. Recognizing and responding to complexity itself requires effort and insight. In our experiment, such effort appears concentrated among more sophisticated participants, making even their mistakes more predictable than the comparatively random errors of less sophisticated ones.[50] Additionally, we find that structure characteristics are more predictive of deviations from optimality among participants who are ambiguity averse, providing further evidence of a link between responses to complexity and ambiguity, as highlighted in the recent literature.

Our experimental design deliberately abstracts from other forces that may influence information demand.

First, we present information structures to participants in a way (using the partition design) that makes their formal characteristics—such as the number and likelihood of signals—transparent and their instrumental value relatively intuitive to understand. This is confirmed in our results: the majority of decisions align with instrumental value, allowing us to study systematic mistakes in an environment where decision making is otherwise highly optimal.[51] As noted earlier, the way we present information structures also effectively removes errors in the *use* of information. We do not claim that such errors are unimportant for information demand. Rather, our point is that even when participants can correctly interpret and use signals, identifying the relative value of information structures remains fundamentally challenging. This difficulty arises because evaluating an information structure requires reasoning through multiple contingencies—assessing how actions

[49] In a strategic setting (i.e., repeated p-beauty contest), Gill & Prowse (2022) find subjects are heterogeneous in their responses to the complexity of strategic situations, with higher level-k (more sophisticated) subjects being more likely to respond strongly to complexity. This is broadly parallel to our finding that more sophisticated subjects' demand for information is more systematically shaped by non-instrumental (complexity) characteristics of information structures.

[50] This perspective resonates with the recent work by Agranov, Schotter & Trevino (2025), who propose that complexity is subjective and demonstrate systematic heterogeneity in how people perceive and engage with decision problems. Our finding that sophisticated subjects make more predictable errors aligns with their framework, as systematic responses to complexity, even erroneous ones, require some degree of cognitive engagement with the problem.

[51] We also focus on the simplest possible decision problem (a guessing task) to study information demand. An interesting direction for future research is to study how the complexity of the decision problem interacts with difficulties in identifying instrumental value.

and payoffs depend on each possible signal realization and integrating these into an overall value. While our design paradigm is not intended to be representative of how information structures are typically encountered in applied settings, it provides a useful benchmark for isolating these difficulties underlying the evaluation of information in its simplest possible case.

Second, the deliberately abstract framing of our experiment offers both advantages and limitations. On the positive side, this abstraction allows us to tightly control several formal characteristics of information structures, most importantly instrumental value, providing the kind of experimental precision that enables relatively sharp conclusions about the drivers of information demand. At the same time, by design, our experiment omits other forces that may shape information demand in applied settings, many of which are likely to be context-dependent. For instance, our results cannot speak to how voters might be drawn to information sources that reinforce their ideological identity, or how employees might seek out feedback that preserves self-esteem rather than maximizes learning. These motivations, while important in many real-world environments, operate alongside the cognitive and instrumental considerations we isolate. Extending our methods to richer settings in which such "non-cognitive" factors also play a role, and comparing their influence to the regularities we identify, seems a promising next step in this line of research.

Despite this, the non-instrumental drivers of informational preferences we identify are intuitive and appear relevant to real-world information environments. Many decision settings, such as a physician choosing among diagnostic tests, an investor selecting between forecasting models, or a teacher deciding which type of student assessment data to rely on, require selecting among information sources that differ not only in expected usefulness but also in how complex they are to evaluate. Our findings suggest that decision makers might be drawn to *simple* information sources and characteristics like posterior dispersion, residual uncertainty, and signal asymmetries may shape which information sources are considered simple and more useful. For example, a doctor may choose a diagnostic test with binary results over a probabilistic one not because the latter is harder to act on, but because its value is harder to evaluate in advance. Similarly, a recent literature on AI-assisted decision making shows that *coarsening* algorithmic recommendations not only improves decision quality but also increases demand for such information (see Dreyfuss & Hoong (2025)). Likewise, sources that leave residual uncertainty, such as diagnostic tests or models that fail to produce definitive results, may appear less valuable, not because they convey less information, but because their partial signals are harder to interpret ex ante. Finally, people may favor information sources which appear highly useful when asymmetries in signal probabilities are ignored. For example, an investor comparing forecasting models may focus on how often each

model "gets it right" across different market conditions, implicitly weighting all conditions equally even though some occur far more frequently than others. Taken together, these examples illustrate how the systematic patterns we observe in a highly controlled experimental setting can shed light on why decision makers in practice may favor simpler or more transparent sources of information, even when more informative alternatives are available.

References

Agranov, M., Schotter, A. & Trevino, I. (2025), Complex for whom? an experimental approach to subjective complexity, working paper.

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

Andries, M. & Haddad, V. (2020), 'Information aversion', *Journal of Political Economy* **128**(5), 1901 – 1939.

Ba, C., Bohren, J. A. & Imas, A. (2023), Over- and underreaction to information, working paper.

Bouchouicha, R., Oprea, R., Vieider, F. M. & Wu, J. (2025), Is Prospect Theory Really a Theory of Choice. Unpublished manuscript.

Caplin, A. & Leahy, J. (2001), 'Psychological expected utility theory and anticipatory feelings', *The Quarterly Journal of Economics* **116**(1), 55–79.

Charness, G., Oprea, R. & Yuksel, S. (2021), 'How do people choose between biased information sources? evidence from a laboratory experiment', *Journal of the European Economic Association* **19**(3), 1656–1691.

Danz, D., Vesterlund, L. & Wilson, A. J. (2022), 'Belief elicitation and behavioral incentive compatibility', *American Economic Review* **112**(9), 2851–83.

de Clippel, G., Moscariello, P., Ortoleva, P. & Rozen, K. (2024), Caution in the face of complexity, working paper.

Dillenberger, D. (2010), 'Preferences for one-shot resolution of uncertainty and allais-type behavior', *Econometrica* **78**(6), 1973–2004.

Dreyfuss, B. & Hoong, R. (2025), Calibrated coarsening: Designing information for ai-assisted decisions, working paper.

Eliaz, K. & Schotter, A. (2007), 'Experimental testing of intrinsic preferences for noninstrumental information', *American Economic Review* **97**(2), 166–169.

Eliaz, K. & Schotter, A. (2010), 'Paying for confidence: An experimental study of the demand for non-instrumental information', *Games and Economic Behavior* **70**(2), 304–324.

Ely, J., Frankel, A. & Kamenica, E. (2015), 'Suspense and surprise', *Journal of Political Economy* **123**(1), 215–260.

Enke, B. (2024), The cognitive turn in behavioral economics, working paper.

Enke, B. & Graeber, T. (2023), 'Cognitive uncertainty', *Quarterly Journal of Economics* **138**(4), 2021–2067.

Enke, B., Graeber, T. & Oprea, R. (2025), 'Complexity and time', *Journal of the European Economic Association* p. jvaf009.

Enke, B., Graeber, T., Oprea, R. & Yang, J. (2024), Behavioral attenuation, Working Paper 32973, National Bureau of Economic Research.

Enke, B. & Shubatt, C. (2024), Quantifying lottery choice complexity, working paper.

Falk, A. & Zimmermann, F. (2024), 'Attention and dread: Experimental evidence on preferences for information', *Management Science* **70**(10), 7090–7100.

Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Stanford University Press.

Fudenberg, D., Kleinberg, J., Liang, A. & Mullainathan, S. (2022), 'Measuring the completeness of economic models', *Journal of Political Economy* **130**(4), 956–990.

Gill, D. & Prowse, V. (2022), 'Strategic complexity and the value of thinking', *The Economic Journal* **133**(650), 761–786.

Golman, R. & Loewenstein, G. (2018), 'Information gaps: A theory of preferences regarding the presence and absence of information', *Decision* **5**(3), 143–164.

Golman, R., Loewenstein, G., Molnar, A. & Saccardo, S. (2022), 'The demand for, and avoidance of, information', *Management Science* **68**(9), 6454–6476.

Gonçalves, D. (2024), Speed, accuracy and complexity, working paper.

Grant, S., Kajii, A. & Polak, B. (1998), 'Intrinsic preference for information', *Journal of Economic Theory* **83**(2), 233–259.

Greiner, B. (2015), 'Subject pool recruitment procedures: organizing experiments with orsee', *Journal of the Economic Science Association* **1**(1), 114–125.

Grether, D. M. & Plott, C. R. (1979), 'Economic theory of choice and the preference reversal phenomenon', *The American Economic Review* **69**(4), 623–638.

Hossain, T. & Okui, R. (2013), 'The binarized scoring rule', *The Review of Economic Studies* **80**(3 (284)), 984–1001.

Jones, M. & Sugden, R. (2001), 'Positive confirmation bias in the acquisition of information', *Theory and Decision* **50**, 59–99.

Koszegi, B. & Rabin, M. (2009), 'Reference-dependent consumption plans', *American Economic Review* **99**(3), 909–36.

Liang, Y. (2023), Boundedly rational information demand, working paper.

Lichtenstein, S. & Slovic, P. (1971), 'Reversals of preference between bids and choices in gambling decisions.', *Journal of experimental psychology* **89**(1), 46.

Llorente-Saguer, A., Oliveros, S. & Zultan, R. (2025), Beyond value: on the role of symmetry in demand for information, working paper.

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

Masatlioglu, Y., Orhun, Y. & Raymond, C. (2023), 'Intrinsic information preferences and skewness', *American Economic Review* **113**(10), 261544.

Montanari, G. & Nunnari, S. (2022), Audi alteram partem: an experiment on selective exposure to information, working paper.

Niederle, M. & Vespa, E. (2023), 'Cognitive limitations: Failures of contingent thinking', *Annual Review of Economics* **15**(Volume 15, 2023), 307–328.

Nielsen, K. (2020), 'Preferences for the resolution of uncertainty and the timing of information', *Journal of Economic Theory* **189**.

Novk, V., Matveenko, A. & Ravaioli, S. (2024), 'The status quo and belief polarization of inattentive agents: Theory and experiment', *American Economic Journal: Microeconomics* **16**(4), 139.

Oprea, R. (2024*a*), Complexity and its measurement, working paper.

Oprea, R. (2024*b*), 'Decisions under risk are decisions under complexity', *American Economic Review* **114**(12), 37893811.

Oprea, R. & Yuksel, S. (2022), 'Social exchange of motivated beliefs', *Journal of the European Economic Association* **20**(2), 667–699.

Palacios-Huerta, I. (1999), 'The aversion to the sequential resolution of uncertainty', *Journal of Risk and Uncertainty* **18**(3), 249–269.

Puri, I. (2025), 'Simplicity and risk', *The Journal of Finance* **80**(2), 1029–1080.

Shubatt, C. & Yang, J. (2024), Tradeoffs and comparison complexity, working paper.

Zimmermann, F. (2015), 'Clumped or piecewise? evidence on preferences for information', *Management Science* **61**(4), 740–753.

Online Appendix for

Beyond Instrumental Value: How Complexity Shapes Information Demand

Menglong Guan Ryan Oprea Sevgi Yuksel

CONTENTS:

A Characteristics of Information structures and Participant-Level Characteristics

Table 2: Characteristics of Information Structures

Structure characteristic	**Description**
Instrumental value (V)	The expected increase in guessing accuracy made possible by the decision maker being able to condition their action on the realized signals from an information structure.
Number of signals	The number of distinct signals that an information structure can generate.
Number of signals with residual uncertainty (n_{ru})	The number of signals inducing non-degenerate posterior beliefs (i.e., beliefs other than 1 or 0) that an information structure can generate.
Number of distinct posteriors (n_{dp})	The number of distinct posterior beliefs that an information structure can generate.
Unweighted guessing accuracy ($\tilde{V}$)	The incorrect measure of expected increase in guessing accuracy that an information structure can produce when assuming the decision maker considers all possible signals realizing with the same chance.
Probability of maximally certain signals	The probability that an information structure generates maximally certain signals (i.e., signals induce posterior of 1 or 0).
Probability of maximally uncertain signals	The probability that an information structure generates maximally uncertain signals (i.e., signals induce posterior of 0.5).
Probability of worse-than-prior signals	The probability that an information structure generates signals inducing posterior beliefs that are less certain (i.e., closer to uniform belief of 0.5) than prior beliefs.
Probability of guessing black	The probability that an information structure generates signals inducing that the Bayesian optimal guess is 'black' – consistent with the Bayesian optimal guess if conditional only on the prior.
Skewness	The third normalized moment of Bayesian posterior $P(b\|s)$.
Informativeness	The Shannon mutual information between prior beliefs and posterior beliefs induced by an information structure.
Blackwell ordering	The Blackwell ordering of information structures.
Refine	An information structure refines another if the former can produce the latter by arranging and combining signals (groups of the 10 balls) of the former in some way.
Value-dissimilarity ratio	Each information structure induces a compound lottery (as each signal generates a lottery over guessing the state correctly). We focus on the first stage lottery over different likelihoods of guessing the state correctly. The value-dissimilarity ratio of a pair of structures is the ratio of the absolute difference in expected value to the CDF distance of the two corresponding lotteries (similar to the value-dissimilarity ratio as defined in Shubatt & Yang (2024), but defined on the first stage lottery over induced guessing accuracies by signals).
Excess dissimilarity	Each information structure induces a compound lottery (as each signal generates a lottery over guessing the state correctly). We focus on the first stage lottery over different likelihoods of guessing the state correctly. The excess dissimilarity of a pair of structures is the difference between the CDF distance and the absolute difference in expected value of the two corresponding lotteries (similar to the excess dissimilarity measure in Enke & Shubatt (2024), but defined on the first stage lottery over induced guessing accuracies by signals).

Notes: For the first 11 structure characteristic measures, we consider both the difference in the corresponding measure and the sum of the measure for a pair of information structures in our LASSO, Decision Tree, and CNN analysis.

Table 3: Characteristics of Information Structures: Summary Statistics

	Structure characteristic	Min	Max	Mean	Median	Variance
Difference	Instrumental value	0.1	0.4	0.1916	0.2	0.0091
	Number of signals	-3	4	0.4694	0	1.7164
	Number of signals with residual uncertainty	-4	2	-1.1315	-1	0.9904
	Number of distinct posterior	-2	3	-0.1234	0	1.0029
	Unweighted guessing accuracy	-0.1778	0.4444	0.161	0.1458	0.0117
	Probability of maximally certain signals	-0.4	1	0.4222	0.4	0.0811
	Probability of maximally uncertain signals	-0.8	0.6	-0.2228	-0.2	0.0828
	Probability of worse-than-prior signals	-0.9	0.7	-0.2661	-0.2	0.092
	Probability of guessing black	-0.7	0.5	-0.0631	-0.1	0.0338
	Skewness	-5.3334	3.0749	-0.5365	-0.4017	1.3655
	Informativeness	-0.1146	0.971	0.4197	0.39	0.0564
Sum	Instrumental value	0.1	0.7	0.3995	0.4	0.029
	Number of signals	3	10	7.727	8	1.8095
	Number of signals with residual uncertainty	1	7	2.8749	3	1.6666
	Number of distinct posterior	3	8	5.3328	5	0.8231
	Unweighted guessing accuracy	0.0111	0.7667	0.4877	0.5	0.0256
	Probability of maximally certain signals	0	1.8	0.9152	0.9	0.1739
	Probability of maximally uncertain signals	0	1.4	0.3276	0.2	0.1052
	Probability of worse-than-prior signals	0	1.6	0.3835	0.4	0.1243
	Probability of guessing black	0.7	1.9	1.2917	1.3	0.0355
	Skewness	-4.1667	3.5396	-0.2905	-0.4547	1.382
	Informativeness	0.0913	1.742	0.9515	0.9768	0.1506
	Blackwell ordering	0	1	0.7766	1	0.1735
	Refine	0	1	0.106	0	0.0948
	Value-dissimilarity ratio	0.4167	1	0.9549	1	0.0133
	Excess dissimilarity	0	0.14	0.0075	0	0.0004

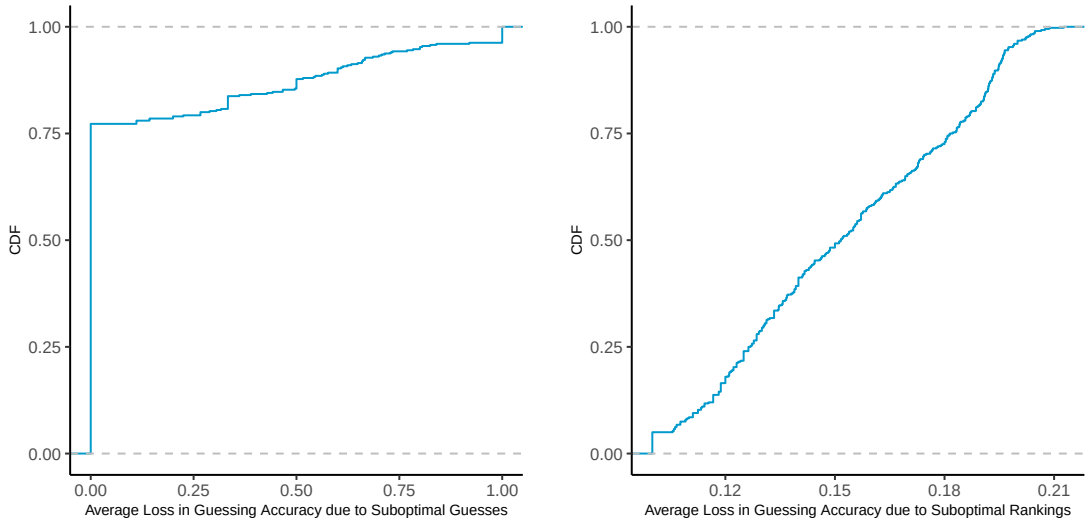
Notes: We focus on the pairwise comparisons of information structures that have a strict difference in instrumental value in our data.

Table 4: Participant Characteristics

Participant characteristic	**Description**
Raven	Scores on three Raven's test questions.
Wason	Score on the Wason question.
CRT	Scores on three cognitive reflection test questions.
BBS	Scores on three belief bias syllogism questions.
Cognitive	Overall scores on the ten cognitive questions.
Risk	Certainty equivalent to a simple lottery that pays $30 with a 50% chance.
Ambiguity	Certainty equivalent to an Ellsberg lottery that has the same expected value as the simple lottery.
Compound	Certainty equivalent to a compound lottery that has the same expected value as the simple lottery.
Confidence in Ranking	Confidence in the optimality of rankings, measuring how certain participants are in the optimality of their rankings of information structures.
Computation of information instrumental value	Computation of the chance that a computer, which always makes optimal guesses, will correctly guess the drawn ball color when receiving information from a given information structure. This is essentially the computation of the instrumental value of the given information structure.
Confidence in computation of information instrumental value	Confidence in correctly computing the instrumental value of the given information structure.
Complex Information computation	Whether a participant faces the relatively more complex information structure in the instrumental value computation question.
Computation of compound lottery value	Computation of the chance of winning the prize of the compound lottery.
Confidence in computation of compound lottery	Confidence in correctly computing the winning chance of the compound lottery.
Mistake in computation of information instrumental value	Mistake in the computation of the instrumental value of the information structure.
Mistake in computation of compound lottery value	Mistake in the computation of the winning chance of the compound lottery.
Absolute mistake in computation of information instrumental value	Absolute mistake in the computation of the instrumental value of the information structure.
Absolute mistake in computation of compound lottery value	Absolute mistake in the computation of the winning chance of the compound lottery.
Ambiguity aversion	How much lower the certainty equivalent to the Ellsberg lottery is compared to the certainty equivalent to the simple lottery. This provides a measure of ambiguity attitude.
Aversion to compound lottery	How much lower the certainty equivalent to the compound lottery is compared to the certainty equivalent to the simple lottery. This provides a measure of attitude to compound risk.

Notes: The table presents the full list of the 20 participant characteristics we include in LASSO and Decision Tree analyses.

B Additional Plots and Tables



(a) Losses due to Suboptimal Guesses (b) Losses due to Suboptimal Ranking

Figure 9: Distribution of Losses due to Suboptimal Guesses and Ranking *Notes: In panel (a) losses in likelihood of winning the bonus due to suboptimal guesses in guessing tasks are computed on the individual level. In panel (b) losses in likelihood of winning the bonus due to suboptimal rankings in ranking tasks are computed on the individual level, assuming optimal guesses.*

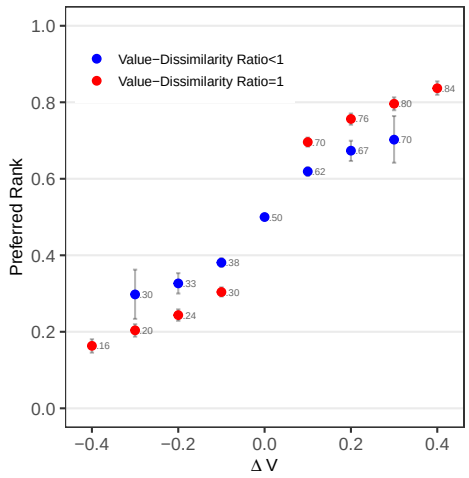


Figure 10: Pairwise Ranking of Information Structures by Value Difference: The Impact of Value-Dissimilarity Ratio *Notes: The figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where the value-dissimilarity ratio (defined in Table 2, inspired by Shubatt & Yang (2024)) between the two structures is either equal to 1 or not. By definition, the value-dissimilarity ratio takes value from [0,1], with a larger value indicating lower complexity of comparing two structures. In our data, value-dissimilarity ratio is equal to 1 in 76% of all pairwise comparisons. Additionally, the excess dissimilarity measure (inspired by Enke & Shubatt (2024)) is by definition closely related to value-dissimilarity ratio. Thus, we do not replicate the same graph for that measure. Gray lines denote 95 percent confidence intervals by bootstrapping.*

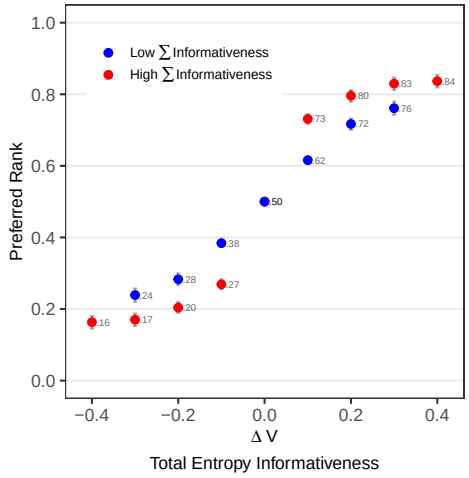


Figure 11: Pairwise Ranking of Information Structures by Value Difference: The Impact of Total Entropy Informativeness *Notes: The figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where the total entropy informativeness of the two structures is either strictly above the median or weakly below. Entropy informativeness of an information structure is defined as the expected reduction of entropy (i.e., a measure of uncertainty in beliefs) from prior to posterior beliefs that are induced by the structure. Gray lines denote 95 percent confidence intervals by bootstrapping.*

Table 5: Impact of Structure Characteristic on Completeness: All and Each Individually

Structure characteristic	Model Completeness
All 26 characteristics	0.85
Δ Number of signals with residual uncertainty (Δn_{ru})	0.56
Δ Informativeness	0.33
Δ Unweighted guessing accuracy ($\Delta \tilde{V}$)	0.32
$\sum$ Informativeness	0.30
Δ Instrumental value	0.29
Δ Probability of maximally certain signals	0.29
$\sum$ Probability of maximally certain signals	0.28
Δ Number of distinct posterior (Δn_{dp})	0.27
$\sum$ Instrumental value ($\sum V$)	0.27
Value-dissimilarity ratio	0.16
$\sum$ Number of signals with residual uncertainty	0.15
Excess dissimilarity	0.15
$\sum$ Unweighted guessing accuracy	0.13
$\sum$ Number of distinct posterior	0.11
$\sum$ Probability of worse-than-prior signals	0.08
Blackwell ordering	0.08
$\sum$ Probability of maximally uncertain signals	0.04
Refine	0.03
Δ Probability of maximally uncertain signals	0.03
Δ Number of signals	0.03
$\sum$ Skewness	0.01
$\sum$ Probability of guessing black	0.00
$\sum$ Number of signals	0.00
Δ Skewness	0.00
Δ Probability of worse-than-prior signals	0.00
Δ Probability of guessing black	0.00

Notes: See Table 1 for how model completeness is defined and computed. Δ denotes the difference in the corresponding measure, and $\sum$ denotes the sum of the measure for a pair of information structures.

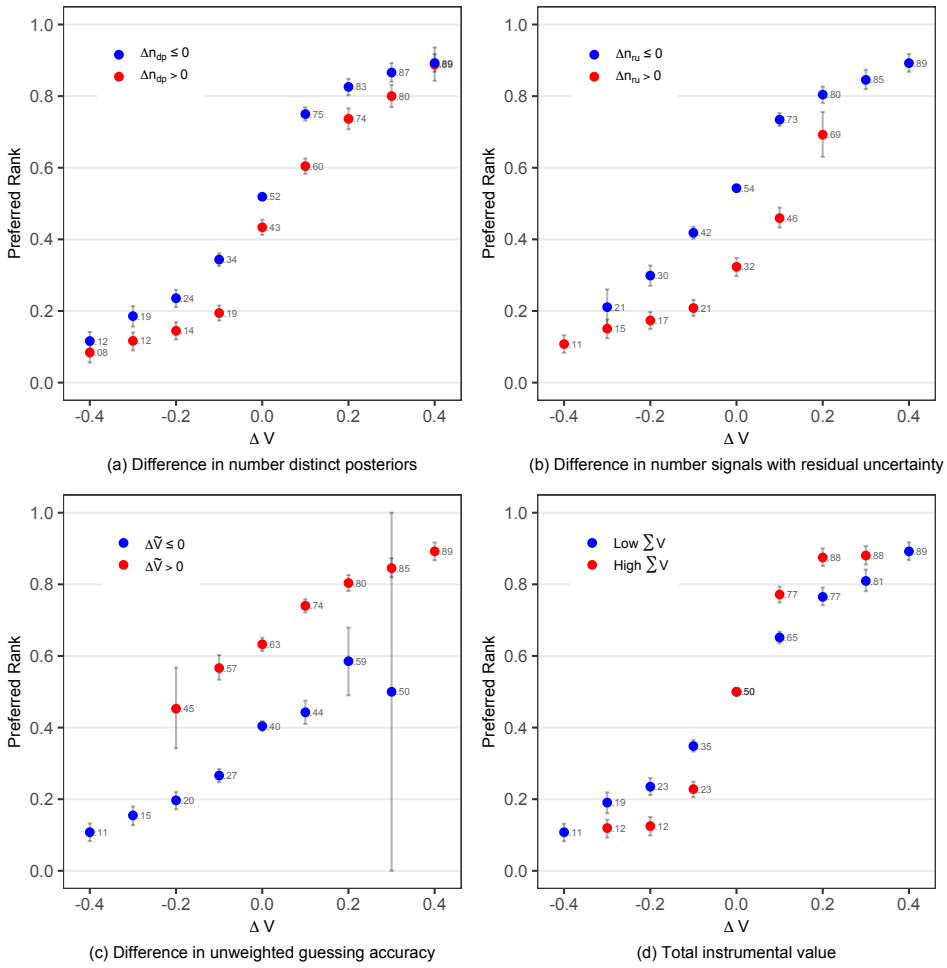


Figure 12: Pairwise Ranking of Information Structures by Value Difference: All Comprehension Questions Answered Correctly at First Try *Notes: This figure replicates Figure 6 but focuses on only participants who answered all comprehension questions correctly at their first try (153 participants). Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

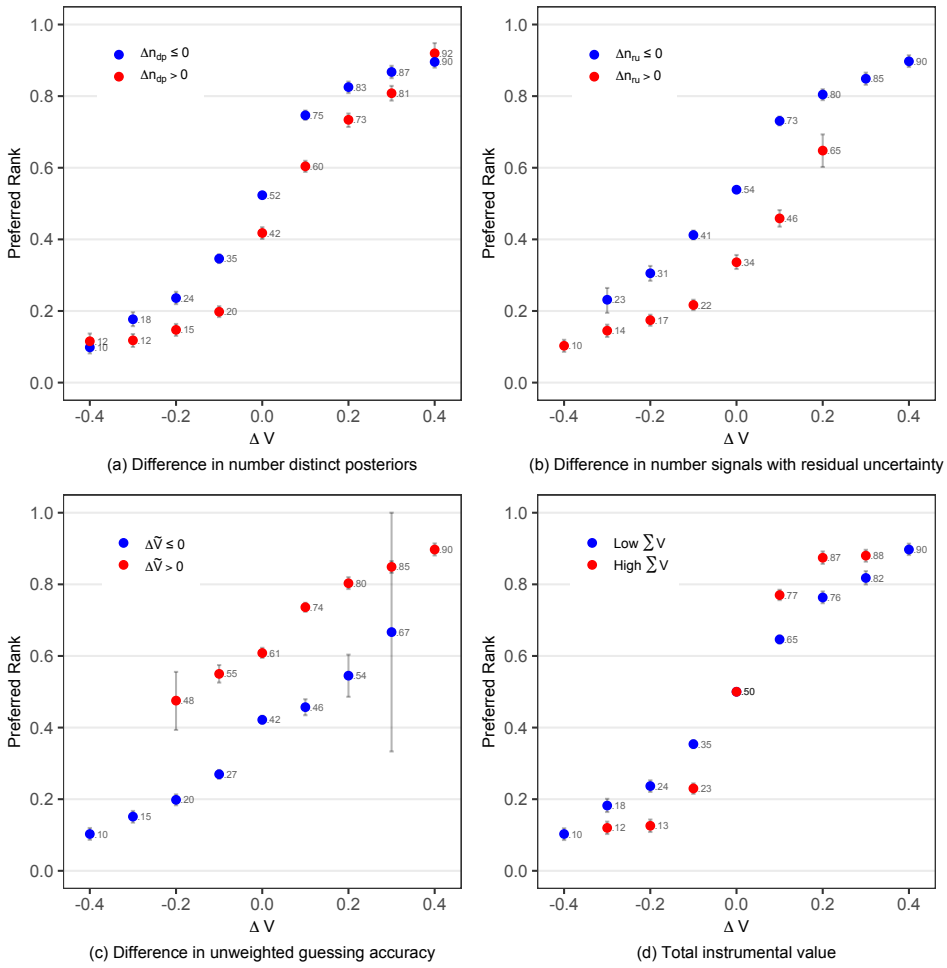


Figure 13: Pairwise Ranking of Information Structures by Value Difference: Guesses Are Always Optimal

Notes: This figure replicates Figure 6 but focuses on only participants who make optimal guesses all the time (309 participants). Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.

Table 6: Impacts of Structure Characteristics on Ranking Preference

	Ranking Preference (Logit)				
	(1)	(2)	(3)	(4)	(5)
Difference in number distinct posteriors (Δn_{dp})	-0.285***				-0.144***
	(0.019)				(0.021)
Difference in number signals with residual uncertainty (Δn_{ru})		-0.420***			-0.082**
		(0.027)			(0.034)
Difference in unweighted guessing accuracy ($\Delta \tilde{V}$)			2.973***		2.917***
			(0.216)		(0.319)
Total instrumental value ($\sum V$)				1.679***	1.297***
				(0.114)	(0.123)
Constant	0.990***	0.571***	0.552***	0.354***	-0.040
	(0.045)	(0.034)	(0.030)	(0.042)	(0.049)
No. of subjects	400	400	400	400	400

*Notes: Analysis is on pairwise comparisons where one structure has a strictly higher instrumental value than the other ($\Delta V > 0$). The dependent variable is whether the structure with a strictly higher instrumental value is ranked as more preferred than the other. Therefore, the constant term captures the average impact of the difference in instrumental value. Standard errors (clustered at the subject level) in parentheses. *** 1%, ** 5%, * 10% significance.*

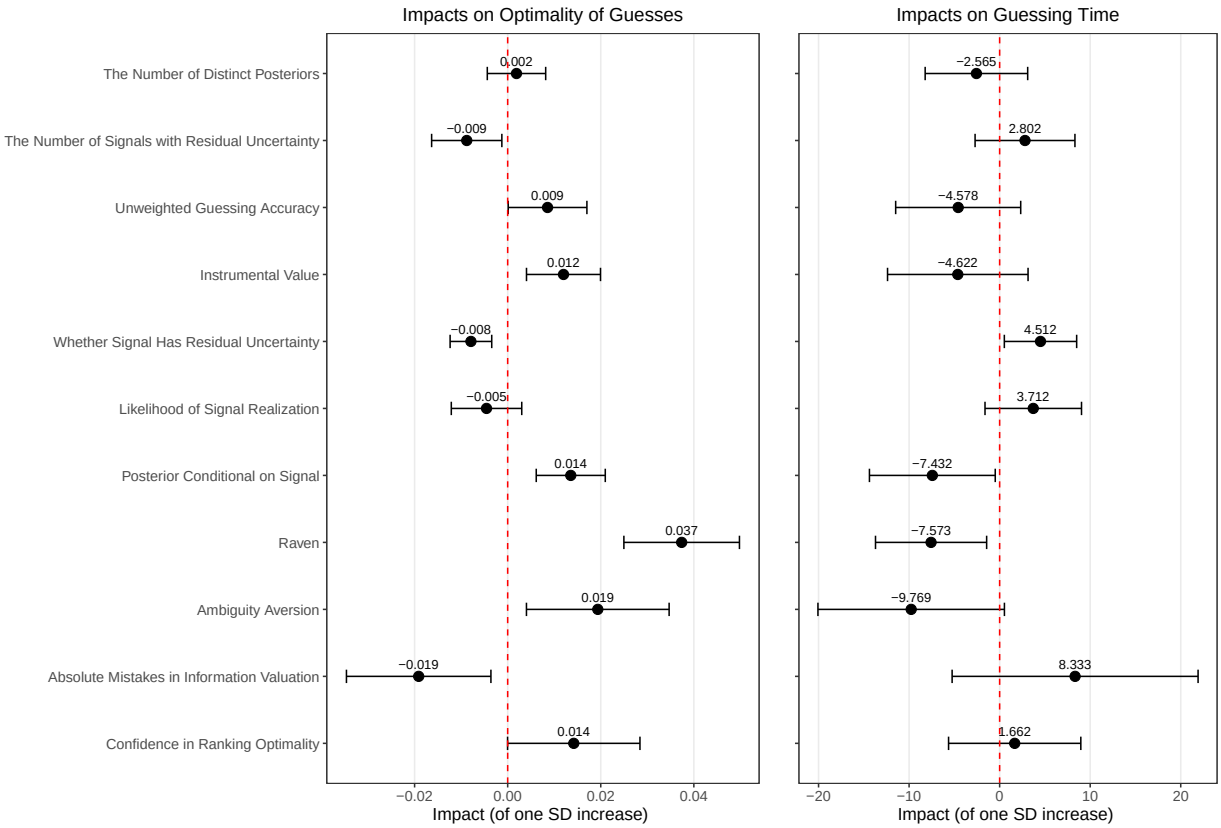


Figure 14: Impact of Structure Characteristics, Signal Characteristics, and Participant Characteristics on Guessing Optimality and Time *Notes: The figure plots the absolute impact of one standard deviation increase in a structure characteristic, signal characteristic, or participant characteristic on guessing optimality and time (in 1/10 seconds). We focus on all pairwise comparisons where one structure is of strictly higher value than the other. Short lines denote 95 percent confidence intervals.*

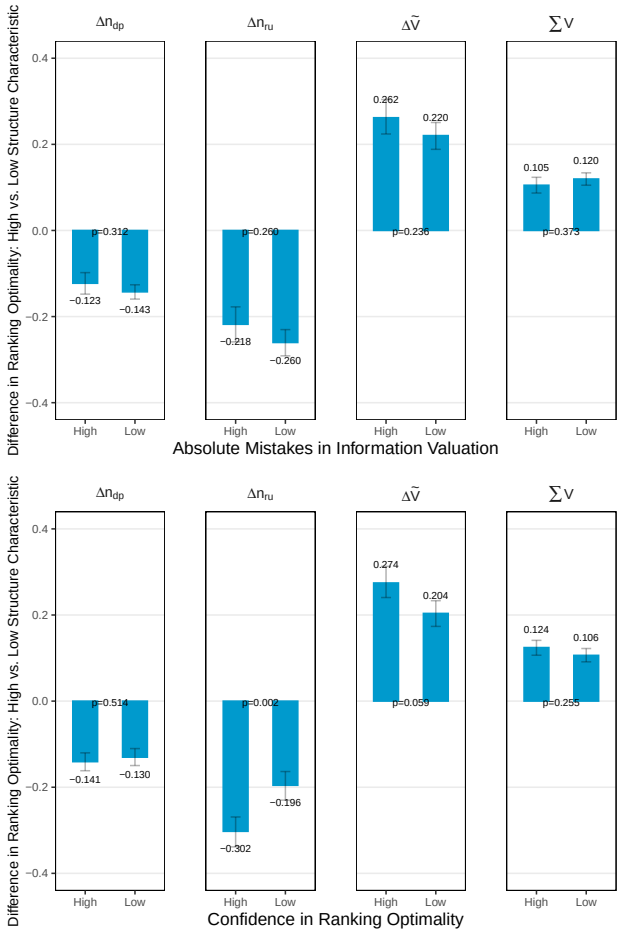


Figure 15: Interaction Between Participant Characteristics and Structure Characteristics *Notes: Analysis is on pairwise rankings where* $\Delta V = 0.1$. *Bars represent the difference in ranking optimality when characteristic* f *is strictly above the median value in our sample versus weakly below, separated by participant type. Gray lines denote 95 percent confidence intervals by bootstrapping. P values are calculated using two-sample t-tests.*

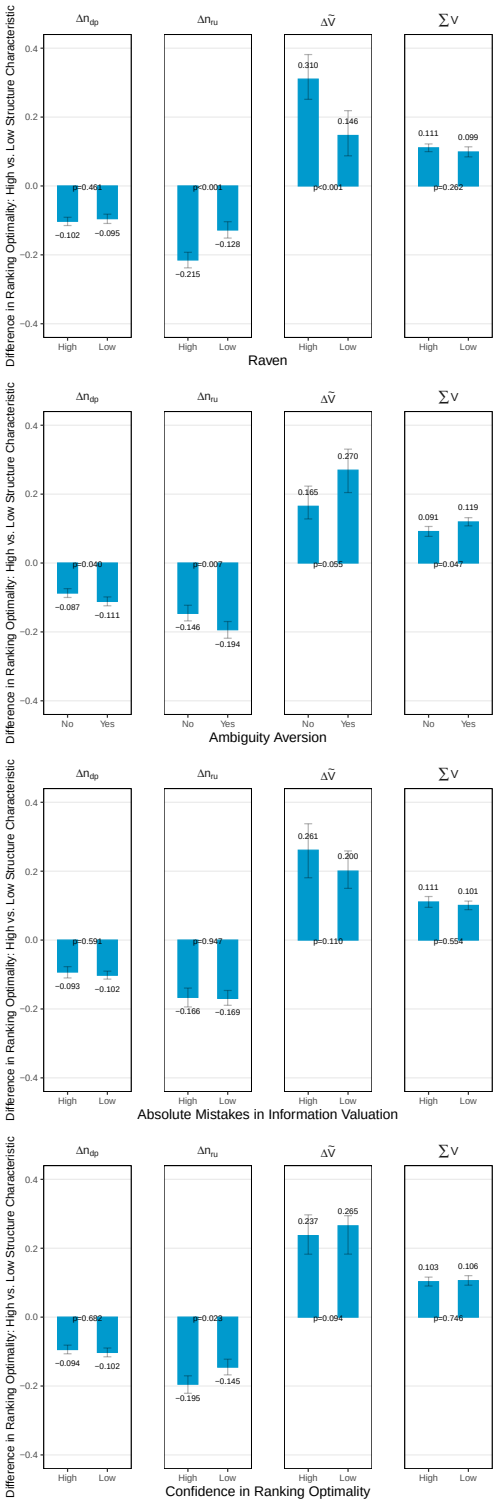


Figure 16: Interaction Between Participant Characteristics and Structure Characteristics (Overall) *Notes: Analysis is on all pairwise rankings for Δn_{dp}, Δn_{ru} and $\Delta \tilde{V}$, and on pairwise rankings where $\Delta V > 0$ for $\sum V$. Bars represent the difference in ranking optimality when characteristic f is strictly above the median value in our sample versus weakly below, separated by participant type. Gray lines denote 95 percent confidence intervals by bootstrapping. P values are calculated using two-sample t-tests.*

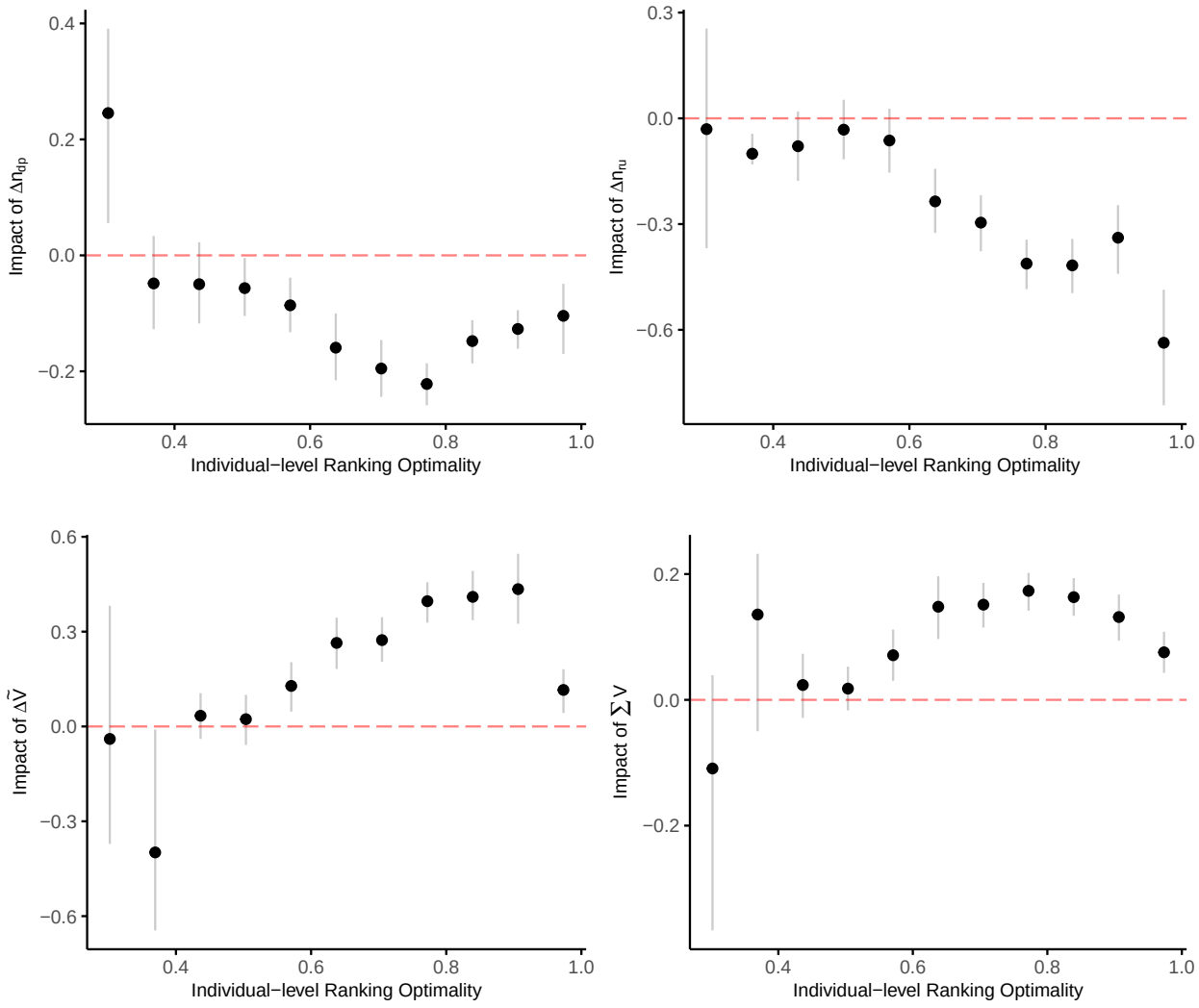


Figure 17: Impact of Selected Characteristics as a Function of Ranking Optimality Rate *Notes: Bin-scatter plots of the impacts of four identified structure characteristics across participants. Analysis is on pairwise rankings where* $\Delta V = 0.1$. *Values (on the y-axis) represent the difference in ranking optimality when characteristic* f *is strictly above the median value in our sample versus weakly below, separated by participants' overall optimality rate in this sample (on the x-axis). Gray lines denote 95 percent confidence intervals by bootstrapping.*

Table 7: Interaction Effects between Structure Characteristics and Ranking Task Complexity on Ranking Preference

	Ranking Preference (Logit)			
	(1)	(2)	(3)	(4)
Difference in number of distinct posteriors (Δn_{dp})	-0.125***			
	(0.021)			
$\Delta n_{dp}\times$ (Number of fully revealing structures - 1)	-0.080**			
	(0.038)			
Difference in number of signals with residual uncertainty (Δn_{ru})		-0.357***		
		(0.029)		
$\Delta n_{ru}\times$ (Number of fully revealing structures - 1)		-0.071		
		(0.045)		
Difference in unweighted guessing accuracy ($\Delta \tilde{V}$)			3.196***	
			(0.251)	
$\Delta \tilde{V}\times$ (Number of fully revealing structures - 1)			0.557*	
			(0.306)	
Total instrumental value ($\sum V$)				0.501***
				(0.128)
$\sum V\times$ (Number of fully revealing structures - 1)				0.334**
				(0.138)
Constants	0.788***	0.481***	0.292***	0.578***
	(0.039)	(0.034)	(0.030)	(0.050)
No. of subjects	400	400	400	400

Notes: The seven information structures in each ranking task are randomly selected without replacement while ensuring each task contains information structures of all possible instrumental values $V \in \{0, 0.1, 0.2, 0.3, 0.4\}$ representing the increase in guessing accuracy. Thus, each ranking task contains at least one information structure that fully reveals the state (i.e., $V = 0.4$). Comparisons between a structure that fully reveals the state and another structure should be straightforward, as it is intuitively easy to identify that the former is more valuable. Therefore, ranking tasks containing more structures that fully reveal the state are arguably less complex, given that there are more pairwise comparisons with a structure that fully reveals the state. We use the number of structures that fully reveal the state (out of seven) as a measure of the simplicity *of a ranking task to study whether this measure influences the main results (i.e., the impacts of identified structure characteristics on ranking preference) reported in the paper. Each regression focuses on pairwise comparisons where one structure has a strictly higher instrumental value than the other ($\Delta V > 0$) and the more valuable structure does not fully reveal the state. The dependent variable is whether the structure with a strictly higher instrumental value is ranked as more preferred than the other. The constant term captures the average impact of the difference in instrumental value (ΔV). Standard errors (clustered at the subject level) in parentheses. *** 1%, ** 5%, * 10% significance.*

Table 8: Impacts of Structure Characteristics on Ranking (Rank-ordered Logit)

	Ranking (1=most preferred, 7=least preferred)					
	(1)	(2)	(3)	(4)	(5)	(6)
Instrumental value (V)	-5.167***	-5.156***	-2.992***	-3.768***	-5.167***	-3.029***
	(0.255)	(0.267)	(0.231)	(0.240)	(0.255)	(0.248)
Number of distinct posteriors (n_{dp})		0.352***				0.302***
		(0.018)				(0.019)
Number of signals with residual uncertainty (n_{ru})			0.401***			0.196***
			(0.029)			(0.034)
Unweighted guessing accuracy ($\tilde{V}$)				-1.703***		-1.306***
				(0.273)		(0.327)
$(\sum_{i=1}^{7} V$ - median$(\sum_{i=1}^{7} V)) \times V$					-0.775	-0.068
					(0.678)	(0.713)
No. of Subjects	400	400	400	400	400	400
No. of Ranking Tasks	3200	3200	3200	3200	3200	3200

*Notes: The unit of analysis is a ranking task. The dependent variable is the ranking of an information structure within the set of seven structures from the same ranking task (1=most preferred, 7=least preferred). A negative (positive) coefficient indicates that structures with higher values of the corresponding characteristic tend to be ranked as more (less) preferred. As discussed in Section 4.3, our primary pairwise comparison analysis examines four identified characteristics: differences in n_{dp}, n_{ru}, $\tilde{V}$, and the sum of V for the two structures being compared. In rank-ordered logit regressions, corresponding to the first three, we directly examine n_{dp}, n_{ru}, and $\tilde{V}$ of each structure. For the fourth characteristic, we use the sum of V of all seven structures in the same ranking task (centered at the median of this measure) and explore whether a lower (higher) $\sum_{i=1}^{7} V$ attenuates (enhances) the impact of instrumental value. Standard errors (clustered at the subject level) in parentheses. *** 1%, ** 5%, * 10% significance.*

C LASSO and Decision Tree

When identifying which information structure characteristics (participant characteristics) are important predictors of information demand using LASSO and Decision Tree models, we focus on the dataset of pairwise comparisons where, in each pair (denoted as structures A and B), structure A has a strictly higher instrumental value than B.

In LASSO analysis, we fit logistic LASSO regressions with the dependent variable being whether structure A was ranked as more preferred than B, and the independent variables being the 26 structure characteristics (20 participant characteristics). We use a grid of regularization parameter (λ) values, specifically $\log\lambda \in \{-15, -14.9, -14.8, ..., -2.2, -2.1, -2\}$. A larger λ indicates a higher strength of the regularization penalty and will result in a simpler model, i.e., more structure characteristics (participant characteristics) have zero coefficients. In decision tree analysis, we fit classification decision trees using the Recursive Partitioning and Regression Trees (rpart) algorithm for a grid of tuning parameters $(minsplit, minbucket, cp) : \{5, 10, ..., 45, 50\} \times \{3, 5, ..., 17, 19\} \times \{0.01, 0.0075, 0.005, 0.0025, 0.001, 0.00075, 0.0005, 0.00025, 0.0001, 0.00005\}$, where *minsplit* controls the minimum node size for splitting, *minbucket* defines the minimum number of observations that are allowed in a terminal node, and *cp* controls the complexity of a tree. In our analysis, *cp* essentially determines structure characteristic (participant characteristic) selection in a decision tree.

Table 9a presents the structure characteristics with non-zero coefficients in the LASSO regressions at different values of λ, while Table 9b presents the structure characteristics included in decision trees at different values of *cp*. Notably, Δn_{ru}, $\Delta\tilde{V}$, and Δn_{dp} are selected in even very simple models under both LASSO and decision tree analyses. Although $\sum V$ is selected in slightly complex models, $\sum Informativeness$, which is highly correlated with $\sum V$, is selected early under both analyses.[52] We focus on $\sum V$ rather than $\sum Informativeness$ as the former has a more intuitive interpretation.

Table 10a presents the participant characteristics with non-zero coefficients in the LASSO regressions, and Table 10b presents the participant characteristics included in decision trees. *Raven, Confidence in ranking*, and *Absolute mistake in computation of information structure value* are selected in simple LASSO models. These three participant characteristics are also selected and appear at the top in decision trees. We additionally include *Ambiguity aversion* as it has been shown

[52] In decision tree models, Δn_{ru}, $\Delta\tilde{V}$, Δn_{dp} and $\sum Informativeness$ appear on the top of the trees (as early splits), meaning they present the strongest predictors.

to have prediction power on decisions under complexity in recent literature. We also note that when excluding confidence measures, *Ambiguity aversion* is selected in relatively simple LASSO models and appears at the top in decision trees (see Table 11).

Table 9: Summary of Structure Characteristic Selection: LASSO and Decision Tree

(a) LASSO

$\log \lambda$	No. of Characteristics	Characteristics with Non-zero Coefficient
-2 to -2.5	0	
-2.6 to -3	1	$\boldsymbol{\Delta n_{ru}}$
-3.1 to -3.4	2	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness
-3.5 to -3.7	5	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$
-3.8	6	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio
-3.9 to -4	7	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$
-4.1 to -4.6	8	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$, $\sum n_{dp}$
-4.7	9	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$, $\sum n_{dp}$, $\sum$ Number of signals
-4.8 to -4.9	10	$\boldsymbol{\Delta n_{ru}}$, $\sum$ Informativeness, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$, $\sum n_{dp}$, $\sum$ Number of signals, Δ Number of signals
...	...	...
-10.5 to -15	26	All 26 characteristics.

(b) Decision Tree

cp	No. of Characteristics	Characteristics
{0.01, 0.0075}	0	
{0.005, 0.0025}	3	$\boldsymbol{\Delta n_{ru}}$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$
{0.001, 0.00075}	6	$\boldsymbol{\Delta n_{ru}}$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum$ Informativeness, $\sum$ Number of signals, Δ Probability of maximally certain signals
0.0005	10	$\boldsymbol{\Delta n_{ru}}$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum$ Informativeness, $\sum$ Number of signals, Δ Probability of maximally certain signals, Δ Skewness, Value-dissimilarity ratio, $\sum n_{ru}$, Δ Probability of maximally uncertain signals
...	...	...
0.00005	26	All 26 characteristics.

Notes: Selected characteristics reported in the paper are highlighted in bold.

Table 10: Summary of Participant Characteristic Selection: LASSO and Decision Tree

(a) LASSO

$\log \lambda$	No. of Characteristics	Characteristics with Non-zero Coefficient
-2 to -2.4	0	
-2.5 to -2.6	1	**Raven**
-2.7 to -2.9	2	**Raven**, Cognitive
-3 to -3.7	3	**Raven**, Cognitive, **Confidence in ranking**
-3.8	4	**Raven**, Cognitive, **Confidence in ranking**, **Absolute mistake in computation of information structure value**
-3.9	5	**Raven**, Cognitive, **Confidence in ranking**, **Absolute mistake in computation of information structure value**, Confidence in computation of information structure value
-4	9	**Raven**, Cognitive, **Confidence in ranking**, **Absolute mistake in computation of information structure value**, Confidence in computation of information structure value, Confidence in computation of compound lottery value, Absolute mistake in computation of compound lottery value, Computation of information structure value, Mistake in computation of information structure value
...	...	...
-5.4 to -7.1	16	**Raven**, Cognitive, **Confidence in ranking**, **Absolute mistake in computation of information structure value**, Confidence in computation of information structure value, Confidence in computation of compound lottery value, Absolute mistake in computation of compound lottery value, Computation of information structure value, Mistake in computation of information structure value, BBS, Aversion to compound lottery, Computation of compound lottery, Mistake in Computation of compound lottery, Compound, Complex information computation, **Ambiguity aversion**
...	...	...
-10.1 to -15	19	All except Risk

(b) Decision Tree

cp	No. of Characteristics	Characteristics
{0.01, 0.0075, 0.005, 0.0025}	0	
{0.001, 0.00075, 0.0005}	13	**Raven**, **Confidence in ranking**, Cognitive, **Absolute mistake in computation of information structure value**, Confidence in computation of compound lottery value, Confidence in computation of information structure value, Compound, Computation of information structure value, Risk, BBS, Absolute mistake in computation of compound lottery value, Aversion to compound lottery, Ambiguity
...	...	...

Notes: Selected characteristics reported in the paper are highlighted in bold.

Table 11: Summary of Participant Characteristic Selection (When Excluding Confidence Measures): LASSO and Decision Tree

(a) LASSO

$\log \lambda$	No. of Characteristics	Characteristics with Non-zero Coefficient
-2 to -2.4	0	
-2.5 to -2.6	1	**Raven**
-2.7 to -3.7	2	**Raven**, Cognitive
-3.8	3	**Raven**, Cognitive, **Absolute mistake in computation of information structure value**
-3.9	4	**Raven**, Cognitive, **Absolute mistake in computation of information structure value**, Computation of information structure value
-4	5	**Raven**, Cognitive, **Absolute mistake in computation of information structure value**, Computation of information structure value, BBS
-4.1 to -4.7	6	**Raven**, Cognitive, **Absolute mistake in computation of information structure value**, Computation of information structure value, BBS, Absolute mistake in computation of compound lottery value
-4.8	9	**Raven**, Cognitive, **Absolute mistake in computation of information structure value**, Computation of information structure value, BBS, Absolute mistake in computation of compound lottery value, **Ambiguity aversion**, Aversion to compound lottery, Wason
...	...	...

(b) Decision Tree

cp	No. of Characteristics	Characteristics
{0.01, 0.0075, 0.005, 0.0025}	0	
0.001	12	**Raven**, **Absolute mistake in computation of information structure value**, **Ambiguity aversion**, Aversion to compound lottery, Computation of information structure value, Computation of compound lottery value, Compound, Complex information computation, Wason, Risk, Ambiguity, Cognitive
...	...	...

Notes: Selected characteristics reported in the paper are highlighted in bold.

D Laboratory Study and Replication Study

D.1 Details of Laboratory Study

We study a total of 16 information structures, depicted in Figure 18, which were selected to independently vary instrumental value and informativeness (as well as other characteristics). For each of these information structures we (i) study how people make use of these information structures (by eliciting guesses for each information structure and signal) and (ii) study how people value these information structures (by eliciting ranking over the 16 structures and willingness-to-pay for these structures).[53,54]

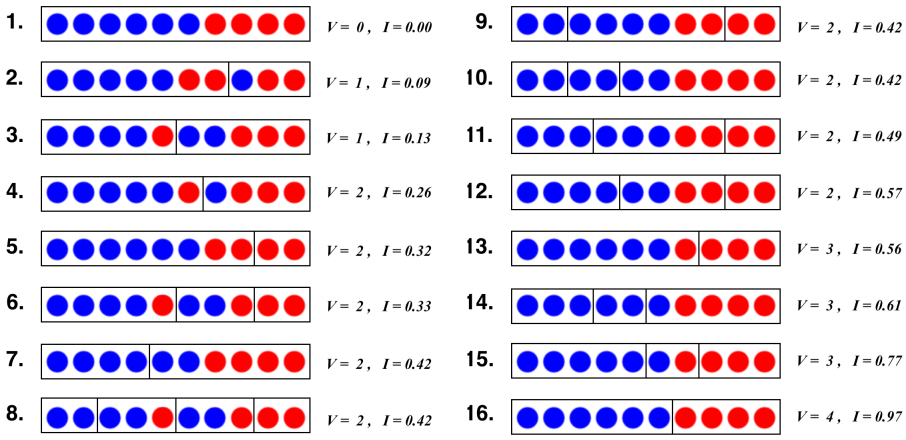


Figure 18: The 16 Information Structures used in the Laboratory Study *Notes: These 16 information structures are a subset of the 237 structures in the Main study. The information structures are listed lexicographically in order of instrumental value (V) and entropy informativeness (I). To make the partitions more salient in the experiment, each cell of the partition was labeled as "Group 1", "Group 2," etc.*

D.2 Treatment Variations in Laboratory Study

In our **Baseline** treatment ($N = 108$ subjects), we ask subjects to perform the Demand section of the experiment (the Ranking and WTP elicitations) first, and the Guesses section afterwards. To this we add two diagnostic treatments that will help us to interpret our results.

First, in our **Reverse** treatment ($N = 53$ subjects), we reversed the order: subjects were assigned Guesses first and Demand afterwards. In these sessions, during the Demand section, subjects were actually shown the guesses they had made earlier, conditional on each possible signal

[53] See Figure 26 in Appendix E for a screenshot of Guessing tasks in the laboratory study.

[54] See Figure 27 in Appendix E for screenshots of Ranking and WTP elicitation tasks in the laboratory study.

for that information structure.[55] This information was not binding, but was designed to remind subjects of how different signal realizations are likely to generate different guessing patterns. The purpose of this treatment was to study whether having already made use of information structures (and being reminded of how they are used) improves subjects' identification of instrumental value in guiding their demand. This is useful for understanding the mechanism behind our results.

Second, in our **No Uncertainty** treatment ($N = 61$ subjects), we removed uncertainty from the design using techniques employed by Martinez-Marquina, Niederle & Vespa (2019) and Oprea (2024*b*). Specifically, instead of drawing only one ball, we informed subjects that we would draw *all balls* and pay subjects based on the accuracy of their guess for each of the balls.[56,57] Everything else in the experiment remains identical. Doing this retains much of the information processing involved in evaluating and interpreting information structures, but removes scope for risk, uncertainty or timing preferences (e.g., preferences for the timing of the resolution of uncertainty). Again, this data is useful for understanding the mechanism behind our results.

Implementation Details

All sessions were conducted at the LITE laboratory at the University of California, Santa Barbara between December 2021 and March 2023. We recruited subjects from across the curriculum to participate in 15 sessions using the ORSEE recruiting software (Greiner 2015). The experiment was conducted using software programmed by the authors in Qualtrics. Between 8 and 21 subjects participated in each session and sessions lasted for 30-40 minutes.

All subjects received a show-up fee of $7. Subjects' earnings depended on a randomly selected section of the experiment. If the Demand section was selected for payment, we randomized between Ranking and WTP. If Ranking was selected, two information structures (from the set of 16) were randomly selected and the subject was assigned to receive information from the one that was ranked higher. If WTP was selected, one information structure (from the set of 16) was randomly selected and whether or not the subject received information from this source was determined according to the BDM mechanism. In either case (regardless of whether Ranking or WTP was selected), at the end of the experiment, the subject was presented with one additional guessing task with

[55] See Figure 28 in Appendix E for a screenshot of how this was displayed to subjects.

[56] In this case, the partition associated with an "information" structure constrains the types of guesses subjects can make by requiring them to make the same guess for all balls in the same cell of the partition.

[57] To achieve a clean comparison to the other treatments, subjects received a prize of $1 for each of the balls for which their guesses were correct. This implies, for example, an information structure that enables a guessing accuracy of 90 percent will generate the same expected bonus payment in all three treatments.

information from the selected information structure and subjects were paid based on this guess.[58] If the Guesses section was selected for payment, a random information structure was picked and the subject's guesses for that case were used to determine payment for a randomly drawn ball. Note that in all cases, whether or not subjects received a bonus payment of $10 ultimately depended on the accuracy of their guess about the color of a randomly selected ball.

D.3 Replication Study

We also ran a **Replication** study on Prolific ($N = 50$ subjects) which used the same 16 information structures from the laboratory study but used the interface of the main study. Our aim was to test whether results would be robust across different subject pools and experimental interfaces. The study was conducted in September 2024. The median time participants spent was about 18 minutes.

D.4 Results

[58] Note that this could mean the subject receives no information depending on their WTP if WTP is selected.

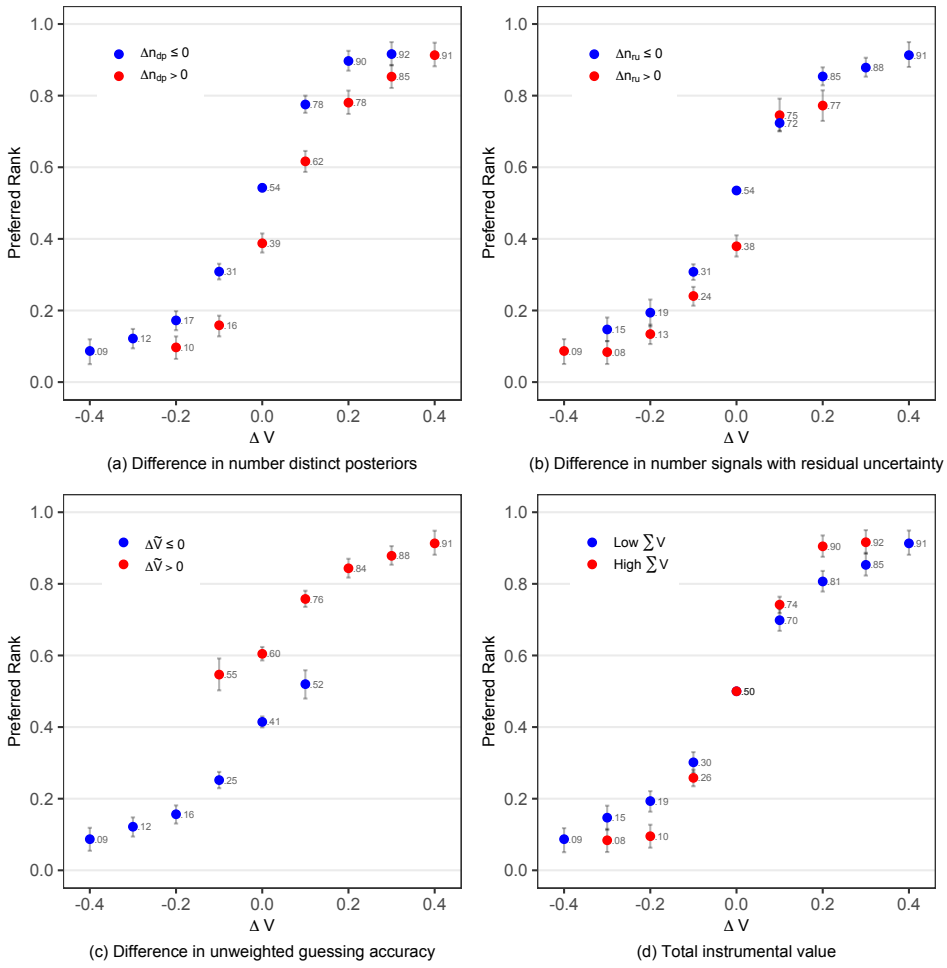


Figure 19: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study *Notes: Both Baseline (Guess First) and Reverse (Guess Second). We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

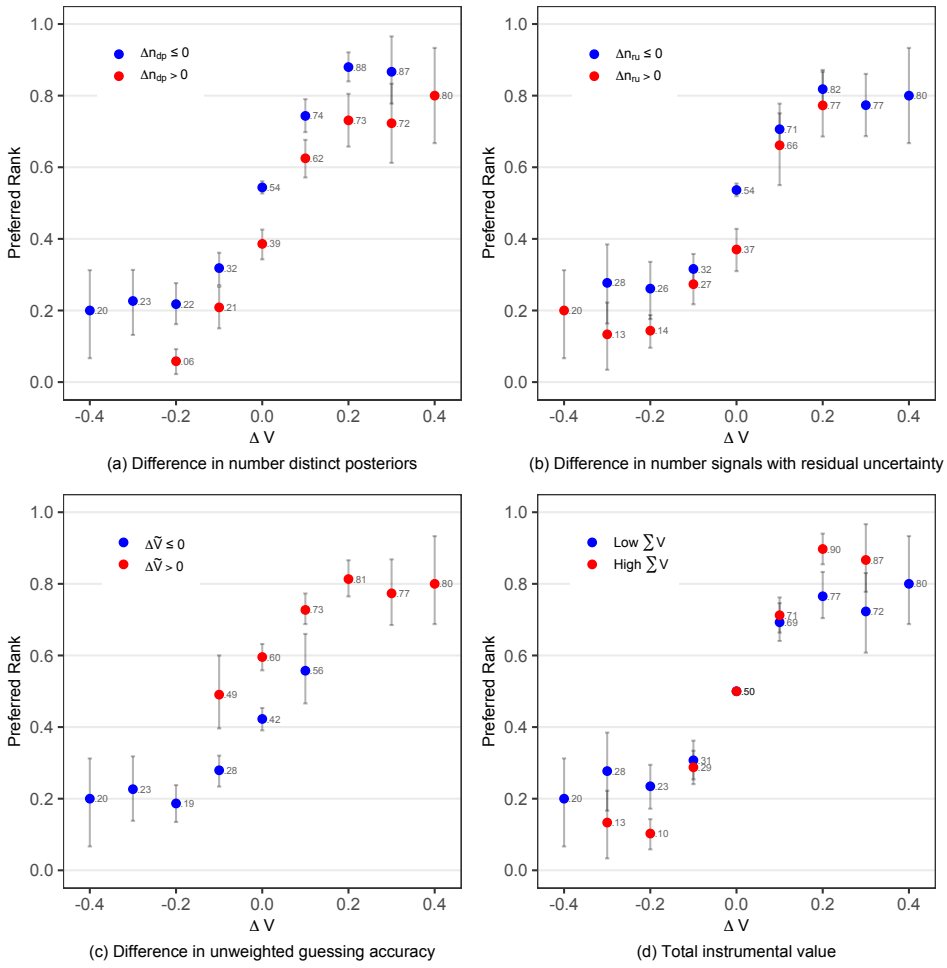


Figure 20: Pairwise Ranking of Information Structures by Value Difference in the Replication Study *Notes: Replication study uses the same experimental interface and subject pool as the main study but focuses on the 16 information structures from the laboratory study. We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

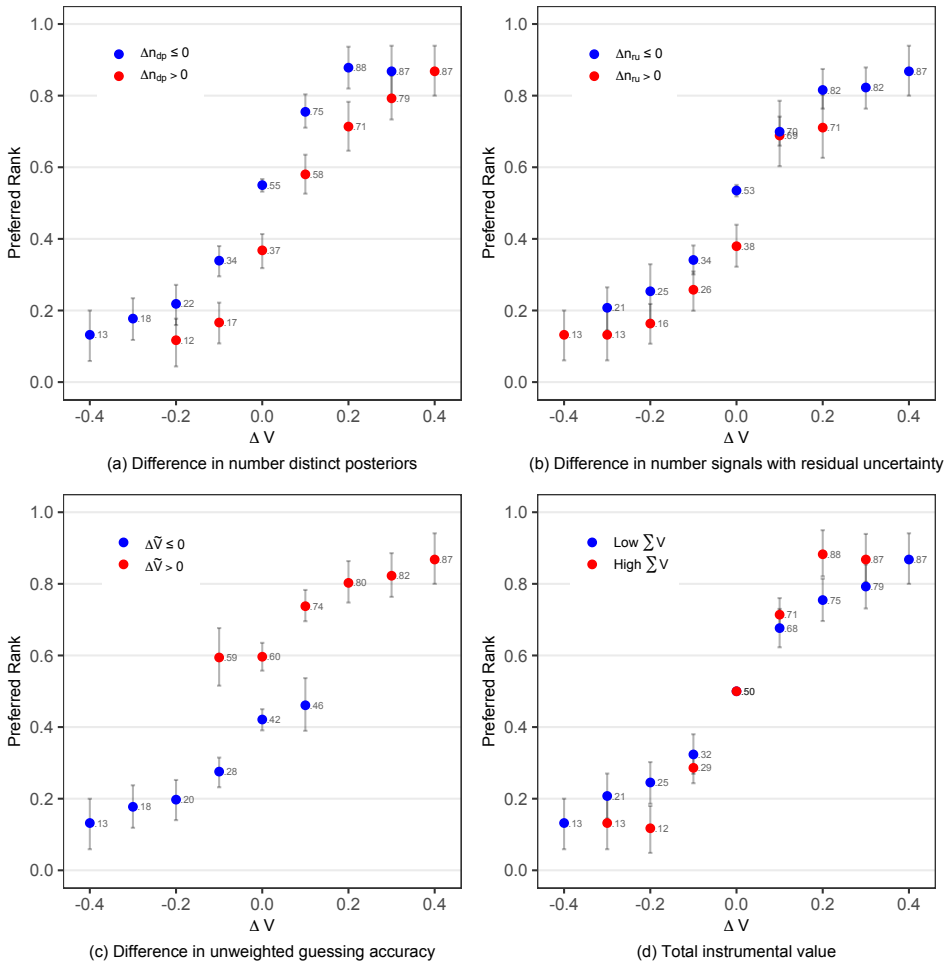


Figure 21: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study (Guess First) *Notes: Subset of participants in the laboratory study who made guesses first (before elicitation for demand for information). We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

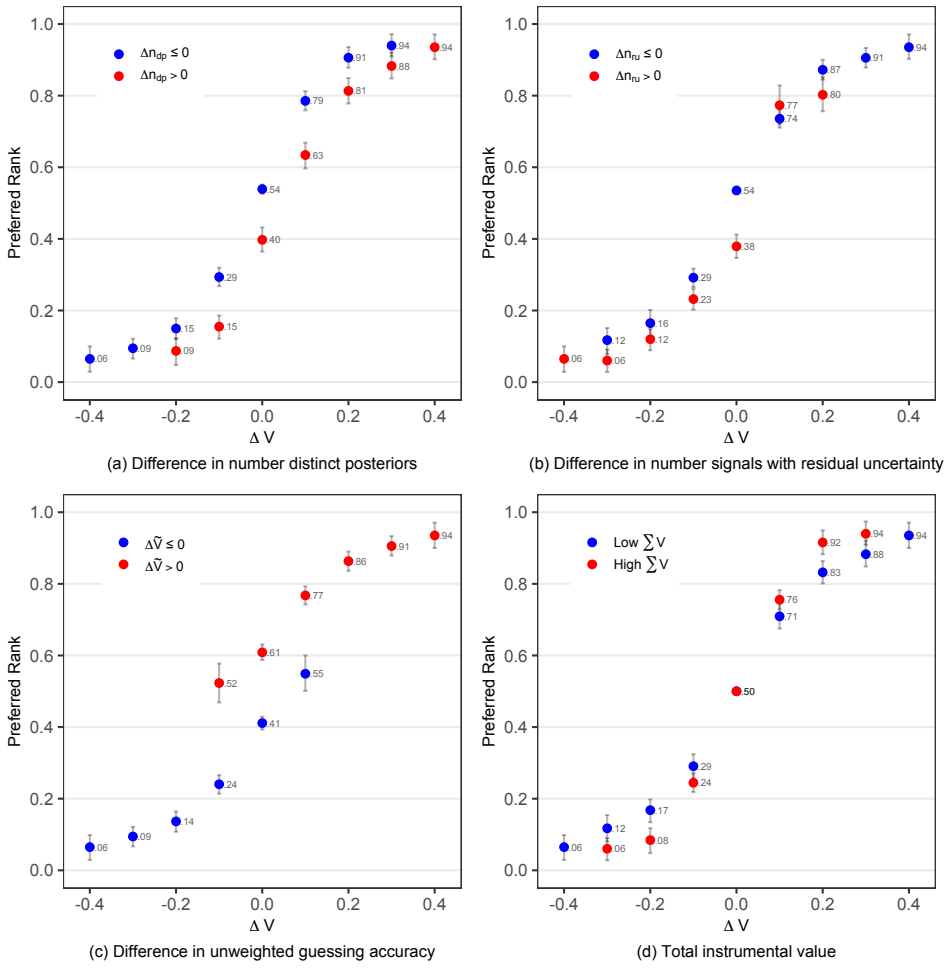


Figure 22: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study (Guess Second) *Notes: Subset of participants in the laboratory study who made guesses second (after elicitation for demand for information). We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

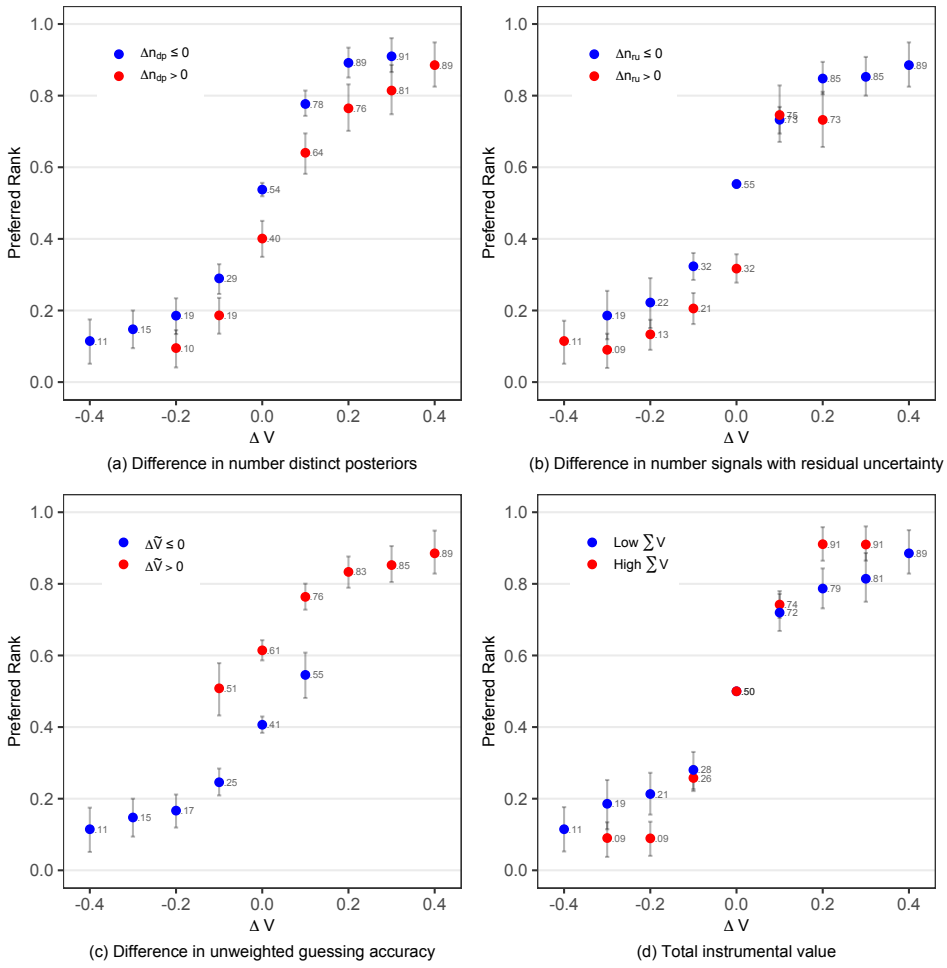


Figure 23: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study (No Uncertainty Treatment) *Notes: The No Uncertainty treatment of the laboratory study. We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

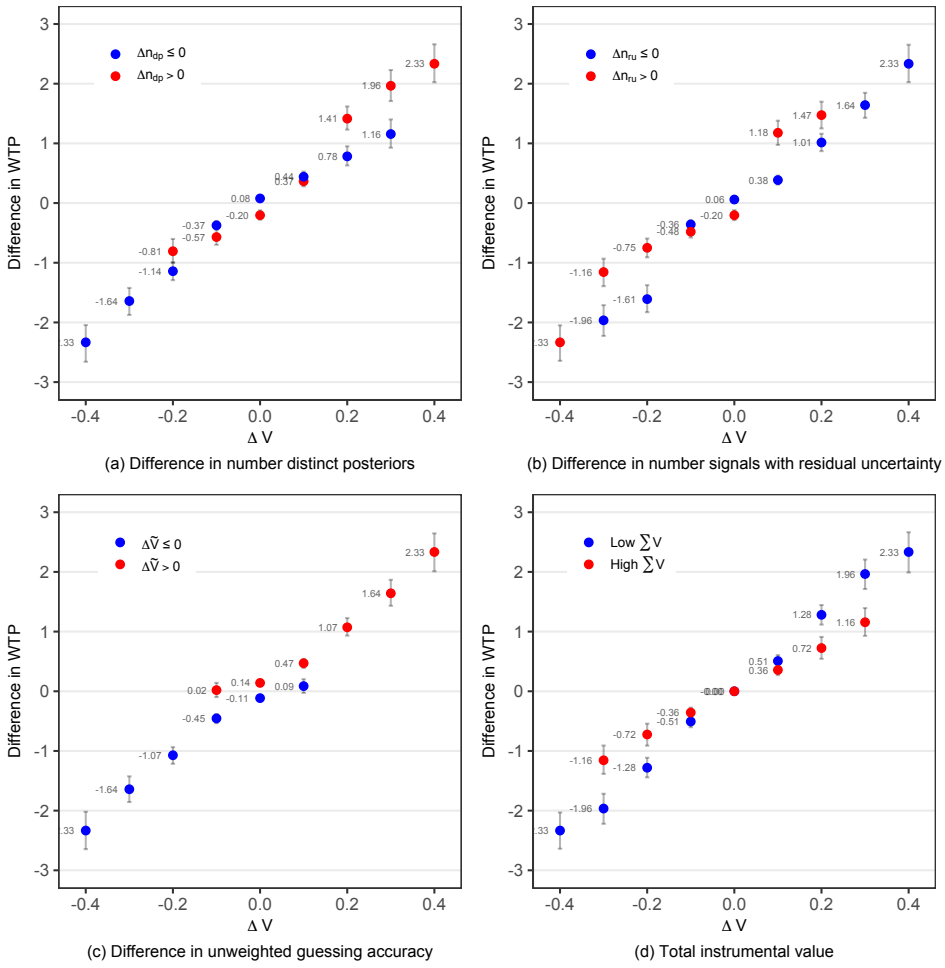


Figure 24: Difference in WTP for Pairwise Information Structures by Value Difference in the Laboratory Study *Notes: Both Baseline (Guess First) and Reverse (Guess Second). We focus on pairwise comparisons. Each figure plots the difference in WTP between a pair of information structures by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

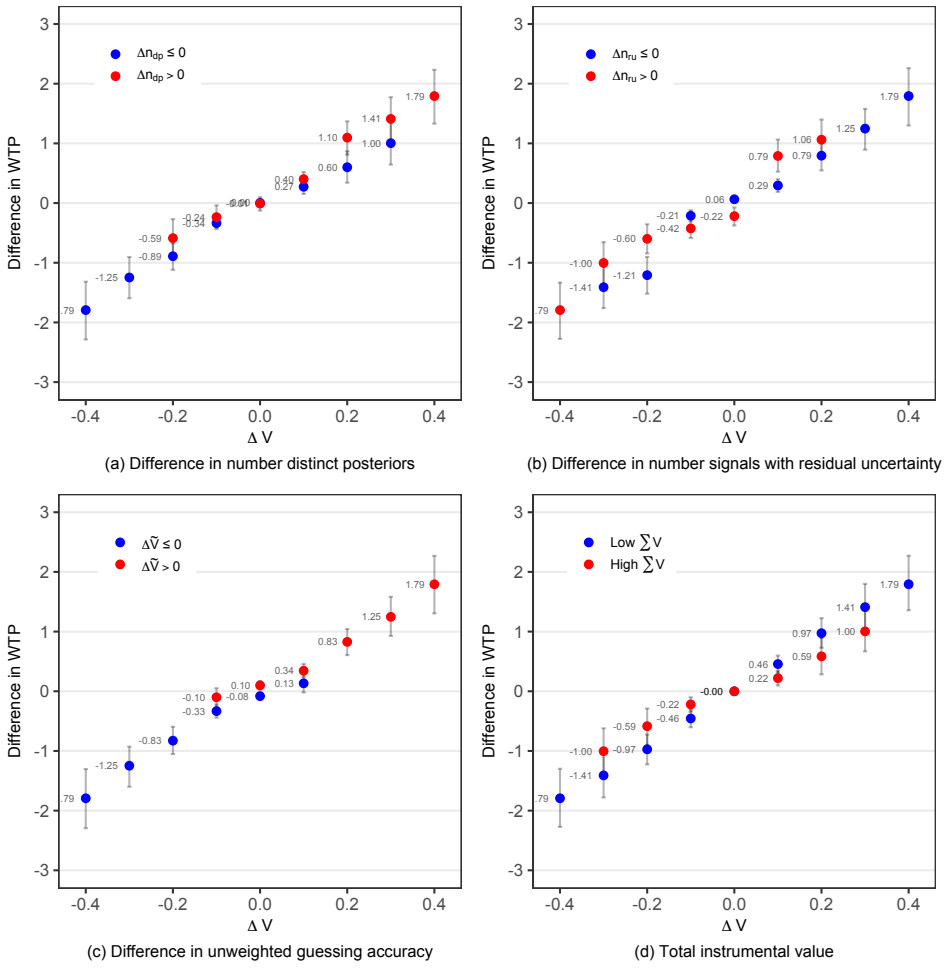
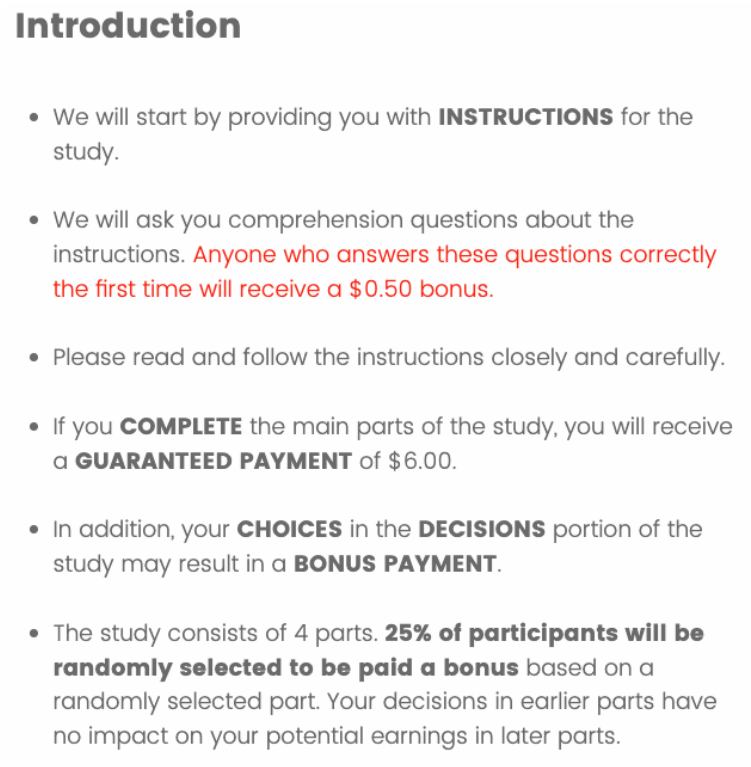


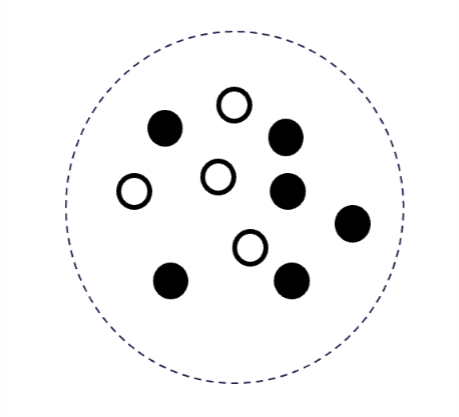
Figure 25: Difference in WTP for Pairwise Information Structures by Value Difference in the Laboratory Study (No Uncertainty Treatment) *Notes: The No Uncertainty treatment of the laboratory study. We focus on pairwise comparisons. Each figure plots the difference in WTP between a pair of information structures by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of distinct posteriors, number of signals with residual uncertainty, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

E Instructions and Screenshots

The following are screenshots of the experiment software.



Part 1: Guessing task



- In each guessing task, there are 10 balls. 6 of them are black, and 4 of them are white.
- One of these balls has been randomly selected by the computer.
- Your task is to guess the color of the ball (black or white) the computer selected.
- You earn a $10 **BONUS** if your guess is correct and $0 if your guess is incorrect.

Each of the following comprehension questions was shown on a new page. Subjects had to answer each question correctly to proceed.

Comprehension question 1: Which statement is INCORRECT?

- ○ 6 of the 10 balls are black.
- ○ 4 of the 10 balls are white.
- ○ The chance that the randomly selected ball is black is 60 percent.
- ○ The chance that the randomly selected ball is white is 40 percent.
- ○ The chance that the randomly selected ball is white is 60 percent.

Comprehension question 2: Which statement is CORRECT about when you win a bonus of $10?

- ○ If your guess about the color of the randomly selected ball is correct.
- ○ If your guess about the color of the randomly selected ball is incorrect.
- ○ If you guess white.
- ○ If you guess black.

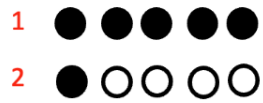
Information

Overview:

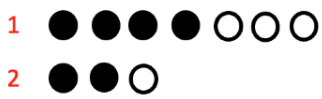
- Before you guess the color of the selected ball, you will receive some **INFORMATION** from an **INFORMATION SOURCE** about which ball the computer selected.
- Each information source divides the ten balls into different groups and informs you about which group the selected ball is in before you have to make a guess.

Description of an information source:

- An information source will divide the 10 balls into different groups like this, where **each row** is a different group and the group number is shown in red. Here is an example:



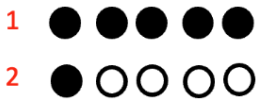
- and here is another:



(these are only examples -- we will show you many different information sources, with many different groupings across questions)

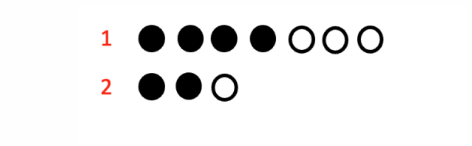
- We will tell you which Group number (shown as red numbers) the randomly selected ball is in **BEFORE** you guess the color of the ball.

For instance, in the following example



- The 10 balls are divided into two groups (two rows, labeled 1 and 2).
- There are 5 balls (all black) in group 1 and 5 balls (1 black and 4 white) in group 2.
- Since the groups are of equal size, the randomly selected ball is equally likely to be in group 1 or 2.
- If the information source reveals that the randomly selected ball is in group 1, then you will know for sure that the ball is black.
- If the information source reveals that the randomly selected ball is in group 2, then you will know that there is a 1/5 chance that it is black and 4/5 chance that it is white.

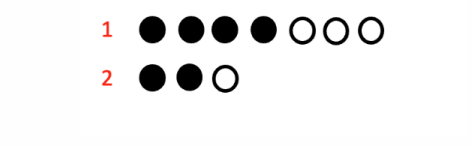
Comprehension question 1: Consider the following information source:



Which statement is CORRECT?

- ○ There is a 70 percent chance that the randomly selected ball is in group 1, and 30 percent chance that it is in group 2.
- ○ There is a 30 percent chance that the randomly selected ball is in group 1, and 70 percent chance that it is in group 2.
- ○ There is a 60 percent chance that the randomly selected ball is in group 1, and 40 percent chance that it is in group 2.
- ○ There is a 40 percent chance that the randomly selected ball is in group 1, and 60 percent chance that it is in group 2.
- ○ There is a 50 percent chance that the randomly selected ball is in group 1, and 50 percent chance that it is in group 2.

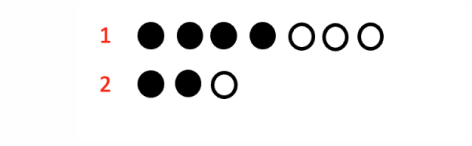
Comprehension question 2: Consider again the following information source:



If the information source reveals that the randomly selected ball is in group 1, what is the chance that is it black?

- ○ 4/7
- ○ 3/7
- ○ 1/3
- ○ 2/3
- ○ 1/2

Comprehension question 3: Consider again the following information source:

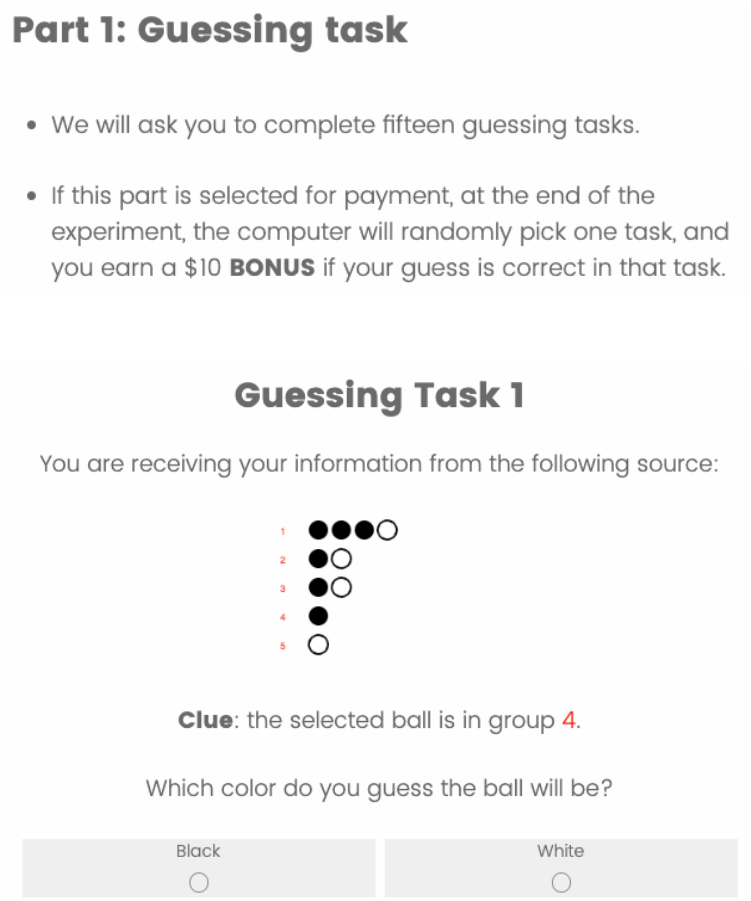


If the information source reveals that the randomly selected ball is in group 2, what is the chance that is it black?

- ○ 4/7
- ○ 3/7
- ○ 1/3
- ○ 2/3
- ○ 1/2

Comprehension question 4: In general, how do you maximize your chances of winning the $10 BONUS after receiving information from an information source?

- ○ Guess the color of the ball to be white.
- ○ Guess the color of the ball to be black.
- ○ Flip a coin to determine whether your guess is white or black.
- ○ Make your best guess about the color of the ball given the information provided to you on which group the selected ball belongs to by the information source.



This is an example of the guessing task. Subjects completed fifteen guessing tasks and then proceeded to part 2.

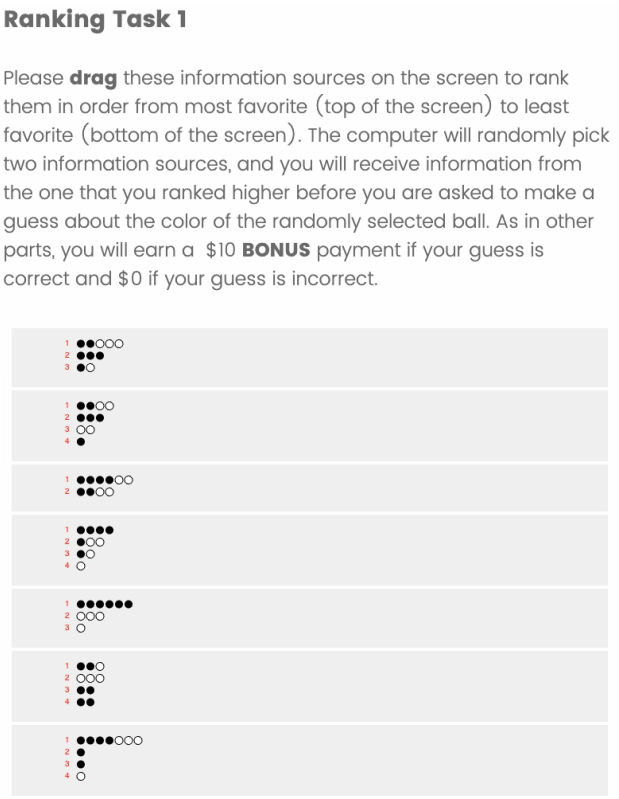
Part 2: Ranking information sources

Overview:

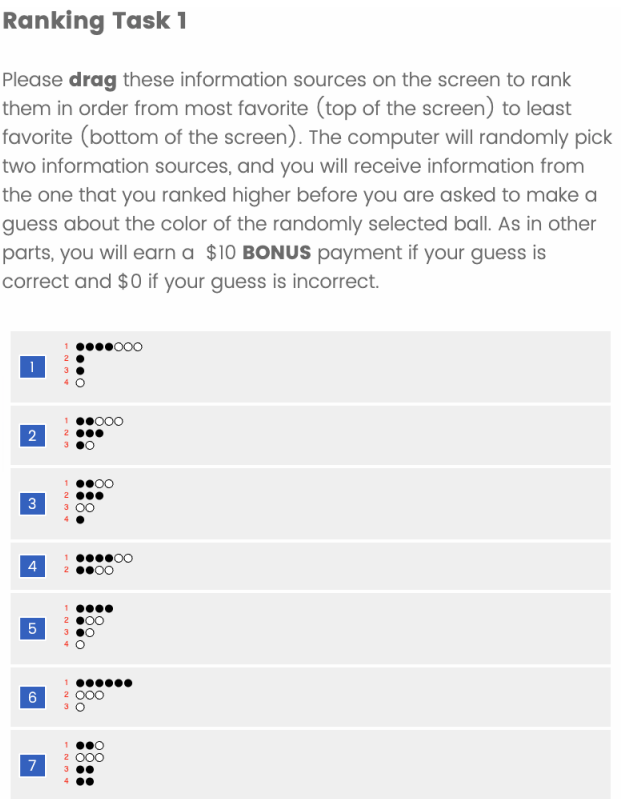
- You will tell us which information sources you would rather receive information from for a future guessing task by **RANKING** different information sources shown to you on your screen.
- Each information source divides the 10 balls (6 black and 4 white) into different groups as shown to you earlier.
- Your ranking (from highest to lowest) reveals which information sources you would rather receive information from before making a guess about the color of the selected ball.
- On each screen we will show you 7 different information sources in random order.
- You will use your mouse to click and drag these sources from most preferred (at the top) to least preferred (at the bottom).
- This part contains eight ranking tasks. If this part is selected for payment, at the end of the experiment, the computer will randomly select two of the information sources from one of the screens, and you will receive information from the one that you ranked higher before you are asked to guess the color of the randomly selected ball.
- Recall that you will earn a $10 **BONUS** payment if your guess is correct and $0 if your guess is incorrect, so your ranking choices could have a big impact on your earnings!

Comprehension question 1: Which statement is CORRECT?

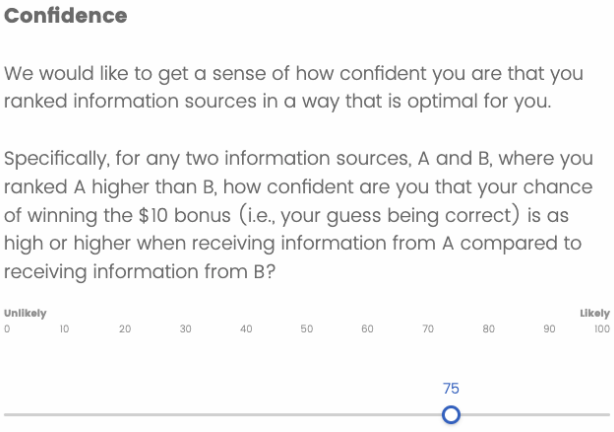
- ○ I am more likely to receive information from the information source I rank higher (above) than the one I rank lower (below).
- ○ I am more likely to receive information from the information source I rank lower (below) than the one I rank higher (above).
- ○ I am equally likely to receive information from any of the information sources.

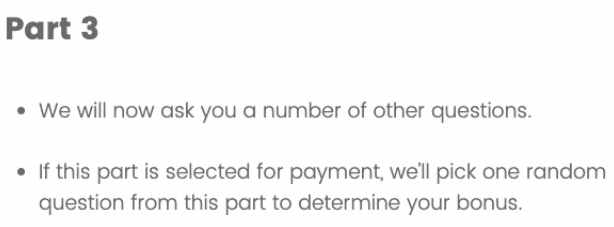


As subjects drag and reorder structures, an automatically updating order number appears to the left of each structure.

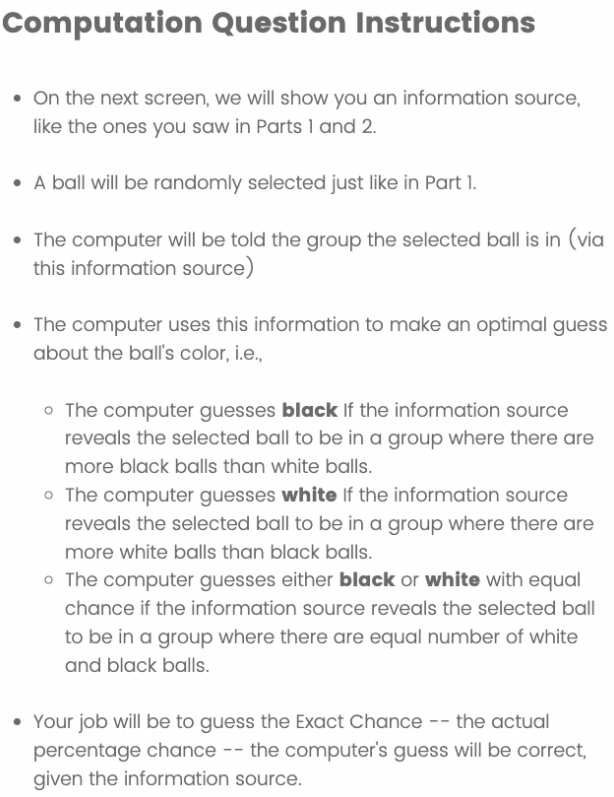


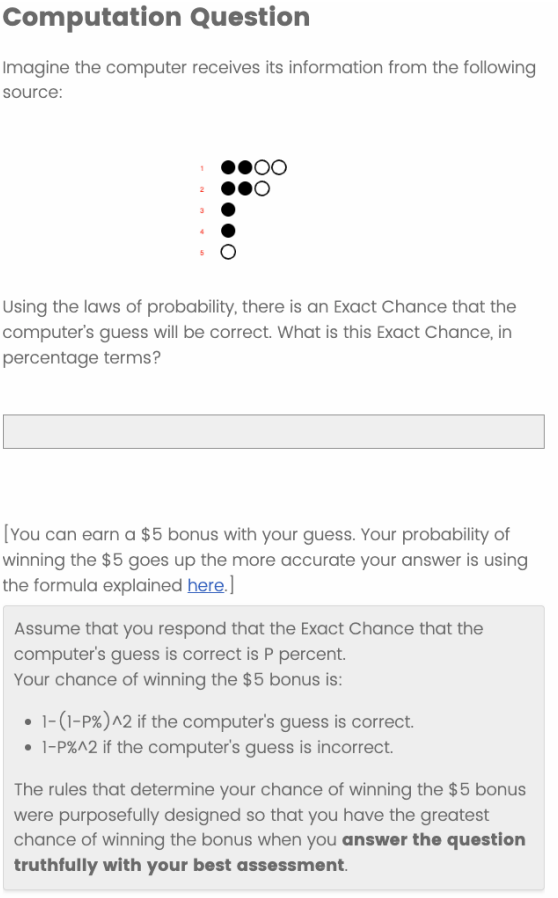
Subjects completed eight ranking tasks and then answered the following confidence question.





Part 3 contains several tasks. Subjects first completed the information structure computation task described below.

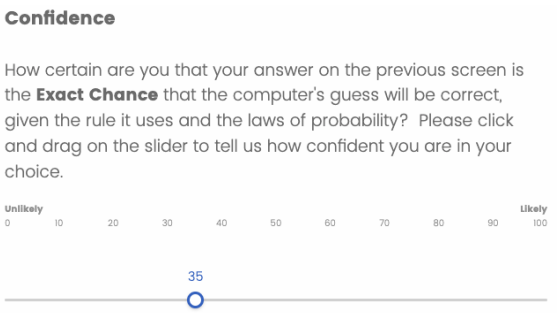




The detailed payoff calculation rule was initially hidden but would show up after subjects clicked the hypertext "here".

This shows one of the two pre-selected information structures used in this question. Another information structure is a partition having two subsets (groups): group 1 contains six black and one white balls; group 2 contains only a white ball.

After answering the above computation question, subjects reported their confidence on the next page.



Subjects then completed three tasks that elicited their certainty equivalents for three lotteries with equal expected value – a simple 50/50 lottery, a compound lottery, and an Ellsberg lottery. Below are the instructions.

Choice Task Instructions

On the next pages, you will be asked to choose between a **fixed amount of money received for sure** or **winning $30 if some event happens**.

For example, suppose there are a number of cards and each card is either Purple or Green. The computer shuffles the cards and then randomly draws one of them. You will be asked to make a choice between two options like below:

$30 if drawn card is Purple	○ ○	**$15** for sure

- If you select the **left option**, you win $30 if the randomly drawn card is Purple and $0 otherwise.
- If you select the **right option**, you win $15 for sure, regardless of the color of the randomly drawn card.

In fact, you will answer a list of questions like this. For example:

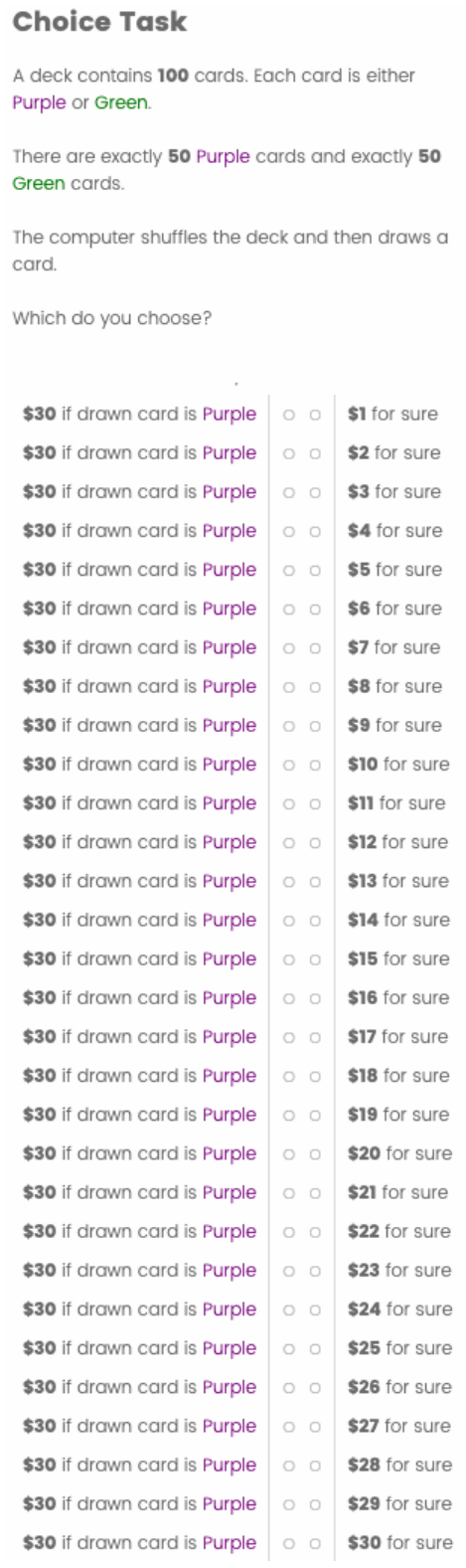
$30 if drawn card is Purple	○ ○	**$10** for sure
$30 if drawn card is Purple	○ ○	**$11** for sure
$30 if drawn card is Purple	○ ○	**$12** for sure
$30 if drawn card is Purple	○ ○	**$13** for sure
$30 if drawn card is Purple	○ ○	**$14** for sure
$30 if drawn card is Purple	○ ○	**$15** for sure
$30 if drawn card is Purple	○ ○	**$16** for sure
$30 if drawn card is Purple	○ ○	**$17** for sure
$30 if drawn card is Purple	○ ○	**$18** for sure
$30 if drawn card is Purple	○ ○	**$19** for sure
$30 if drawn card is Purple	○ ○	**$20** for sure

The option on **the left does not change**, while the option on **the right gets better as you go down the list**.

You must make a choice in all rows, but for simplicity, you only have to **click once on the row where you want to switch from left to right**.

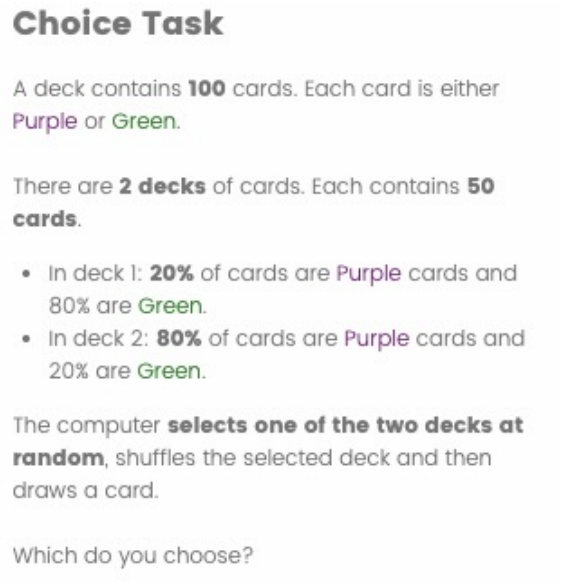
Intuitively, you can think about how much you would pay for a bet that would win you $30 if the randomly drawn card is Purple. Then, you switch to the sure amount of money as soon as it is above the amount you would pay for the bet.

At the end of the experiment, if this question is selected to determine your bonus payment, the computer will **randomly pick one row from the list** and pay you based on your choice in that row. This means you should **carefully consider your choice in each row** as any row could determine your bonus payment.

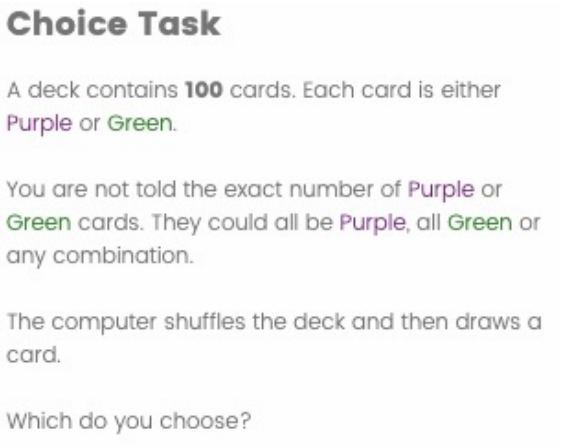


This is the elicitation of certainty equivalent for the simple 50/50 lottery. The other two elicitation tasks used the same multiple price lists. The order of the three elicitation tasks was randomized at the subject level.

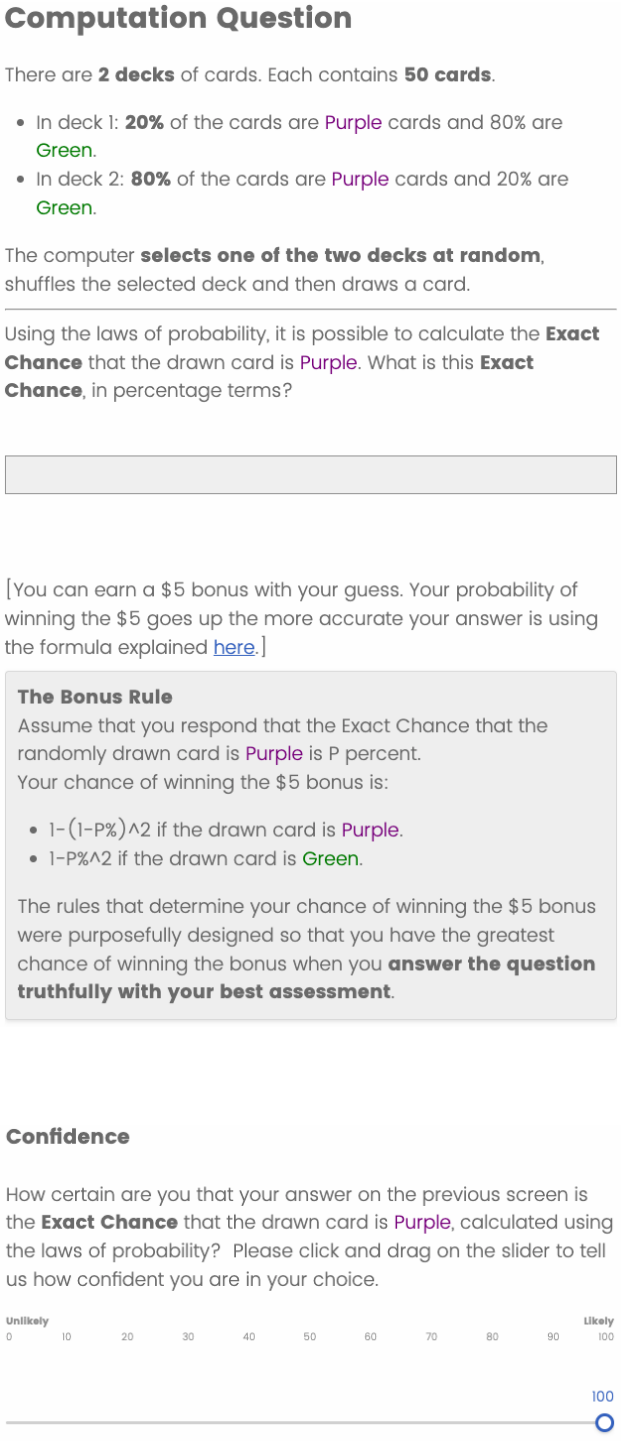
Compound lottery



Ellsberg lottery



Subjects then completed the following compound lottery computation question and reported their confidence on the next page.



Part 4 contains ten cognitive questions, each displayed on a different page in the experiment.

Part 4

In this Part we are going to ask you a set of 10 questions. If this part is selected for payment, you will earn 50 cents for each correct answer.

Question 1

Below are four cards. Each has a symbol on one side and a number on the other side. You are given the following statement which may or may not be true of all four cards: "If a card has a star on one side, then it has a number larger than 10 on the other side."

Your task is to name those cards, and only those cards that need to be turned over in order to determine whether the statement is true or false.

- ☐ Star card
- ☐ Sun card
- ☐ 7 card
- ☐ 13 card

Question 2

A farmer had 15 sheep and all but 8 died. How many are left?

- ○ none
- ○ 7
- ○ 8
- ○ 10

Question 3

If you're running a race and you pass the person in second place, what place are you in?

- ○ first
- ○ second
- ○ third
- ○ fourth

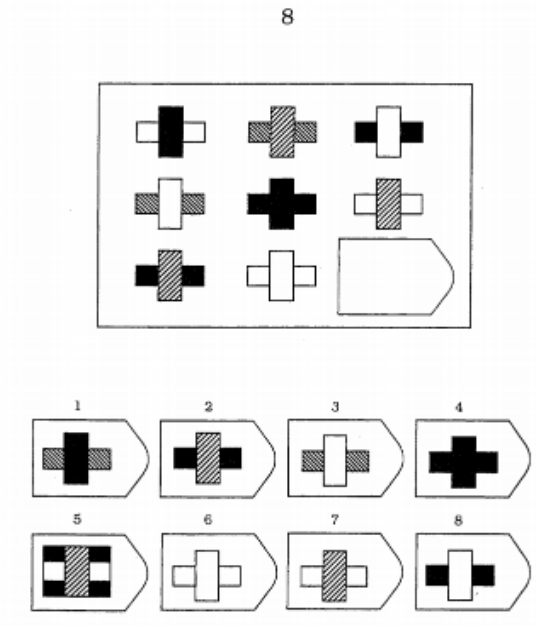
Question 4

Emily's father has three daughters. The first two are named April and May. What is the third daughter's name?

- ○ Can't know from information given
- ○ June
- ○ May
- ○ Emily

In each of the next three questions (5-7), there is a pattern with a piece missing and a number of pieces below the pattern. You have to choose which of the pieces below is the right one to complete the pattern.You will see 8 pieces that might complete the pattern. In every case, one and only one of these pieces is the right one to complete the pattern.

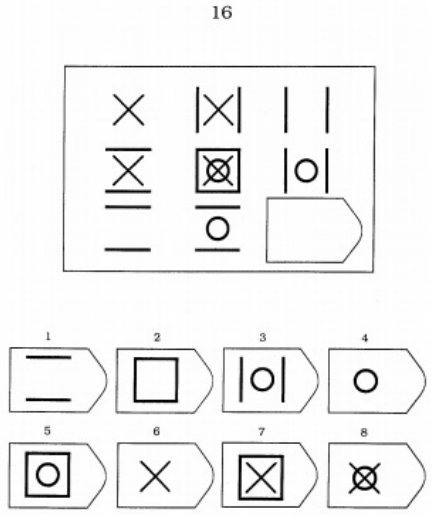
Question 5



Which piece completes the pattern?

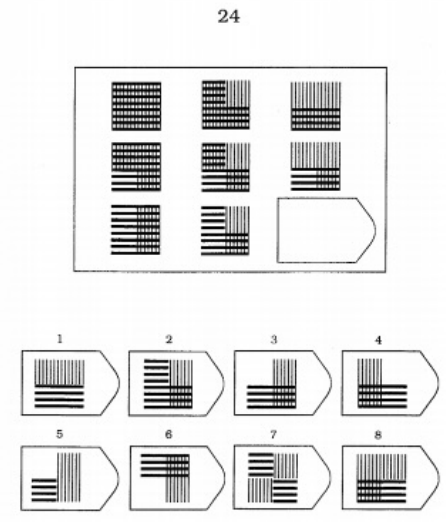
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8

Question 6



Which piece completes the pattern?

Question 7



Which piece completes the pattern?

For each of the following three problems (8-10), decide if the given conclusion follows logically from the premises. Select YES if, and only if, you judge that the conclusion can be derived unequivocally from the given premises. Otherwise select NO.

Question 8

Premise 1: All flowers have petals.
Premise 2: Roses have petals.
Conclusion: Roses are flowers.

Does the conclusion follow from the premises?

- ○ YES
- ○ NO

Question 9

Premise 1: All mammals walk.
Premise 2: Whales are mammals.
Conclusion: Whales walk.

Does the conclusion follow from the premises?

- ○ YES
- ○ NO

Question 10

Premise 1: All animals like water.
Premise 2: Cats do not like water.
Conclusion: Cats are not animals.

Does the conclusion follow from the premises?

- ○ YES
- ○ NO

Guessing Question 1

There are 6 blue balls, and 4 red balls. One of these balls will be randomly selected and you earn $10 if you correctly guess the color (blue or red):

You will learn which of the following Groups the ball is in before you guess:

Please tell us what color you will guess if you learn the ball is in each of these possible Groups:

If I learn the ball is in **Group 1**, I will guess the color to be:	○ **Blue** ○ **Red**
If I learn the ball is in **Group 2**, I will guess the color to be:	○ **Blue** ○ **Red**
If I learn the ball is in **Group 3**, I will guess the color to be:	○ **Blue** ○ **Red**

(Remember,these choices determine your actual guess and therefore your payment!)

Figure 26: Guessing Task in Laboratory Study

(a) Ranking (b) WTP

Figure 27: Ranking and WTP Elicitation Tasks in Laboratory Study

Ranking Information Sources

Please drag these information sources on the screen to rank them in order from most favorite (top of the screen) to least favorite (bottom of the screen). The computer will randomly pick two information sources, and you will receive information from the one that you ranked higher before you are asked to make a guess about the color of the randomly selected ball. As in other parts, you will earn a $10 BONUS payment if your guess is correct and $0 if your guess is incorrect.

(To help with your decision, for each Information Source, we include below each Group the guess you made for that Group earlier in a corresponding Guessing Question.)

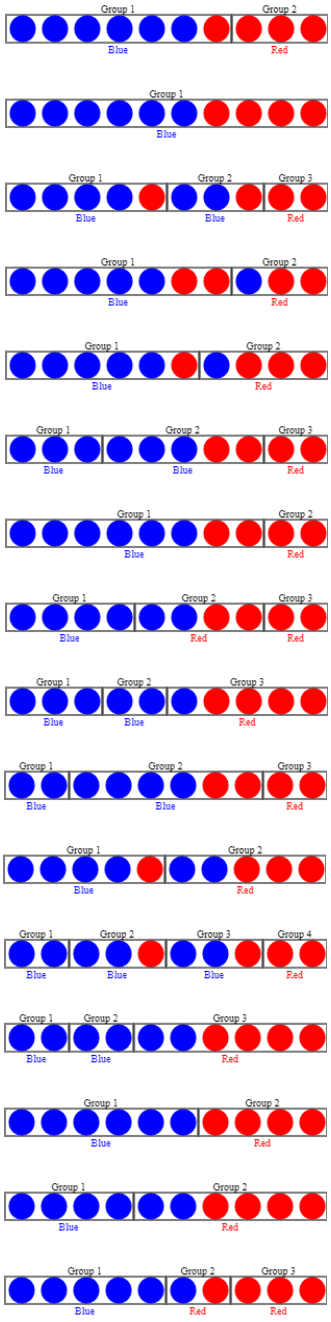


Figure 28: Ranking Task from Reverse Treatment (Guess First) in Laboratory Study